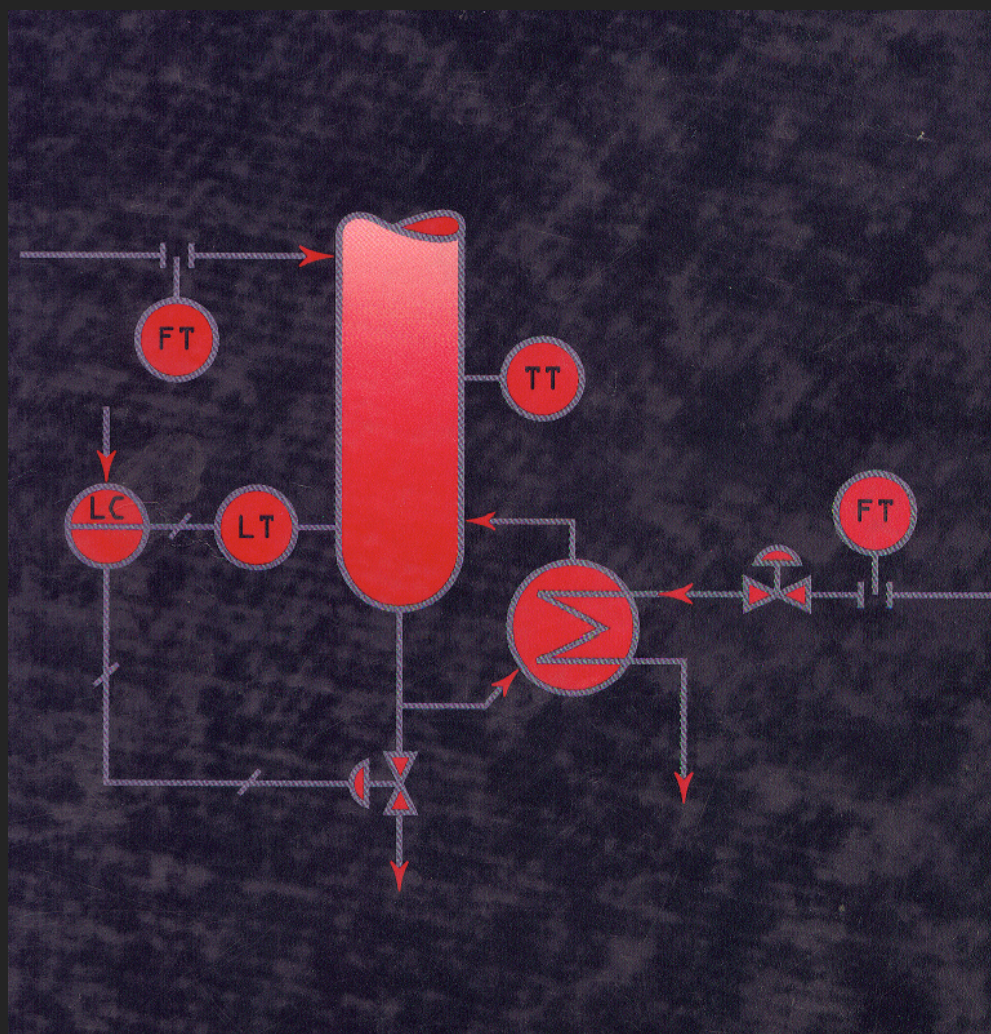# Principles and Practice of Automatic Process Control

Second Edition

Carl& A. Smith

Armando B. Corripio

SELECTED TABLES AND FIGURES

Principles and Practice of Automatic Process Control

Second Edition

Carlos A. Smith, Ph.D., P.E.

University of South Florida

Armando B. Corripio, Ph.D., P.E.

Louisiana State University

This work is dedicated with all our love to The Lord our God,
for all his daily blessings made this book possible

The Smiths:
Cristina, Carlos A. Jr., Tim, Cristina M., and Sophia C. Livingston,
and Mrs. Rene M. Smith,

my four grandsons:
Nicholas, Robert, Garrett and David

and to our dearest homeland, Cuba

Preface

This edition is a major revision and expansion to the first edition. Several new subjects have been added, notably the z-transform analysis and discrete controllers, and several other subjects have been reorganized and expanded. The objective of the book, however, remains the same as in the first edition, "to present the practice of automatic process control along with the fundamental principles of control theory." A significant number of applications resulting from our practice as part-time consultants have also been added to this edition.

Twelve years have passed since the first edition was published, and even though the principles are still very much the same, the "tools" to implement the controls strategies have certainly advanced. The use of computer-based instrumentation and control systems is the norm.

Chapters 1 and 2 present the definitions of terms and mathematical tools used in process control. In this edition Chapter 2 stresses the determination of the quantitative characteristics of the dynamic response, settling time, frequency of oscillation, and damping ratio, and de-emphasizes the exact determination of the analytical response. In this way the students can analyze the response of a dynamic system without having to carry out the time-consuming evaluation of the coefficients in the partial fraction expansion. Typical responses of first-, second-, and higher-order systems are now presented in Chapter 2.

The derivation of process dynamic models from basic principles is the subject of Chapters 3 and 4. As compared to the first edition, the discussion of process modelling has been expanded. The discussion, meaning, and significance of process nonlinearities has been expanded as well. Several numerical examples are presented to aid in the understanding of this important process characteristic. Chapter 4 concludes with a presentation of integrating, inverse-response, and open-loop unstable processes.

Chapter 5 presents the design and characteristics of the basic components of a control system: sensors and transmitters, control valves, and feedback controllers. The presentation of control valves and feedback controllers has been expanded. Chapter 5 should be studied together with Appendix C where practical operating principles of some common sensors, transmitters, and control valves are presented.

The design and tuning of feedback controllers are the subjects of Chapters 6 and 7. Chapter 6 presents the analysis of the stability of feedback control loops. In this edition we stress the direct substitution method for determining both the ultimate gain and period of the loop. Routh's test is deemphasized, but still presented in a separate section. In keeping with the spirit of Chapter 2, the examples and problems deal with the determination of the characteristics of the response of the closed loop, not with the exact analytical response of the loop. Chapter 7 keeps the same tried-and-true tuning methods from the first edition. A new section on tuning controllers for integrating processes, and a discussion of the Internal Model Control (IMC) tuning rules, have been added.

Chapter 8 presents the root locus technique, and Chapter 9 presents the frequency response techniques. These techniques are principally used to study the stability of control systems.

The additional control techniques that supplement and enhance feedback control have been distributed among Chapters 10 through 13 to facilitate the selection of their coverage in university courses. Cascade control is presented first, in Chapter 10, because it is so commonly a part of the other schemes. Several examples are presented to help understanding of this important and common control technique.

Chapter 11 presents different computing algorithms sometimes used to implement control schemes. A method to scale these algorithms, when necessary, is presented. The chapter also presents the techniques of override, or constraint, control, and selective control. Examples are used to explain the meaning and justification of them.

Chapter 12 presents and discusses in detail the techniques of ratio and feedforward control. Industrial examples are also presented. A significant number of new problems have been added.

Multivariable control and loop interaction are the subjects of Chapter 13. The calculation and interpretation of the relative gain matrix (RGM) and the design of decouplers, are kept from the first edition. Several examples have been added, and the material has been reorganized to keep all the dynamic topics in one section.

Finally Chapters 14 and 15 present the tools for the design and analysis of sampled-data (computer) control systems. Chapter 14 presents the z-transform and its use to analyze sampled-data control systems, while Chapter 15 presents the design of basic algorithms for computer control and the tuning of sampled-data feedback controllers. The chapter includes sections on the design and tuning of dead-time compensation algorithms and model-reference control algorithms. Two examples of Dynamic Matrix Control (DMC) are also included.

As in the first edition, Appendix A presents some symbols, labels, and other notations commonly used in instrumentation and control diagrams. We have adopted throughout the book the ISA symbols for conceptual diagrams which eliminate the need to differentiate between pneumatic, electronic, or computer implementation of the various control schemes. In keeping with this spirit, we express all instrument signals in percent of range rather than in mA or psig. Appendix B presents several processes to provide the student/reader an opportunity to design control systems from scratch.

During this edition we have been very fortunate to have received the help and encouragement of several wonderful individuals. The encouragement of our students, especially Daniel Palomares, Denise Farmer, Carl Thomas, Gene Daniel, Samuel Peebles, Dan Logue, and Steve Hunter, will never be forgotten. Thanks are also due to Dr. Russell Rhinehart of Texas Tech University who read several chapters when they were in the initial stages. His comments were very helpful and resulted in a better book. Professors Ray Wagonner, of Missouri Rolla, and G. David Shilling, of Rhode Island, gave us invaluable suggestions on how to improve the first edition. To both of them we are grateful. We are also grateful to Michael R. Benning of Exxon Chemical Americas who volunteered to review the manuscript and offered many useful suggestions from his industrial background.

In the preface to the first edition we said that "To serve as agents in the training and development of young minds is certainly a most rewarding profession." This is still our conviction and we feel blessed to be able to do so. It is with this desire that we have written this edition.

CARLOS A. SMITH
Tampa, Florida, 1997

ARMANDO B. CORRIPIO
Baton Rouge, Louisiana, 1997

Contents

Chapter 13 Multivariable Process Control — 545

Chapter 14   Mathematical Tools for Computer Control Systems — 599

Chapter 1

Introduction

The purpose of this chapter is to present the need for automatic process control and to motivate you, the reader, to study it. Automatic process control is concerned with maintaining process variables, temperatures, pressures, flows, compositions, and the like at some desired operating value. As we shall see, processes are dynamic in nature. Changes are always occurring, and if appropriate actions are not taken in response, then the important process variables-those related to safety, product quality, and production rates-will not achieve design conditions.

This chapter also introduces two control systems, takes a look at some of their components, and defines some terms used in the field of process control. Finally, the background needed for the study of process control is discussed.

In writing this book, we have been constantly aware that to be successful, the engineer must be able to apply the principles learned. Consequently, the book covers the principles that underlie the successful practice of automatic process control. The book is full of actual cases drawn from our years of industrial experience as full-time practitioners or part-time consultants. We sincerely hope that you get excited about studying automatic process control. It is a very dynamic, challenging, and rewarding area of process engineering.

1-1 A PROCESS CONTROL SYSTEM

To illustrate process control, let us consider a heat exchanger in which a process stream is heated by condensing steam; the process is sketched in Fig. 1-1.1. The purpose of this unit is to heat the process fluid from some inlet temperature $T_i(t)$ up to a certain desired outlet temperature T(t). The energy gained by the process fluid is provided by the latent heat of condensation of the steam.

In this process there are many variables that can change, causing the outlet temperature to deviate from its desired value. If this happens, then some action must be taken to correct the deviation. The objective is to maintain the outlet process temperature at its desired value.

One way to accomplish this objective is by measuring the temperature T(t), comparing it to the desired value, and, on the basis of this comparison, deciding what to do to correct any deviation. The steam valve can be manipulated to correct the deviation. That is, if the temperature is above its desired value, then the steam valve can be
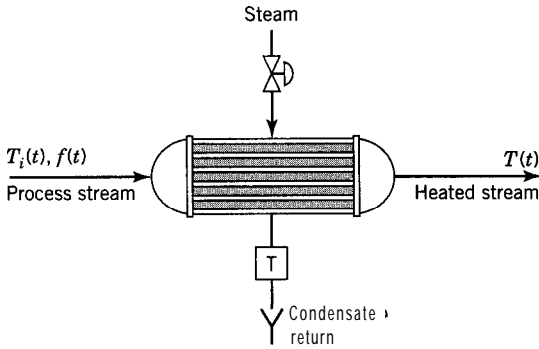
Figure 1-1.1 Heat exchanger.

throttled back to cut the steam flow (energy) to the heat exchanger. If the temperature is below the desired value, then the steam valve can be opened more to increase the steam flow to the exchanger. All of this can be done manually by the operator, and the procedure is fairly straightforward. However, there are several problems with such *manual control.* First, the job requires that the operator look at the temperature frequently to take corrective action whenever it deviates from the desired value. Second, different operators make different decisions about how to move the steam valve, and this results in a less than perfectly consistent operation. Third, because in most process plants there are hundreds of variables that must be maintained at some desired value, manual correction requires a large number of operators. As a result of these problems, we would like to accomplish this control automatically. That is, we would like to have systems that control the variables without requiring intervention from the operator. This is what is meant by *automatic process control.*

To achieve automatic process control, a *control system* must be designed and implemented. A possible control system for our heat exchanger is shown in Fig. 1-1.2. (Ap-
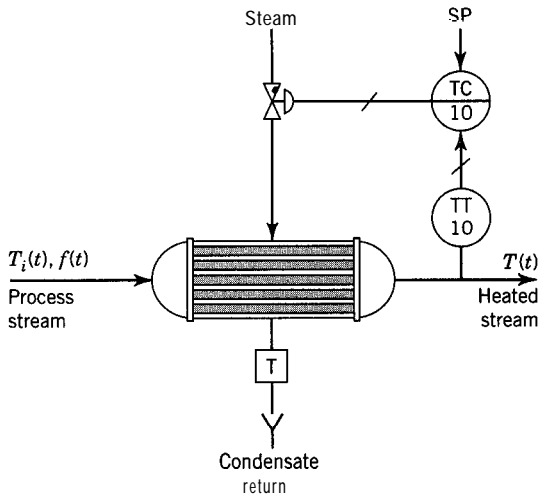
Figure l-l.2 Heat exchanger control system.

pendix A presents the symbols and identifications for different devices.) The first thing to do is measure the outlet temperature of the process stream. This is done by a **sensor** (thermocouple, resistance temperature device, filled system thermometer, thermistor, or the like). Usually this sensor is physically connected to a **transmitter,** which takes the output from the sensor and converts it to a signal strong enough to be transmitted to a **controller.** The controller then receives the signal, which is related to the temperature, and compares it with the desired value. Depending on the result of this comparison, the controller decides what to do to maintain the temperature at the desired value. On the basis of this decision, the controller sends a signal to **the final control element,** which in turn manipulates the steam flow. This type of control strategy is known as **feedback control.**

Thus the three basic components of all control systems are

1. **Sensor/transmitter** Also often called the primary and secondary elements.
2. **Controller** The "brain" of the control system.
3. **Final control element** Often a control valve but not always. Other common final control elements are variable-speed pumps, conveyors, and electric motors.

These components perform the three basic operations that **must** be present in every control system. These operations are

1. **Measurement(M)** Measuring the variable to be controlled is usually done by the combination of sensor and transmitter. In some systems, the signal from the sensor can be fed directly to the controller, so there is no need for the transmitter.
2. **Decision (D)** On the basis of the measurement, the controller decides what to do to maintain the variable at its desired value.
3. **Action (A)** As a result of the controller's decision, the system must then take an action. This is usually accomplished by the final control element.

These three operations, M, D, and A, are always present in every type of control system, and it is imperative that they be in a loop. That is, on the basis of the measurement a decision is made, and on the basis of this decision an action is taken. **The action taken must come back and affect the measurement; otherwise, it is a major flaw in the design, and control will not be achieved.** When the action taken does not affect the measurement, an open-loop condition exists and control will not be achieved. The decision making in some systems is rather simple, whereas in others it is more complex; we will look at many systems in this book.

1-2 IMPORTANT TERMS AND THE OBJECTIVE OF AUTOMATIC PROCESS CONTROL

At this time it is necessary to define some terms used in the field of automatic process control. The **controlled variable** is the variable that must be maintained, or controlled, at some desired value. In our example of the heat exchanger, the process outlet temperature, $T(t)$, is the controlled variable. Sometimes the term **process variable** is also used to refer to the controlled variable. The **set point** (SP) is the desired value of the controlled variable. Thus the job of a control system is to maintain the controlled variable at its set point. **The manipulated variable** is the variable used to maintain the controlled variable at its set point. In the example, the steam valve position is the

manipulated variable. Finally, any variable that causes the controlled variable to deviate from the set point is known as a *disturbance* or *upset*. In most processes there are a number of different disturbances. In the heat exchanger shown in Fig. 1-1.2, possible disturbances include the inlet process temperature, $T_i(t)$, the process flow, $f(t)$, the energy content of the steam, ambient conditions, process fluid composition, and fouling. It is important to understand that disturbances are always occurring in processes. Steady state is not the rule, and transient conditions are very common. It is because of these disturbances that automatic process control is needed. If there were no disturbances, then design operating conditions would prevail and there would be no need to "monitor" the process continuously.

The following additional terms are also important. *Manual control* is the condition in which the controller is disconnected from the process. That is, the controller is not deciding how to maintain the controlled variable at set point. It is up to the operator to manipulate the signal to the final control element to maintain the controlled variable at set point. *Closed-loop control* is the condition in which the controller is connected to the process, comparing the set point to the controlled variable and determining and taking corrective action.

Now that we have defined these terms, we can express the objective of an automatic process control system meaningfully: *The objective of an automatic process control system is to adjust the manipulated variable to maintain the controlled variable at its set point in spite of disturbances.*

Control is important for many reasons. Those that follow are not the only ones, but we feel they are the most important. They are based on our industrial experience, and we would like to pass them on. Control is important to

1. Prevent injury to plant personnel, protect the environment by preventing emissions and minimizing waste, and prevent damage to the process equipment. *SAFETY* must always be in everyone's mind; it is the single most important consideration.
2. Maintain product quality (composition, purity, color, and the like) on a continuous basis and with minimum cost.
3. Maintain plant production rate at minimum cost.

Thus process plants are automated to provide a safe environment and at the same time maintain desired product quality, high plant throughput, and reduced demand on human labor.

1-3 REGULATORY AND SERVO CONTROL

In some processes, the controlled variable deviates from set point because of disturbances. Systems designed to compensate for these disturbances *exert regulatory control*. In some other instances, the most important disturbance is the set point itself. That is, the set point may be changed as a function of time (typical of this is a batch reactor where the temperature must follow a desired profile), and therefore the controlled variable must follow the set point. Systems designed for this purpose exert *servo control*.

Regulatory control is much more common than servo control in the process indus-

tries. However, the same basic approach is used in designing both. Thus the principles in this book apply to both cases.

1-4 TRANSMISSION SIGNALS, CONTROL SYSTEMS, AND OTHER TERMS

Three principal types of signals are used in the process industries. The *pneumatic signal,* or air pressure, normally ranges between 3 and 15 psig. The usual representation for pneumatic signals in process and instrumentation diagrams (P&IDs) is —#————#—. The *electrical signal* normally ranges between 4 and 20 mA. Less often, a range of 10 to 50 mA, 1 to 5 V, or 0 to 10 V is used. The usual representation for this signal in P&IDs is a series of dashed lines such as ▪▪▪▪-. The third type of signal is the *digital,* or *discrete, signal* (zeros and ones). In this book we will show such signals as —+————+— (see Fig. 1-1.2), which is the representation proposed by the Instrument Society of America (ISA) when a control concept is shown without concern for specific hardware. The reader is encouraged to review Appendix A, where different symbols and labels are presented. Most times we will refer to signals as percentages instead of using psig or mA. That is, 0%- 100% is equivalent to 3 to 15 psig or 4 to 20 mA.

It will help in understanding control systems to realize that signals are used by de-vices-transmitters, controllers, final control elements, and the like-to communicate. That is, signals are used to *convey information.* The signal from the transmitter to the controller is used by the transmitter to inform the controller of the value of the controlled variable. This signal is not the measurement in engineering units but rather is a mA, psig, volt, or any other signal that is proportional to the measurement. The relationship to the measurement depends on the calibration of the sensor/transmitter. The controller uses its output signal to tell the final control element what to do: how much to open if it is a valve, how fast to run if it is a variable-speed pump, and so on.

It is often necessary to change one type of signal into another. This is done by a *transducer,* or *converter.* For example, there may be a need to change from an electrical signal in milliamperes (mA) to a pneumatic signal in pounds per square inch, gauge (psig). This is done by the use of a current (I) to pneumatic (P) transducer (I/P); see Fig. 1-4.1. The input signal may be 4 to 20 mA and the output 3 to 15 psig. An analog-to-digital converter (A to D) changes from a mA, or a volt signal to a digital signal. There are many other types of transducers: digital-to-analog (D to A), pneumatic-to-current (P/I), voltage-to-pneumatic (E/P), pneumatic-to-voltage (P/E), and so on.

The term *analog* refers to a controller, or any other instrument, that is either pneu-matic or electrical. Most controllers, however, *are computer-based,* or *digital.* By com-puter-based we don't necessarily mean a main-frame computer but anything starting from a microprocessor. In fact, most controllers are microprocessor-based. Chapter 5 presents different types of controllers and defines some terms related to controllers and control systems.
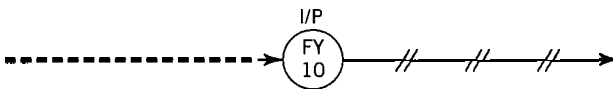
Figure   1-4.1   I/P   transducer.

1-5   CONTROL   STRATEGIES

1-5.1 Feedback   Control

The control scheme shown in Fig. l-l.2 is referred to as ***feedback control*** and is also called **afeedback control loop.** One must understand the working principles of feedback control to recognize its advantages and disadvantages; the heat exchanger control loop shown in Fig. l-l.2 is presented to foster this understanding.

If the inlet process temperature increases, thus creating a disturbance, its effect must propagate through the heat exchanger before the outlet temperature increases. Once this temperature changes, the signal from the transmitter to the controller also changes. It is then that the controller becomes aware that a deviation from set point has occurred and that it must compensate for the disturbance by manipulating the steam valve. The controller signals the valve to close and thus to decrease the steam flow. Fig. 1-5.1 shows graphically the effect of the disturbance and the action of the controller.

It is instructive to note that the outlet temperature first increases, because of the increase in inlet temperature, but it then decreases even below set point and continues to oscillate around set point until the temperature finally stabilizes. This oscillatory response is typical of feedback control and shows that it is essentially a trial-and-error operation. That is, when the controller "notices" that the outlet temperature has increased above the set point, it signals the valve to close, but the closure is more than required. Therefore, the outlet temperature decreases below the set point. Noticing *this,*
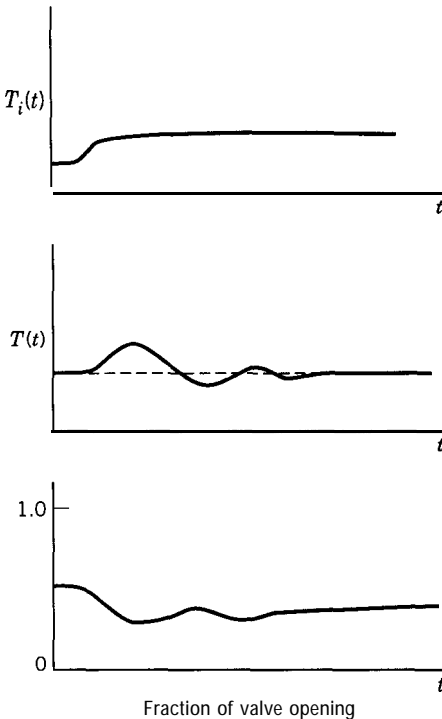
Figure 1-5.1 Response of a heat exchanger to a disturbance: feedback control.

the controller signals the valve to open again somewhat to bring the temperature back up. This trial-and-error operation continues until the temperature reaches and remains at set point.

The *advantage* of feedback control is that it is a very simple technique that compensates for all disturbances. Any disturbance affects the controlled variable, and once this variable deviates from set point, the controller changes its output in such a way as to return the temperature to set point. The feedback control loop does not know, nor does it care, which disturbance enters the process. It tries only to maintain the controlled variable at set point and in so doing compensates for all disturbances. The feedback controller works with minimum knowledge of the process. In fact, the only information it needs is in which direction to move. How much to move is usually adjusted by trial and error. *The disadvantage* of feedback control is that it can compensate for a disturbance only after the controlled variable has deviated from set point. That is, the disturbance must propagate through the entire process before the feedback control scheme can initiate action to compensate for it.

The job of the engineer is to design a control scheme that will maintain the controlled variable at its set point. Once this is done, the engineer must adjust, or tune, the controller so that it minimizes the amount of trial and error required. Most controllers have up to three terms (also known as parameters) used to tune them. To do a creditable job, the engineer must first know the characteristics of the process to be controlled. Once these characteristics are known, the control system can be designed and the controller tuned. Process characteristics are explained in Chapters 3 and 4, Chapter 5 presents the meaning of the three terms in the controllers, and Chapter 7 explains how to tune them.

14.2 Feedforward Control

Feedback control is the most common control strategy in the process industries. Its simplicity accounts for its popularity. In some processes, however, feedback control may not provide the required control performance. For these processes, other types of control strategies may have to be designed. Chapters 10, 11, 12, 13, and 15 present additional control strategies that have proved profitable. One such strategy is feedforward control. The objective of feedforward control is to measure disturbances and compensate for them before the controlled variable deviates from set point. When feedforward control is applied correctly, deviation of the controlled variable is minimized.

A concrete example of feedforward control is the heat exchanger shown in Fig. 1-1.2. Suppose that "major" disturbances are the inlet temperature, $T_i(t)$, and the process flow, $f(t)$. To implement feedforward control, these two disturbances must first be measured, and then a decision must be made about how to manipulate the steam valve to compensate for them. Fig. 1-5.2 shows this control strategy. The feedforward controller makes the decision about how to manipulate the steam valve to maintain the controlled variable at set point, depending on the inlet temperature and process flow.

In Section 1-2 we learned that there are a number of different disturbances. The feedforward control system shown in Fig. 1-5.2 compensates for only two of them. If any of the others enter the process, this strategy will not compensate for it, and the result will be a permanent deviation of the controlled variable from set point. To avoid this deviation, some feedback compensation must be added to feedforward control; this is shown in Fig. 1-5.3. Feedforward control now compensates for the "major" distur-
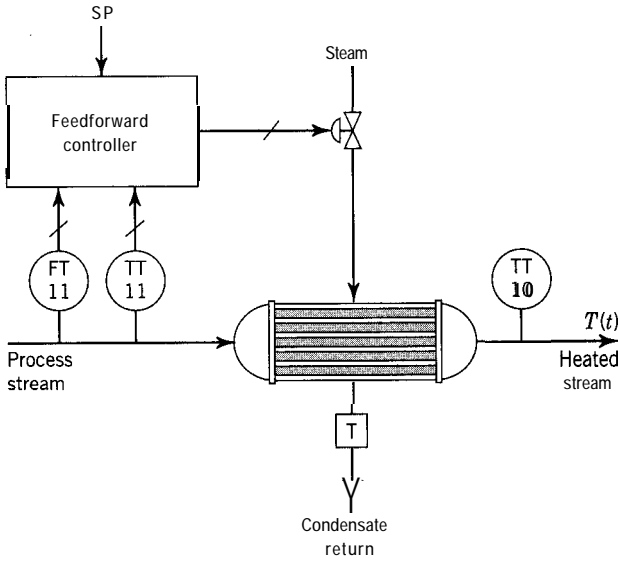
Figure l-S.2 Heat exchanger feedforward control system.

bances, while feedback control compensates for all other disturbances. Chapter 12 presents the development of the feedforward controller. Actual industrial cases are used to discuss this important strategy in detail.

It is important to note that the three basic operations, M, D, A, are still present in this more "advanced" control strategy. Measurement is performed by the sensors and transmitters. Decision is made by both the feedforward and the feedback controllers. Action is taken by the steam valve.

The advanced control strategies are usually more costly than feedback control in
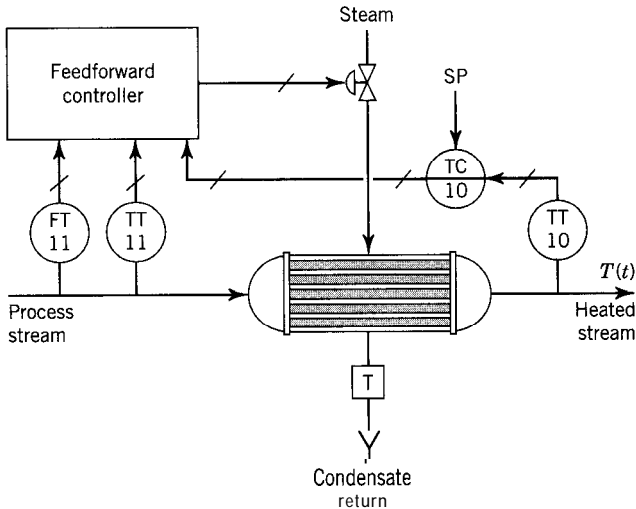
Figure 1-5.3 Heat exchanger feedforward control with feedback compensation.

hardware, computing power, and the effort involved in designing, implementing, and maintaining them. Therefore, the expense must be justified before they can be implemented. The best procedure is first to design and implement a simple control strategy, keeping in mind that if it does not prove satisfactory, then a more advanced strategy may be justifiable. It is important, however, to recognize that these advanced strategies still require some feedback compensation.

1-6 BACKGROUND NEEDED FOR PROCESS CONTROL

To be successful in the practice of automatic process control, the engineer must first understand the principles of process engineering. Therefore, this book assumes that the reader is familiar with the basic principles of thermodynamics, fluid flow, heat transfer, separation processes, reaction processes, and the like.

For the study of process control, it is also fundamental to understand how processes behave dynamically. Thus it is necessary to develop the set of equations that describes different processes. This is called *modeling*. To do this requires knowledge of the basic principles mentioned in the previous paragraph and of mathematics through differential equations. **Laplace** transforms are used heavily in process control. This greatly simplifies the solution of differential equations and the dynamic analysis of processes and their control systems. Chapter 2 of this book is devoted to the development and use of the **Laplace** transforms, along with a review of complex-number algebra. Chapters 3 and 4 offer an introduction to the modeling of some processes.

1-7 SUMMARY

In this chapter, we discussed the need for automatic process control. *Industrial processes are not static but rather very dynamic; they are continuously changing as a result of many types of disturbances.* It is principally because of this dynamic nature that control systems are needed to continuously and automatically watch over the variables that must be controlled.

The working principles of a control system can be summarized with the three letters M, D, and A. M refers to the measurement of process variables. D refers to the decision made on the basis of the measurement of those process variables. Finally, A refers to the action taken on the basis of that decision.

The fundamental components of a process control system were also presented: sensor/ transmitter, controller, and final control element. The most common types of **signals**— pneumatic, electrical, and digital-were introduced, along with the purpose of transducers.

Two control strategies were presented: feedback and feedforward control. The advantages and disadvantages of both strategies were briefly discussed. Chapters 6 and 7 present the design and analysis of feedback control loops.

PROBLEMS

1-l. For the following automatic control systems commonly encountered in daily life, identify the devices that perform the measurement (M), decision (D), and action

(A) functions, and classify the action function as "On/Off" or "Regulating." Also draw a process and instrumentation diagram (P&ID), using the standard ISA symbols given in Appendix A, and determine whether the control is feedback or **feed-** forward.

    (a) House air conditioning/heating
    (b) Cooking oven
    (c) Toaster
    (d) Automatic sprinkler system for fires
    (e) Automobile cruise speed control
    (f) Refrigerator

1-2. Instrumentation Diagram: Automatic Shower Temperature Control. Sketch the process and instrumentation diagram for an automatic control system to control the temperature of the water from a common shower-that is, a system that will automatically do what you do when you adjust the temperature of the water when you take a shower. Use the standard ISA instrumentation symbols given in Appendix A. Identify the measurement (M), decision (D), and action (A) devices of your control system.

Chapter 2

Mathematical Tools for Control Systems Analysis

This chapter presents two mathematical tools that are particularly useful for analyzing process dynamics and designing automatic control systems: Laplace transforms and linearization. Combined, these two techniques allow us to gain insight into the dynamic responses of a wide variety of processes and instruments. In contrast, the technique of computer simulation provides us with a more accurate and detailed analysis of the dynamic behavior of specific systems but seldom allows us to generalize our findings to other processes.

Laplace transforms are used to convert the differential equations that represent the dynamic behavior of process output variables into algebraic equations. It is then possible to isolate in the resulting algebraic equations what is characteristic of the process, the *transfer function,* from what is characteristic of the input forcing functions. Because the differential equations that represent most processes are nonlinear, linearization is required to approximate nonlinear differential equations with linear ones that can then be treated by the method of Laplace transforms.

The material in this chapter is not just a simple review of Laplace transforms but is a presentation of the tool in the way it is used to analyze process dynamics and to design control systems. Also presented are the responses of some common process transfer functions to some common input functions. These responses are related to the parameters of the process transfer functions so that the important characteristics of the responses can be inferred directly from the transfer functions without having to re-invert them each time. Because a familiarity with complex numbers is required to work with Laplace transforms, we have included a brief review of complex-number algebra as a separate section. We firmly believe that a knowledge of Laplace transforms is essential for understanding the fundamentals of process dynamics and control systems design.

2-1 THE **LAPLACE** TRANSFORM

This section reviews the definition of the Laplace transform and its properties.

11

2-1.1   Definition of the Laplace Transform

In the analysis of process dynamics, the process variables and control signals are functions of time, t. The **Laplace** transform of a function of time, f(t), is defined by the formula

$$F(s) = \mathcal{L}[f(t)] = \int_0^{\infty} f(t)e^{-st}\, dt \qquad\qquad \textbf{(2-1.1)}$$

where

$\quad$ F(s) = the **Laplace** transform off(t)
$\qquad s$ = the **Laplace** transform variable, time-'

The **Laplace** transform changes the function of time, $f(t)$, into a function in the **Laplace** transform variable, F(s). The limits of integration show that the **Laplace** transform contains information on the function $f(t)$ for positive time only. This is perfectly acceptable, because in process control, as in life, nothing can be done about the past (negative time); control action can affect the process only in the future. The following example uses the definition of the **Laplace** transform to develop the transforms of a few common forcing functions.

EXAMPLE 2-1.1

The four signals shown in Fig. 2-1.1 are commonly applied as inputs to processes and instruments to study their dynamic responses. We now use the definition of the **Laplace** transform to derive their transforms.

(a) UNIT STEP FUNCTION
$\quad$ This is a sudden change of unit magnitude as sketched in Fig. $2\text{-}1.1a$. Its algebraic representation is

$$u(t) = \begin{cases} 0 & t < 0 \\ 1 & t \geq 0 \end{cases}$$

Substituting into Eq. 2-1.1 yields

$$\mathcal{L}[u(t)] = \int_6^{\infty} u(t)e^{-st}\, dt = -\frac{1}{s}e^{-st}\Big|_0^{\infty}$$

$$= -\frac{1}{s}(0 - 1) = \frac{1}{s}$$
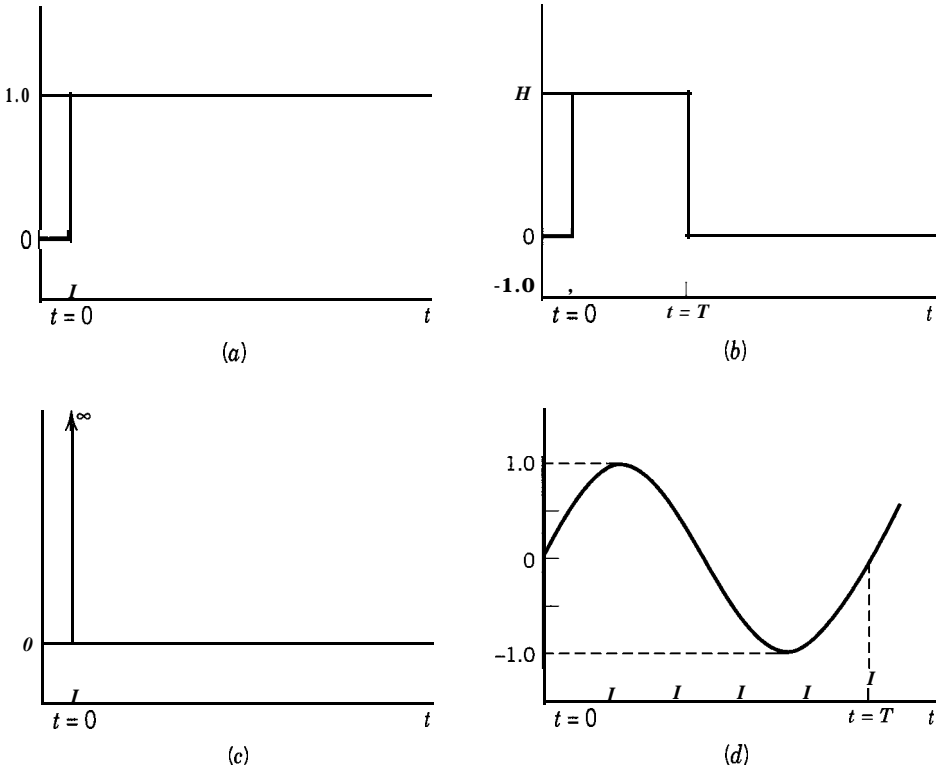
Figure 2-1.1 Common input signals for the study of control system response. (a) Unit step function, u(t). (b) Pulse. (c) Unit impulse function, $\delta(t)$. (d) Sine wave, $\sin \omega t$ ($\omega = 2\pi/T$).

(b) *A PULSE OF MAGNITUDE HAND DURATION T*

The pulse sketched in Fig. 2-1.1 *b* is represented by

$$f(t) = \begin{cases} 0 & t < 0, t \geq T \\ H & 0 \leq t < T \end{cases}$$

Substituting into Eq. 2-1.1 yields

$$\mathcal{L}[f(t)] = \int_0^\infty f(t)e^{-st}\, dt = \int_0^T He^{-st}\, dt$$

$$= -\frac{H}{s} e^{-st} \Big|_0^T = -\frac{H}{s}(e^{-sT} - 1)$$

$$= \frac{H}{s}(1 - e^{-sT})$$

(c) *A UNIT IMPULSE FUNCTION*

This function, also known as the *Dirac delta function* and represented by $\delta(t)$, is

sketched in Fig. 2-1.1c. It is an ideal pulse with zero duration and unit area. All of its area is concentrated at time zero. Because the function is zero at all times except at zero, and because the term e^{-st} in Eq. 2-1.1 is equal to unity at $t = 0$, the **Laplace** transform is

$$\mathcal{L}[\delta(t)] = \int_0^\infty \delta(t)e^{-st}\, dt = 1$$

Note that the result of the integration, 1, is the area of the impulse. The same result can be obtained by substituting $H = 1/T$ in the result of part (b), so that $HT = 1$, and then taking limits as T *goes* to zero.

(d) *A SINE WAVE OF UNITY AMPLITUDE AND FREQUENCY* ω
The sine wave is sketched in Fig. 2-1. 1d and is represented in exponential form by

$$\sin\, \omega t = \frac{e^{i\omega t} - e^{-i\omega t}}{2i}$$

where $i = \sqrt{-1}$ is the unit of imaginary numbers. Substituting into Eq. 2-1.1 yields

$$\mathcal{L}[\sin\, \omega t] = \int_0^\infty \frac{e^{i\omega t} - e^{-i\omega t}}{2i} e^{-st}\, dt$$

$$= \frac{1}{2i} \int_0^\infty [e^{-(s-i\omega)t} - e^{-(s+i\omega)t}]\, dt$$

$$= \frac{1}{2i} \left[\frac{e^{-(s-i\omega)t}}{s - i\omega} + \frac{e^{-(s+i\omega)t}}{s + i\omega} \Big|_0^\infty \right.$$

$$= \frac{1}{2i} \left[\frac{0 - 1 + 0 - 1}{s - i\omega \quad s + i\omega} \right]$$

$$= \frac{1}{2i} \frac{2iw}{s^2 + \omega^2}$$

$$= \frac{\omega}{s^2 + \omega^2}$$

The preceding example illustrates some algebraic manipulations required to derive the **Laplace** transform of various functions using its definition. Table 2- 1.1 contains a short list of the **Laplace** transforms of some common functions.

2-1.2   Properties of the **Laplace** Transform

This section presents the properties of **Laplace** transforms in order of their usefulness in analyzing process dynamics and designing control systems. Linearity and the real differentiation and integration theorems are essential for transforming differential equations into algebraic equations. The final value theorem is useful for predicting the final

Table 2-1.1 Laplace Transforms of Common
Functions

$f(t)$	$F(s) = \mathcal{L}[f(t)]$
$\delta(t)$	1
$u(t)$	$\dfrac{1}{s}$
t	$\dfrac{1}{s^2}$
t^n	$\dfrac{n!}{s^{n+1}}$
e^{-at}	$\dfrac{1}{s+a}$
te^{-at}	$\dfrac{1}{(s+a)^2}$
$t^n e^{-at}$	$\dfrac{n!}{(s+a)^{n+1}}$
$\sin \omega t$	$\dfrac{\omega}{s^2 + \omega^2}$
$\cos \omega t$	$\dfrac{s}{s^2 + \omega^2}$
$e^{-at} \sin \omega t$	$\dfrac{\omega}{(s+a)^2 + \omega^2}$
$e^{-at} \cos \omega t$	$\dfrac{s+a}{(s+a)^2 + \omega^2}$

steady-state value of a time function from its Laplace transform, and the real translation theorem is useful for dealing with functions delayed in time. Other properties are useful for deriving the transforms of complex functions from the transforms of simpler functions such as those listed in Table 2-1.1.

Linearity

It is very important to realize that the Laplace transform is a linear operation. This means that if a is a constant, then

$$\mathcal{L}[af(t)] = a\mathcal{L}[f(t)] = aF(s) \qquad \text{(2-1.2)}$$

The distributive property of addition also follows from the linearity property:

$$\mathcal{L}[af(t) + bg(t)] = aF(s) + bG(s) \qquad \text{(2-1.3)}$$

where a and b are constants. You can easily derive both formulas by application of Eq. 2- 1.1, the definition of the Laplace transform.

Real   Differentiation   Theorem

This theorem, which establishes a relationship between the Laplace transform of a function and that of its derivatives, is most important in transforming differential equations into algebraic equations. It states that

$$\mathscr{L}\left[\frac{df(t)}{dt}\right] = sF(s) - f(0) \tag{2-1.4}$$

Proof From the definition of the Laplace transform, Eq. 2-1.1,

$$\mathscr{L}\left[\frac{df(t)}{dt}\right] = \int_0^\infty \frac{df(t)}{dt} e^{-st}\, dt$$

Integrate by parts.

$$u = e^{-st} \qquad dv = \frac{df(t)}{dt}\, dt$$

$$du = -se^{-st}\, dt \qquad v = f(t)$$

$$\mathscr{L}\left[\frac{df(t)}{dt}\right] = [f(t)e^{-st}]_0^\infty - \int_0^\infty f(t)(-se^{-st}\, dt)$$

$$= [0 - f(0)] + s\int_0^\infty f(t)e^{-st}\, dt$$

$$= sF(s) - f(O) \qquad \text{q.e.d.}$$

The extension to higher derivatives is straightforward.

$$\mathscr{L}\left[\frac{d^2f(t)}{dt^2}\right] = \mathscr{L}\left[\frac{d}{dt}\left(\frac{df(t)}{dt}\right)\right]$$

$$= s\mathscr{L}\left[\frac{df(t)}{dt}\right] - \frac{df}{dt}\bigg|_{t=0}$$

$$= s[sF(s) - f(0)] - \frac{df}{dt}\bigg|_{t=0}$$

$$= s^2F(s) - sf(0) - \frac{df}{dt}\bigg|_{t=0}$$

In general,

$$\mathscr{L}\left[\frac{d^nf(t)}{dt^n}\right] = s^nF(s) \quad s^{n-1}f(0) - \ldots - \frac{d^{n-1}f}{dt^{n-1}}\bigg|_{t=0} \tag{2-1.5}$$

In process control, it is normally assumed that the initial conditions are at steady state (time derivatives are zero) and that the variables are deviations from initial conditions (initial value is zero). For this very important case, the preceding expression reduces to

$$\mathscr{L}\left[\frac{d^n f(t)}{dt^n}\right] = s^n F(s)$$

(2-1.6)

This means that for the case of zero initial conditions at steady state, the **Laplace** transform of the derivative of a function is obtained by simply substituting variable s for the "d/dt" operator, and $F(s)$ for $f(t)$.

Real   Integration   Theorem

This theorem establishes the relationship between the **Laplace** transform of a function and **that** of its integral. It states that

$$\mathscr{L}\left[\int_0^t f(t)\, dt\right] = \frac{1}{s} F(s)$$

(2-1.7)

The proof of this theorem is carried out by integrating the definition of the **Laplace** transform by parts. This proof is similar to that of the real differentiation theorem and is left as an exercise. The **Laplace** transform of the nth integral of a function is the transform of the function divided by s^n.

Real   Translation   Theorem

This theorem deals with the translation of a function in the time axis, as shown in Fig. 2-1.2. The translated function is the original function delayed in time. As we shall see in Chapter 3, time delays are caused by transportation lag, a phenomenon also known as *dead time.* The theorem states that

$$\mathscr{L}[f(t - t_0)] = e^{-st_0} F(s)$$

(2-1.8)

Because the **Laplace** transform does not contain information about the original function for negative time, the delayed function must be zero for all times less than the time delay (see Fig. 2- 1.2). This condition is satisfied if the process variables are expressed as deviations from initial steady-state conditions.

Proof. From the definition of the **Laplace** transform, Eq. 2- 1.1,

$$\mathscr{L}[f(t - t_0)] = \int_0^\infty f(t - t_0) e^{-st}\, dt$$
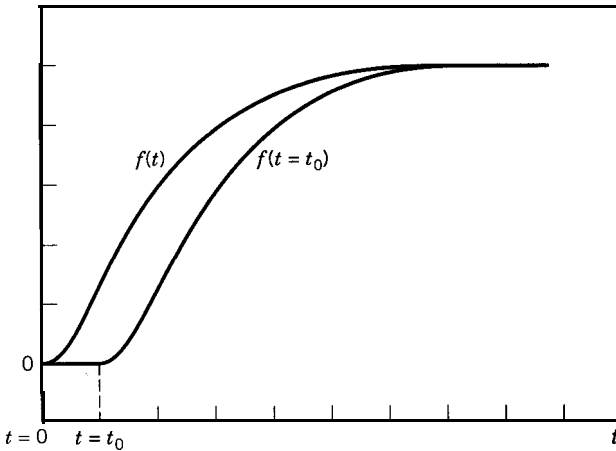
Figure 2-1.2 Function delayed in time is zero for all times less than the time delay t_0.

Let $\tau = t - t_0$ (or $t = t_0 + \tau$) and substitute.

$$\mathscr{L}[f(t - t_0)] = \int_{\tau=-t_0}^{\infty} f(\tau)e^{-s(t_0+\tau)} \, d(t_0 + \tau)$$

$$= \int_{\tau=0}^{\infty} f(\tau)e^{-st_0}e^{-s\tau} \, d\tau$$

$$= e^{-st_0} \int_{0}^{\infty} f(\tau)e^{-s\tau} \, d\tau$$

$$= e^{-st_0}F(s) \qquad \text{q.e.d.}$$

Note that in this proof, we made use of the fact that $f(\tau) = 0$ for $\tau < 0$ ($t < t_0$).

Final   Value   Theorem

This theorem allows us to figure out the final, or steady-state, value of a function from its transform. It is also useful in checking the validity of derived transforms. If the limit of f(t) as $t \rightarrow \infty$ exists, then it can be found from its Laplace transform as follows:

$$\lim_{t\to\infty} f(t) = \lim_{s\to 0} sF(s)$$

(2-1.9)

The proof of this theorem adds little to our understanding of it.

The last three properties of the Laplace transform, to be presented next without proof, are not used as often in the analysis of process dynamics as are the ones already presented.

Complex Differentiation Theorem

This theorem is useful for evaluating the transforms of functions that involve powers of the independent variable, t. It states that

$$\mathcal{L}[tf(t)] = -\frac{d}{ds} F(s) \tag{2-1.10}$$

Complex Translation Theorem

This theorem is useful for evaluating transforms of functions that involve exponential functions of time. It states that

$$\mathcal{L}[e^{at}f(t)] = F(s - a) \tag{2-1.11}$$

Initial Value Theorem

This theorem enables us to calculate the initial value of a function from its transform. It would provide another check of the validity of derived transforms were it not for the fact that in process dynamic analysis, the initial conditions of the variables are usually zero. The theorem states that

$$\lim_{t \to 0} f(t) = \lim_{s \to \infty} sF(s) \tag{2-1.12}$$

The following examples illustrate the use of the properties of Laplace transforms we have just discussed.

EXAMPLE 2-1.2

Derive the Laplace transform of the differential equation

$$9 \frac{d^2y(t)}{dt^2} + 6 \frac{dy(t)}{dt} + y(t) = 2x(t)$$

with initial conditions of zero at steady state-that is, $y(0) = 0$ and $dy/dt|_{t=0} = 0$.

SOLUTION

By application of the linearity property, Eq. 2-1.3, take the Laplace transform of each term.

$$9\mathcal{L}\left[\frac{d^2y(t)}{dt^2}\right] + 6\mathcal{L}\left[\frac{dy(t)}{dt}\right] + \mathcal{L}[y(t)] = 2\mathcal{L}[x(t)]$$

Then apply the real differentiation theorem, Eq. 2-1.6.

$$9s^2Y(s) + 6sY(s) + Y(s) = 2X(s)$$

Finally, solve for Y(s).

$$Y(s) = \frac{2}{9s^2 + 6s + 1}X(s)$$

The preceding example shows how the **Laplace** transform converts the original differential equation into an algebraic equation that can then be rearranged to solve for the dependent variable Y(s). Herein lies the great usefulness of the **Laplace** transform, because algebraic equations are a lot easier to manipulate than differential equations.

EXAMPLE 2-1.3

Obtain the **Laplace** transform of the following function:

$$c(t) = u(t - 3)[1 - e^{-(t-3)/4}]$$

Note: The term $u(t - 3)$ in this expression shows that the function is zero for $t < 3$. We recall, from Example 2-1.1(a), that $u(t - 3)$ is a change from zero to one at $t = 3$, which means that the expression in brackets is multiplied by zero until $t = 3$ and is multiplied by unity after that. Thus the presence of the unit step function does not alter the rest of the function for $t \geq 3$.

SOLUTION

Let

$$c(t) = f(t - 3) = u(t - 3)[1 - e^{-(t-3)/4}]$$

Then

$$f(t) = u(t)[1 - e^{-t/4}]$$
$$= u(t) - u(t)e^{-t/4}$$

Apply Eq. 2-1.3, the linearity property, and use entries from Table 2-1.1 with $a = 1/4$.

$$F(s) = \frac{1}{s} - \frac{1}{s + \frac{1}{4}} = \frac{1}{s(4s + 1)}$$

Next apply the real translation theorem, Eq. 2-1.8.

$$C(s) = \mathcal{L}[f(t - 3)] = e^{-3s}F(s)$$

$$C(s) = \frac{e^{-3s}}{s(4s + 1)}$$

We can check the validity of this answer by using the final value theorem, Eq. 2-1.9.

$$\lim_{t \to \infty} c(t) = \lim_{t \to \infty} u(t - 3) \left[1 - e^{-(t-3)/4}\right] = 1$$

$$\lim_{s \to 0} sC(s) = \lim_{s \to 0} s \frac{e^{-3s}}{s(4s + 1)} = 1 \qquad \text{Check!}$$

2-2 SOLUTION OF DIFFERENTIAL EQUATIONS USING THE LAPLACE TRANSFORM

This section presents the use of the Laplace transform to solve the differential equations that represent the dynamics of processes and their control systems. Because our objective is to find out how the output signals respond to input forcing functions, we will always assume that the initial conditions are at steady state (zero time derivatives). We will also define all variables as deviations from their initial values. This forces the initial values of the deviation variables also to be zero.

2-2.1 Laplace Transform Solution Procedure

The procedure for solving a differential equation by Laplace transforms consists of three steps:

1. Transform the differential equation into an algebraic equation in the Laplace transform variable s.
2. Solve for the transform of the output (or dependent) variable.
3. Invert the transform to obtain the response of the output variable with time, t.

Consider the following second-order differential equation:

$$a_2 \frac{d^2 y(t)}{dt^2} + a_1 \frac{dy(t)}{dt} + a\&t) = bx(t) \qquad (2\text{-}2.1)$$

The problem of solving this equation can be stated as follows: Given the constant coefficients a_1, a_1, a_2, and b, the initial conditions $y(0)$ and $dy/dt \mid_{t=0}$, and the function $x(t)$, find the function $y(t)$ that satisfies the differential equation.

We call the function $x(t)$ the "forcing function" or input variable, and we call $y(t)$ the "output" or dependent variable. In process control systems, a differential equation such as Eq. 2-2.1 usually represents how a particular process or instrument relates its output signal, $y(t)$, to its input signal, $x(t)$. Our approach is that of British inventor James

Watt (1736–1819), who considered process variables as signals and processes as signal processors.

The first step is to take the Laplace transform of Eq. 2-2.1. We do this by applying the linearity property of Laplace transforms, Eq. 2-1.3, which allows us to take the Laplace transform of each term separately:

$$a_2 \mathcal{L}\left[\frac{d^2 y(t)}{dt^2}\right] + a_1 \mathcal{L}\left[\frac{dy(t)}{dt}\right] + a_0 \mathcal{L}[y(t)] = b\mathcal{L}[x(t)] \qquad (2\text{-}2.2)$$

Assuming for the moment that the initial conditions are not zero, the indicated Laplace transforms are obtained by using the real differentiation theorem, Eq. 2-1.5.

$$\mathcal{L}\left[\frac{d^2 y(t)}{dt^2}\right] = s^2 Y(s) - sy(0) - \left.\frac{dy}{dt}\right|_{t=0}$$

$$\mathcal{L}\left[\frac{dy(t)}{dt}\right]_1 = sY(s) - y(0)$$

Next we substitute these terms into Eq. 2-2.2 and rearrange it to obtain

$$(a_2 s^2 + a_1 s + a_0)Y(s) - (a_2 s + a_1)y(0) - a_2 \left.\frac{dy}{dt}\right|_{t=0} = bX(s)$$

The second step is to manipulate this algebraic equation to solve for the transform of the output variable, Y(s).

$$Y(s) = \frac{bX(s) + (a_2 s + a_1)y(0) + a_2 \left.\dfrac{dy}{dt}\right|_{t=0}}{a_2 s^2 + a_1 s + a_,} \qquad (2\text{-}2.3)$$

This equation shows the effect of the input variable, X(s), and of the initial conditions on the output variable. Our objective is to study how the output variable responds to the input variable, so the presence of the initial conditions complicates our analysis. To avoid this unnecessary complication, we assume that the initial conditions are at steady state, $dy/dt\,|_{t=0} = 0$, and define the output variable as the *deviation* from its initial value, thus forcing $y(0) = 0$. We will show in the next section how this can be done without loss of generality. With zero initial conditions, the equation is reduced to

$$Y(s) = \left[\frac{b}{a_2 s^2 + a_1 s + a_1}\right] X(s) \qquad (2\text{-}2.4)$$

The form of Eq. 2-2.4 allows us to break the transform of the output variable into the product of two terms: the term in brackets, known as the *transfer function,* and the transform of the input variable, X(s). The transfer function and its parameters charac-

terize the process or device and determine how the output variable responds to the input variable. The concept of transfer function is described in more detail in Chapter 3.

The third and final step is to invert the transform of the output to obtain the time function $y(t)$, which is the response of the output. Inversion is the opposite operation to taking the Laplace transform. Before we can invert, we must select a specific input function for $x(t)$. A common function, because of its simplicity, is the unit step function, u(t), which was introduced in Example 2-1.1. From that example, or from Table 2-1.1, we learn that for x(t) = $u(t)$, $X(s)$ = 1/s. We substitute into Eq. 2-2.4 and invert to obtain

$$y(t) = \mathcal{L}^{-1}\left[\frac{b}{a_2 s^2 + a_1 s + a,}\frac{1}{s}\right] \qquad (2\text{-}2.5)$$

where the symbol $\mathcal{L}^{-1}$ stands for the inverse Laplace transform. The response to a step input is called the **step response** for short.

The inversion could easily be carried out if we could find the expression within the brackets in Table 2-1.1 or in a more extensive table of Laplace transforms. Obviously, we will not be able to find complex expressions in such a table. The mathematical technique of partial fractions expansion, to be introduced next, is designed to expand the transform of the output into a sum of simpler terms. We can then invert these simpler terms separately by matching entries in Table 2-1.1.

2-2.2   Inversion by Partial Fractions Expansion

The mathematical technique of partial fractions expansion was introduced by the British physicist Oliver Heaviside (1850– 1925) as part of his revolutionary "operational calculus." The first step in expanding the transform, Eq. 2-2.5, into a sum of fractions is to factor its denominator, as follows:

$$(a_2 s^2 + a_1 s + a_0)s = a_2(s - r_1)(s - r_2)s \qquad (2\text{-}2.6)$$

where r_1 and r_2 are the roots of the quadratic term-that is, the values of s that satisfy the equation

$$a_2 s^2 + a_1 s + a, = 0$$

For a quadratic, or second-degree, polynomial, the roots can be calculated by the standard quadratic formula:

$$r_{1,2} = \frac{-a_1 \pm \sqrt{a_1^2 - 4a_2 a_0}}{2a_2} \qquad (2\text{-}2.7)$$

For higher-degree polynomials, the reader is referred to any numerical methods text for a root-finding procedure. Most electronic calculators are now able to find the roots of

third- and higher-degree polynomials. Computer programs such as Mathcad[1] and MATLAB[2] provide functions for finding the roots of polynomials of any degree.

Once the denominator is factored into first-degree terms, the transform is expanded into partial fractions as follows:

$$Y(s) = \frac{A_1}{s - r_1} + \frac{A_2}{s - r_2} + \frac{A_3}{s} \qquad (2\text{-}2.8)$$

provided that the roots, r_1, r_2, and $r_3 = 0$, are not equal to each other. For this case of unrepeated roots, the constant coefficients are found by the formula

$$A, = \lim_{s \to r_k} (s - r_k)Y(s) \qquad (2\text{-}2.9)$$

We can now carry out the inversion of Eq. 2-2.8 by matching each term to entries in Table 2-1 .1; in this case the first two terms match the exponential function with a $= - r_k$, and the third term matches the unit step function. The resulting inverse function is

$$y(t) = A_1 e^{r_1 t} + A_2 e^{r_2 t} + A_3 u(t)$$

Repeated   Roots

For the case of *repeated roots,* say $r_1 = r_2$, the expansion is carried out as follows:

$$Y(s) = \frac{\Lambda_1}{(s - r_1)^2} + \frac{A_2}{s\text{-}r,} + \frac{A_3}{s} \qquad (2\text{-}2.10)$$

Coefficient *A,* is calculated as before, but coefficients *A,* and *A,* must be calculated by the following formulas:

$$A, = \lim_{s \to r_1} (s - r_1)^2 Y(s)$$

$$A_2 = \lim_{s \to r_1} \frac{1}{1!} \frac{d}{ds} [(s - r_1)^2 Y(s)]$$

Again, we carry out the inversion of Eq. 2-2.10 by matching terms in Table 2-1.1. The first term matches the sixth term in the table with a $= - r_1$, to give the inverse.

$$y(t) = A_1 t e^{r_1 t} + A_2 e^{r_1 t} + A_3 u(t) \qquad (2\text{-}2.11)$$

---

[1] *Mathcad* User's *Guide,* by MathSoft, Inc., 201 Broadway, Cambridge, MA, 02139, 1992.
[2] *MATLAB* User's *Guide,* The MathWorks, Inc., 24 Prime Park Way, Natick, MA, 01760, 1992.

In general, if root r_1 is repeated m times, the expansion is carried out as follows:

$$Y(s) = \frac{A_1}{(s - r_1)^m} + \frac{A_2}{(s - r_1)^{m-1}} + \cdots + \frac{A_m}{s - r_1} + \cdots \qquad (2\text{-}2.12)$$

The coefficients are calculated by

$$A_1 = \lim_{s \to r_1} (s - r_1)^m Y(s)$$

$$A_k = \lim_{s \to r_1} \frac{1}{(k - 1)!} \frac{d^{k-1}}{ds^{k-1}} [(s - r_1)^m Y(s)] \qquad (2\text{-}2.13)$$

for $k = 2, \ldots, m.$ The inverse function is then:

$$y(t) = \left[\frac{A_1 t^{m-1}}{(m - 1)!} + \frac{A_2 t^{m-2}}{(m - 2)!} + \cdots + A_m \, e^{r_1 t} + \cdots \right. \qquad (2\text{-}2.14)$$

The following example is designed to illustrate numerically the partial fractions expansion procedure and the entire inversion process. Three cases are considered: unrepeated real roots, repeated roots, and complex conjugate roots.

EXAMPLE 2-2.1

Given the quadratic differential equation considered in the preceding discussion, Eq. 2-2.1, with zero steady-state initial conditions, we will obtain the unit step response of the output variable $y(t)$ for three different sets of parameters.

(a) *UNREPEATED REAL ROOTS*

Let $a_2 = 9$, a, $= 10$, a, $= 1$, and $b = 2$, in Eq. (2-2.1). Then the unit step response is, from Eq. 2-2.5,

$$y(t) = \mathcal{L}^{-1} \left[\frac{2}{9s^2 + 10s + 1} \frac{1}{s} \right]$$

The roots, from the quadratic equation, are $r_1 = -1/9$, $r_2 = -1$. The denominator is factored as follows:

$$Y(s) = \frac{2}{9 \left(s + \dfrac{1}{9} \right) (s + 1)} \frac{1}{s}$$

$$= \frac{A_1}{s + \dfrac{1}{9}} + \frac{A_2}{s + 1} + \frac{A_3}{s}$$

The coefficients are calculated using Eq. 2-2.9.

$$A, = \lim_{s \to -1/9} \frac{2}{9(s + 1)s} = -2.25$$

$$A_2 = \lim_{s \to -1} \frac{2}{9\left(s + \dfrac{1}{9}\right)s} = 0.25$$

$$A_3 = \lim_{s \to 0} \frac{2}{9\left(s + \dfrac{1}{9}\right)(s + 1)} = 2$$

Invert by matching entries in Table 2-1.1 to obtain the step response.

$$y(t) = -2.25e^{-t/9} + 0.25e^{-t} + 2u(t)$$

(b) *REPEATED ROOTS*
Let a, = 6, and let the other parameters be as before. The roots, from the quadratic formula, are $r_1 = r_2 = -1/3$, and the **Laplace** transform of the output response is

$$Y(s) = \frac{2}{9\left(s + \dfrac{1}{3}\right)^2 s}$$

$$= \frac{A_1}{\left(s + \dfrac{1}{3}\right)^2} + \frac{A_2}{s + \dfrac{1}{3}} + \frac{A_3}{s}$$

The coefficients are, from Eq. 2-2.13,

$$A, = \lim_{s \to -1/3} \frac{2}{9s} = -\frac{2}{3}$$

$$A_2 = \lim_{s \to -1/3} \frac{1}{1!} \frac{d}{d} \left[\frac{2}{9s}\right]$$

$$= \lim_{s \to -1/3} -\frac{2}{9s^2} = -2$$

and A, = 2, as before. The step response is then obtained by matching entries in Table 2-1.1.

$$y(t) = \left(-\frac{2}{3}t - 2\right)e^{-t/3} + 2u(t)$$

(c) *PAIR OF COMPLEX CONJUGATE ROOTS*

Let a, = 3, and let the other parameters be as before. The roots, from the quadratic formula, are $r_{1,2} = -0.167 \pm i0.289$, where $i = \sqrt{-1}$ is the unit of the imaginary numbers. The transform of the output is then

$$Y(s) = \frac{2}{9(s + 0.167 \quad i0.289)(s + 0.167 + i0.289)s}$$

$$\blacksquare \quad \frac{A_1}{s + 0.167 - i0.289} + \frac{A_2}{s + 0.167 + i0.289} + \frac{A_3}{s}$$

Once more we calculate the coefficients by Eq. 2-2.9.

$$A, = \lim_{s \to -0.167 + i0.289} \frac{2}{9(s + 0.167 + i0.289)s} = -1 + i0.577$$

$$A, = \lim_{s \to -0.167 - i0.289} \frac{2}{9(s + 0.167 - i0.289)\sim} = -1 - i0.577$$

and A, = 2, as before. The inverse response is again obtained by matching entries in Table 2-1.1. Note that the fact that the numbers are complex does not affect this part of the procedure.

$$y(t) = (-1 + i0.577)e^{(-0.167 + i0.289)t} + \qquad (-1 - i0.577)e^{(-0.167 - i0.289)t} + \qquad 2u(t)$$

It is evident from the preceding example that calculating the coefficients of the partial fractions expansion can be difficult, especially when the factors of the transform are complex numbers. As we shall see in the next section, the roots of the denominator of the transfer function contain most of the significant information about the response. Consequently, in analyzing the response of process control systems, it is seldom necessary to calculate the coefficients of the partial fractions expansion. This is indeed fortunate.

2-2.3 Handling Time Delays

The technique of partial fractions expansion is restricted to use with Laplace transforms that can be expressed as the ratio of two polynomials. When the response contains time delays, by the real translation theorem, Eq. 2-1.8, an exponential function of s appears in the transform. Because the exponential is a transcendental function, we must appropriately modify the inversion procedure.

If the denominator of the transform contains exponential functions of s, it cannot be factored because the exponential function introduces an infinite number of factors. On the other hand, we can handle exponential terms in the numerator of the transform, as we shall now see.

Consider the case in which there is a single exponential term that can be factored as follows:

$$Y(s) = Y_1(s)e^{-st_0} \qquad (2\text{-}2.15)$$

The correct procedure is to expand in partial fractions the portion of the transform that does not contain the exponential term.

$$Y_1(s) = \frac{A_1}{s - r_1} + \frac{A_2}{s - r_2} + \cdots + \frac{A_n}{s - r_n} \qquad (2\text{-}2.16)$$

Then invert this expression.

$$y_1(t) = A_1 e^{r_1 t} + A_2 e^{r_2 t} + \ldots + A_n e^{r_n t} \qquad (2\text{-}2.17)$$

Now invert Eq. 2-2.15, making use of the real translation theorem, Eq. 2-1.8.

$$y(t) = \mathcal{L}^{-1}[e^{-st_0}Y_1(s)] = y_1(t - t_0)$$
$$= A_1 e^{r_1(t-t_0)} + A_2 e^{r_2(t-t_0)} + \ldots + A_n e^{r_n(t-t_0)} \qquad (2\text{-}2.18)$$

It is important to realize that the exponential term must be excluded from the partial fractions expansion procedure. Although inclusion of the exponential term in the partial fractions expansion may give the correct result in some special cases, doing so is fundamentally *incorrect*.

Next let us consider the case of multiple delays. When there are more delay terms than one in the numerator of the transform, proper algebraic manipulation will convert the transform into a sum of terms, each having a single exponential function:

$$Y(s) = Y_1(s)e^{-st_{01}} + Y_2(s)e^{-st_{02}} + \ldots \qquad (2\text{-}2.19)$$

Expand each of the sub-transforms-$Y_1(s)$, $Y_2(s)$, and so on-in partial fractions and invert them separately, leaving out the exponential terms. Finally, apply Eq. 2-2.18 to each term to produce the result

$$y(t) = y_1(t - t_{01}) + y_2(t - t_{02}) + \ldots \qquad (2\text{-}2.20)$$

The following example illustrates this procedure.

EXAMPLE 2-2.2

Given the differential equation

$$\frac{dc(t)}{dt} + 2c(t) = f(t)$$

with $c(0) = 0$, find the response of the output for

(a)

(b)
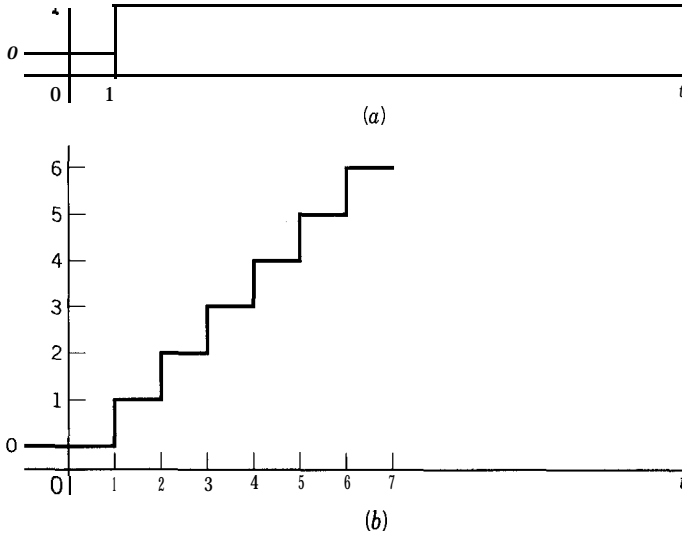
Figure 2-2.1 Input functions for Example 2-2.2. (a) Delayed unit step, $u(t - 1)$. (b) Staircase of unit steps.

(a) A unit step change at $t = 1$: $f(t) = u(t \quad 1)$
(b) A staircase function of unit steps at every unit of time

$$f(t) = u(t - 1) + u(t - 2) + u(t - 3) + \cdots$$

The functions are sketched in Fig. 2-2.1

SOLUTION

(a) Transform the differential equation, solve for C(s), and substitute $F(s) = (1/s)e^{-s}$.

$$\mathbf{C(s)} = \frac{1}{s + 2} F(s) = \frac{1}{s + 2} \frac{1}{s} e^{-s}$$

Let $C(s) = C_1(s)e^{-s}$, and then invert C,(s).

$$C_1(s) = \frac{11}{s + 2 \; s} = \frac{A_1}{s + 2} + \frac{A_2}{s}$$

$$A, = \lim_{s \to -2} (s + 2) \frac{1}{(s + 2)s} = -\frac{1}{2}$$

$$A_2 = \lim_{s \to 0} s \frac{1}{(s + 2)s} = \frac{1}{2}$$

Invert by matching entries in Table 2-1.1.

$$c_1(t) = -\frac{1}{2} e^{-2t} + \frac{1}{2} u(t)$$

$$= \frac{1}{2} u(t)[1 - e^{-2t}]$$

Apply Eq. 2-2.18.

$$c(t) = \mathcal{L}^{-1}[C_1(s)e^{-s}] = c_1(t - 1) = \frac{1}{2} u(t - 1)[1 - e^{-2(t-1)}]$$

Note that the unit step $u(t - 1)$ must multiply the exponential term to show that $c(t) = 0$ for $t < 1$.

(b) For the staircase function,

$$C(s) = \left[\frac{1}{s + 2}\right] \left(\frac{e^{-s}}{s} \quad \frac{e^{-2s}}{s} \quad \frac{e^{-3s}}{s} \quad \cdots\right)$$

$$= \left[\frac{1}{(s + 2)s}\right] (e^{-s} + e^{-2s} + e^{-3s} + \cdots)$$

$$= C_1(s)e^{-s} + C_1(s)e^{-2s} + C_1(s)e^{-3s} + \ldots$$

We note that $C_1(s)$ is the same as for part (a), and therefore $c_1(t)$ is the same. Applying Eq. 2-2.18 to each term results in

$$c(t) = c_1(t - 1) + c_1(t - 2) + c_1(t - 3) + \ldots$$

$$= \frac{1}{2} u(t - 1)[1 - e^{-2(t-1)}] + \frac{1}{2} u(t - 2)[1 - e^{-2(t-2)}]$$

$$+ \frac{1}{2} u(t - 3)[1 - e^{-2(t-3)}] + \ldots$$

The preceding example illustrates how to handle time delays in the input function. The same procedure can be applied when the time delay appears in the transfer function of the system. This situation arises in Chapter 3 in the models of processes with transportation lag.

2-3   CHARACTERIZATION OF PROCESS RESPONSE

In the preceding section, we learned that we can express the Laplace of the process output variable as the product of two terms: a transfer function, which is characteristic of the process, and the transform of the input signal. A major objective of this section is to relate the characteristics of the response of the output variable to the parameters of the process transfer function and, in particular, to the roots of the denominator of

the transfer function. We will see that most of the important information about the process response can be obtained from these roots, and it is not in general necessary to obtain the exact solution to each problem.

The relevant questions about the output response are the following:

- Is the response stable? That is, will it remain bound when forced by a bound input?
- If stable, what will be its final steady-state value?
- Is the response monotonic or oscillatory?
- If monotonic and stable, how long will it take for the transients to die out?
- If oscillatory, what is the period of oscillation and how long will it take for the oscillations to die out?

We will see that we can obtain the answers to all these questions from the parameters of the transfer function of the system. But first, let us formally define deviation variables and see how their use, combined with the assumption of steady-state initial conditions, allows us to eliminate the effect of the initial conditions on the response.

2-3.1 Deviation Variables

In Section 2-2 we saw that the response of the output variable is affected not only by the input variables but also by its initial conditions. Because we are interested in studying the response of processes and their control systems to the input variables (disturbances and manipulated variables, defined in Chapter 1), we want to eliminate the effect of the initial conditions on the response. To do this, we assume that the initial conditions are at steady state. This makes the initial values of the time derivatives equal to zero, but not the initial value of the output itself. To eliminate the initial value of the output, we replace the output variable with its deviation from the initial value. This gives rise to *deviation variables,* which we defined as

$$Y(t) = y(t) \quad y(0) \tag{2-3.1}$$

where

$$Y(t) = \text{deviation variable}$$
$$y(t) = \text{total value of the variable}$$

In the balance of this book, deviation variables will be represented by capital letters and absolute variables by lower-case letters, whenever possible. From the definition of a deviation variable, its initial value is always zero: $Y(O) = y(0) - y(0) = 0$.

To illustrate the simplifications that result from the use of deviation variables, consider the nth-order linear differential equation:

$$a_n \frac{d^n y(t)}{dt^n} + a_{n-1} \frac{d^{n-1} y(t)}{dt^{n-1}} + \ldots + a \& t)$$

$$= b_m \frac{d^m x(t)}{dt^m} + b_{m-1} \frac{d^{m-1} x(t)}{dt^{m-1}} + \ldots + b_0 x(t) + c \tag{2-3.2}$$

where $n > m$, y(t) is the output variable, x(t) is the input variable, and c is a constant. At the initial steady state, all the time derivatives are zero, and we can write

$$a_0 y(0) = b_0 x(0) + c \qquad (2\text{-}3.3)$$

Subtracting Eq. 2-3.3 from Eq. 2-3.2 results in

$$a_n \frac{d^n Y(t)}{dt^n} + a_{n-1} \frac{d^{n-1} Y(t)}{dt^{n-1}} + \ldots + a_0 Y(t)$$

$$= b_m \frac{d^m X(t)}{dt^m} + b_{m-1} \frac{d^{m-1} X(t)}{dt^{m-1}} + \ldots + b_0 X(t) \ (2\text{-}3.4)$$

where $Y(t) = y(t) - y(O)$, $X(t) = x(t) - x(O)$, and the deviation variables can be directly substituted for the respective variables in the derivative terms because they differ only by a constant bias.

$$\frac{d^k Y(t)}{dt^k} = \frac{d^k [y(t) - y(0)]}{dt^k} = \frac{d^k y(t)}{dt^k} - \frac{d^k y(0)}{dt^k} = \frac{d^k y(t)}{dt^k}$$

Note that Eq. 2-3.4 in the deviation variables is essentially the same as Eq. 2-3.2 in the original variables except for constant c, which cancels out. This result is general.

2-3.2 Output Response

To show the relationship between the output response and the roots of the denominator of the transfer function, let us **Laplace** transform the nth-order differential equation in the deviation variables, Eq. 2-3.4, and solve for the transform of the output.

$$Y(s) = \left[\frac{b_m s^m + b_{m-1} s^{m-1} + \ldots + b_0}{a_n s^n + a_{n-1} s^{n-1} + \ldots + a,} \right] X(s) \qquad (2\text{-}3.5)$$

where we have made use of the fact that all the initial conditions are zero. The expression in brackets is the transfer function; its denominator can be factored into n first-degree terms, one for each of its roots.

$$Y(s) = \left[\frac{b_m s^m + b_{m-1} s^{m-1} + \ldots + b_0}{a_n (s - r_1)(s - r_2), \ldots (s - r_n)} \right] X(s) \qquad (2\text{-}3.6)$$

where $r_1, r_2, \ldots, r_n$, are the roots of the denominator polynomial. Besides the n factors shown in Eq. 2-3.6, there are additional factors introduced by the input variable X(s) that depend on the type of input (step, pulse, ramp, and so on). Next we expand the transform in partial fractions.

$$Y(s) = \frac{\Lambda_1}{s - r_1} + \frac{A_2}{s - r_2} + \ldots + \frac{A_n}{s - r_n} + [\text{terms of X(s)}] \qquad (2\text{-}3.7)$$

Finally, we invert the transform by matching entries in Table 2-1.1 to obtain the response as a function of time. If there are no repeated roots, the inverse is

$$Y(t) = A_1 e^{r_1 t} + A_2 e^{r_2 t} + \cdots + A_n e^{r_n t} + (\text{terms of } X) \qquad (2\text{-}3.8)$$

The first n terms on the right-hand side come from the transfer function, and the rest of the terms differ depending on the input function X(t).

If any of the roots is repeated m times, then its coefficient is replaced by a polynomial in t of degree $m - 1$, as shown in the preceding section. The total number of terms will of course be n, counting the terms in the polynomial of t.

Let us next answer the questions posed at the beginning of this section by analyzing Eq. 2-3.8. We consider first the case in which all the roots are real and then the possibility of complex conjugate pairs of roots.

All Real Roots

If all the roots are real, then the terms of Eq. 2-3.8 are simple exponential functions of time that can only grow with time if the root is positive or decay to zero if the root is negative. Therefore, real roots cannot cause the response to oscillate. Furthermore, if any of the roots is positive, the response will grow exponentially without bound, so it will be unstable. You might ask, what if the coefficient of the term with the positive root is zero? The system is just as unstable; a zero coefficient merely means that for a particular input it may not run away (like a pencil standing on its sharpened point), but the slightest deviation from equilibrium will cause it to run away from that position.

Thus, in answer to our initial questions, if all the roots of the denominator of the transfer function are real, then

- The response is monotonic (nonoscillatory).
- It is stable only if all the roots are negative.

Figures 2-3.1*a* and *b* are, respectively, examples of stable and unstable monotonic responses.

Regarding the time it takes for the transients to die out, we can see that each exponential term starts at unity ($e^0 = 1$) and, if the root is negative, decays to zero with time. Theoretically, an exponential never reaches zero, so we have to define a threshold below which the transient can be considered gone. Let us say we define the threshold for each term of the response as less than 1% of its initial value. To use a good round number, let $e^{rt} = e^{-5} = 0.0067$, or 0.67%, which is less than 1%. Then the time required for the kth exponential term to reach 0.67% of its initial value is

$$t_k = \frac{-5}{r_k} \qquad (2\text{-}3.9)$$

Thus the root with the smallest absolute value (least negative) will take the longest to die out. Such a root is called the *dominant root* of the response.
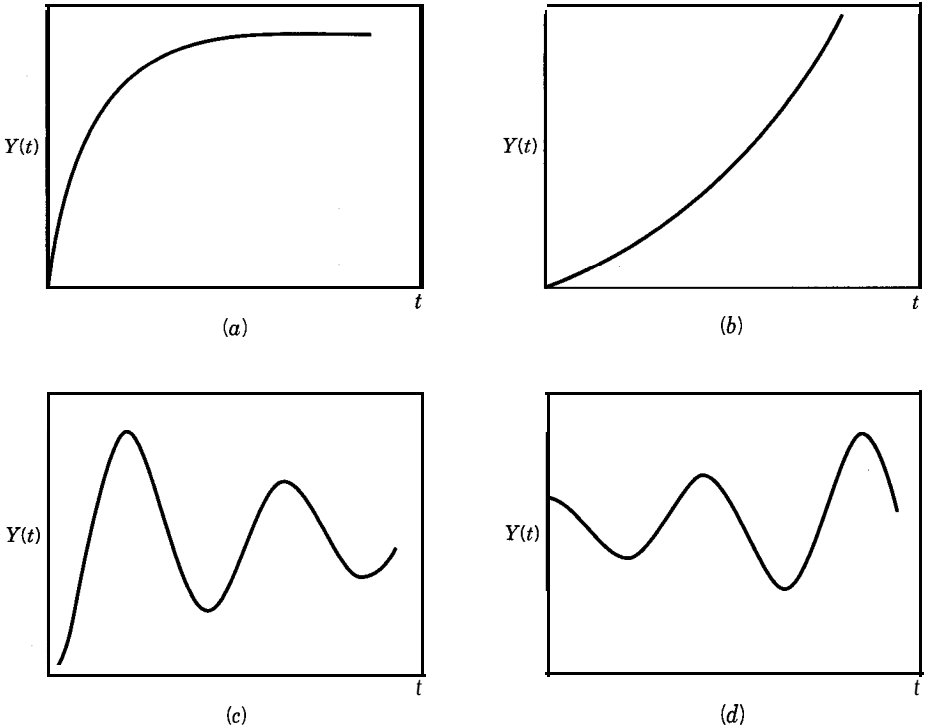
Figure 2-3.1 Examples of responses. (a) Stable, negative real root. (b) Unstable, positive real root. (c) Oscillatory stable, complex roots with negative real part. (d) Oscillatory unstable, complex roots with positive real part.

Pair of Complex Conjugate Roots

Complex roots of a real polynomial come in complex conjugate pairs, such as

$$r_1 = p + i\omega \qquad r_2 = p - i\omega$$

where p is the real part and ω is the imaginary part. For the case of one such pair of roots, the expanded transform of the output is

$$
\begin{aligned}
Y(s) &= \frac{A_1}{s - p - h} + \frac{A_2}{s - p + i\omega} + \cdots \\
&= \frac{(A_1 + A_2)(s - p)}{(s - p)^2 + \omega^2} + \frac{i(A_1 - A_2)\omega}{(s - p)^2 + \omega^2} + \cdots \\
&= \frac{B(s - p)}{(s - p)^2 + \omega^2} + \frac{C\,0}{(s - p)'} + 3 + \cdots
\end{aligned}
\tag{2-3.10}
$$

where

$$B = A, + A,$$
$$C = i(A_1 \quad A,)$$

It can be shown that A, and A, are complex numbers and are conjugates of each other. Consequently, B and C are real numbers. We can now invert Eq. 2-3.10 by matching the last two entries in Table 2-1.1, with $a = - \rho$.

$$Y(t) = Be^{\rho t} \cos \omega t + Ce^{\rho t} \sin \omega t$$

$$= e^{\rho t}[B \cos \omega t + C \sin \omega t] + \ldots$$

This equation can be further simplified by using the trigonometric identity

$$\sin(\omega t + \theta) = \sin \theta \cos \omega t + \cos \theta \sin \omega t$$

The result is

$$Y(t) = De^{\rho t} \sin(\omega t + \theta) + \ldots \cdot \tag{2-3.11}$$

where

$D = \sqrt{B^2 + C^2}$ is the initial amplitude

$\theta = \tan' \dfrac{B}{C}$ is the phase angle, in radians[3]

This result shows that the response is oscillatory, because it contains the sine wave. The amplitude of the sine wave varies with time according to the exponential term $e^{\rho t}$, which is initially zero but can grow with time if ρ is positive or decay to zero if ρ is negative. Thus, for the case of one or more pairs of complex conjugate roots, we can further answer the questions at the beginning of this section as follows:

- The response is oscillatory.
- The oscillations grow with time (unstable) if any of the pairs of complex roots has a positive real part.

Figures 2-3.1c and d show, respectively, examples of stable and unstable oscillatory responses.

Equation 2-3.11 shows that the frequency of the sine wave is equal to the imaginary part of the roots, ω, in radians per unit time. The period of the oscillations is the time it takes for a complete cycle-that is, the time it takes for the argument of the sine wave, $(\omega t + \theta)$, to increase by 2π radians. Thus the period is

$$T = \frac{2\pi}{\omega} \tag{2-3.12}$$

---

[3] For the formulas derived here, the argument of the trigonometric functions must be in *radians,* which means that your calculator must be in radian mode when you invoke the trigonometric functions. The trigonometric functions of most computer languages (including FORTRAN, Pascal, C, and BASIC), and of spreadsheets, work in radians.

The SI unit for frequency is the hertz (Hz), which is the number of cycles per second, or the reciprocal of the period in seconds. Our formulas, however, require that the frequency be in radians per unit time.

Whereas the period of the oscillations is determined by ω, the imaginary part of the roots, the time it takes for the oscillations to die out is controlled by the real part of the roots, ρ. As with the real roots, the time it takes for the oscillations to decay to less than 1% of the initial amplitude, specifically $e^{-5} = 0.0067$, or 0.67%, is

$$t_s = \frac{-5}{\rho} \tag{2-3.13}$$

where t_s is approximately the 1% *settling time.*

Perhaps a better measure of the decay of the oscillations is the decay *ratio,* or the ratio at which the amplitude of the oscillations decays in one period. This number is

$$\text{Decay ratio} = e^{\rho T} = e^{2\pi\rho/\omega} \tag{2-3.14}$$

Final Steady-State Value

The only question left to be answered is determination of the final steady-state, or equilibrium, value of the output after the transients die out. For a final steady state to exist, the input variable, $X(t)$, must remain steady for some time. The easiest way to analyze the response to find this final steady-state value is to use the final value theorem of Laplace transforms and assume a step input, $X(t) = \text{Ax } u(t)$, or $X(s) = \Delta x/s$. Substituting into Eq. 2-3.5 and applying the final value theorem, Eq. 2-1.9, yields

$$AY = \lim_{s \to 0} s \left[\frac{b_m s^m + b_{m-1}s^{m-1} + \cdots + b_0}{a_n s^n + a_{n-1}s^{n-1} + \cdots + a_0} \right] \frac{\Delta x}{s} = \frac{b_0}{a_0} \Delta x \tag{2-3.15}$$

In this section, we have derived several formulas for computing important parameters of the output response. All these formulas except the formula for the final steady-state value, Eq. 2-3.15, are based on the roots of the denominator polynomial of the transfer function. None of these parameters depends on the values of the coefficients of the partial fractions expansion.

Table 2-3.1 summarizes the relationships we have established in this section between the output response and the roots of the denominator polynomial in the transfer function. The following example illustrates the application of the ideas discussed in this section.

EXAMPLE 2-3.1

Characterize the responses described by the differential equations that follow. Assume that the time is measured in minutes and the variables are deviations from initial steady-state conditions.

Table 2-3.1 Relationship Between the Laplace Transform $Y(s)$ and Its Inverse $Y(t)$

Denominator of $Y(s)$	Partial Fraction Term	Term of $Y(t)$
Umepeated real root	$\dfrac{A}{s - r}$	Ae^{rt}
Pair of complex conjugate roots where $D = \sqrt{B^2 + C^2}$ $\theta = \tan^{-1}\dfrac{B}{C}$	$\dfrac{Bs + C}{(s - p)^2 + \omega^2}$	$De^{pt}\sin(\omega t + \theta)$
Real root repeated m times	$\displaystyle\sum_{j=1}^{m}\dfrac{A_j}{(s - r)^j}$	$e^{rt}\displaystyle\sum_{j=1}^{m}\dfrac{A_j t^{j-1}}{(j - 1)!}$

(a) $30\,\dfrac{d^3Y(t)}{dt^3} + 43\,\dfrac{d^2Y(t)}{dt^2} + 14\,\dfrac{dY(t)}{dt} + Y(t) = 2.5X(t)$

Laplace transform and solve for $Y(s)$.

$$Y(s) = \frac{2.5}{30s^3 + 43s^2 + 14s + 1}X(s)$$

The roots of the denominator are -0.1, -0.333, and -1.0. Because the roots are all real and negative, the response is monotonic and stable. It is

$$Y(t) = A_1 e^{-0.1t} + A_2 e^{-0.333t} + A_3 e^{-t} + \text{(terms of } X)$$

The times required for these terms to decay to 0.67% of their initial value are, respectively, 50 (= $5/-0.1$), 15, and 5 minutes, so the first term dominates the response. For a step change in $X(t)$, the final steady-state value is $2.5/1.0 = 2.5$ times the amplitude of the step.

(b) $\dfrac{d^3Y(t)}{dt^3} + 5\,\dfrac{d^2Y(t)}{dt^2} + 11\,\dfrac{dY(t)}{dt} + 15Y(t) = 12X(t)$

Laplace transform and solve for $Y(s)$.

$$Y(s) = \frac{12}{s^3 + 5s^2 + 11s + 15}X(s)$$

The roots of the denominator are $-1 \pm i2$ and -3. Because there is a pair of complex conjugate roots, the response is oscillatory with a frequency of 2 radians/minute and a period of $2\pi/2 = 3.14$ minutes. The response is stable because the real part of the complex roots is negative and so is the real root. It is

$$Y(t) = De^{-t}\sin(2t + \theta) + A_3 e^{-3t} + \text{(terms of } X)$$

The sine wave decays to 0.67% of its initial amplitude in $- 5/- 1 = 5$ minutes, whereas the term with the real root decays in $- 5/- 3 = 1.67$ minutes. Therefore, the sine wave term is the dominant term. The decay ratio of the sine wave is

$$\text{Decay ratio} = e^{(-1)(3.14)} = 0.043$$

This means that the amplitude is reduced to 4.3% of its value during one cycle. The final steady-state change in Y(t) is 0.8 (= $12/15$) times the size of the sustained change in X(t).

EXAMPLE 2-3.2   *RESPONSE OF A PENDULUM*

Grandfather and cuckoo clocks use for a timing device a pendulum-that is, a weight suspended by a rod that can oscillate around its equilibrium value, which is the vertical position. Determine which parameters of the pendulum (weight, length, shape, and so on) determine its period of oscillation.

SOLUTION

A horizontal force balance on the pendulum, neglecting for the moment the resistance of the air and assuming that the angle of oscillation is small, results in the following differential   equation:

$$M \frac{d^2x(t)}{dt^2} = - Mg \frac{x(t)}{L} + f(t)$$

where $x(t)$ is the horizontal position of the weight in meters (m) from the equilibrium position, M is the mass of the weight in kilograms (kg), L is the length of the rod in m, $g = 9.8 \text{ m/s}^2$ is the acceleration of, gravity, and $f(t)$ is the force in newtons (N) required to start the pendulum in motion, usually a short pulse or impulse. Assuming the pendulum is originally at equilibrium, x(O) = 0, Laplace transform the equation and solve for $X(s)$ to obtain

$$X(s) = \left[\frac{1}{Ms^2 + \dfrac{Mg}{L}} \right] F(s)$$

The roots of the denominator are pure imaginary numbers.

$$r_{1,2} = + i\sqrt{\frac{g}{L}}$$

From Eq. 2-3.11, the response is

$$x(t) = D \sin\left(\sqrt{\frac{g}{L}}\, t + \theta\right) + [\text{terms of } f(t)]$$

That means that the pendulum will oscillate forever with a frequency that is independent of its weight and shape and is a function only of its length and the local acceleration of gravity. The period of oscillation is

$$T = 2\pi \sqrt{\frac{L}{g}}$$

Thus, if your grandfather clock is gaining time, you must lower the weight along the rod; if it is losing time, you must raise the weight. For example, a pendulum with a length of 1.0 m will have a period of $2\pi(1/9.8)^{0.5} = 2.0$ s.

Because of the resistance of the air, the pendulum does not oscillate forever. The weight and shape of the pendulum affect the air resistance. Clocks are equipped with a weight or spring mechanism to overcome the resistance of the air. This action could be incorporated into the external force $f(t)$. How could we incorporate the resistance of the air in the equation of motion to show that left to itself, a pendulum will eventually stop?

Note that we solved the preceding example without having to evaluate the coefficients of the partial fractions expansion or specify the input function.

2-3.3 Stability

Stability is the ability of the response to remain bound (remain within limits) when subjected to bound inputs. From the discussion in the preceding section, we conclude that the roots of the denominator of the transfer function of a process or device determine the stability of its response to input signals. That discussion can be summarized by the following condition of stability for linear systems: *A system is stable if all the roots of the denominator of its transfer function are either negative real numbers or complex numbers with negative real parts.* This condition of stability will be discussed further in Chapter 6, where we will see that stability is a very important constraint on the operation and tuning of feedback control loops.

2-4 RESPONSE OF FIRST-ORDER SYSTEMS

As we shall see in Chapter 3, the dynamic response of many processes and control system components can be represented by linear first-order differential equations. We refer to these processes as *first-order systems*. This section presents the response of first-order systems to three different types of input signals: a step function, a ramp, and a sine wave. Our objective is to learn how the parameters of first-order systems affect their response so that later we can infer the important characteristics of the response of

a system by simply examining its transfer function. First-order systems are also important because many higher-order systems can be treated as combinations of first-order systems in series and parallel.

Consider the linear first-order differential equation:

$$a_1 \frac{dy(t)}{dt} + a_0 y(t) = bx(t) + c \qquad \textbf{(2-4.1)}$$

where $y(t)$ is the output or dependent variable; $x(t)$ is the input variable; t is time, the independent variable; and the parameters a,, a,, b, and c are constant. We can write the equation at the initial steady state-that is, before any change in input $x(t)$ takes place.

$$a_0 y(0) = bx(0) + c \qquad (2\text{-}4.2)$$

Note that this equation establishes a relationship between the initial values of x and y. Subtracting Eq. 2-4.2 from Eq. 2-4.1 results in

$$a, \frac{dY(t)}{dt} + a_0 Y(t) = bX(t) \qquad (2\text{-}4.3)$$

where

$$Y(t) = y(t) - y(0)$$
$$X(t) = x(t) - x(0)$$

are the deviation variables, and we have made use of the fact that $dy(t)/dt = dY(t)/dt$, because they differ by only the constant bias y(O). Note that the constant c cancels out.

Equation 2-4.3 is the general linear first-order differential equation in terms of the deviations of the input and output variables from their initial steady-state values. It has three coefficients (a,, a,, and b), but without loss of generality, we can divide the equation by one of the three so that we can characterize the equation by just two parameters. In process control it is customary to divide by the coefficient of the output variable, a,, provided it is not zero. Such an operation results in the following equation, which we shall call the *standard form* of the linear first-order differential equation.

$$\boxed{\tau \frac{dY(t)}{dt} + Y(t) = KX(t)} \qquad (2\text{-}4.4)$$

where

$$\tau = \frac{a_1}{a_0} \text{ is the } \textit{time constant}$$

$$K = \frac{b}{a_0} \text{ is the } \textit{steady-state gain}$$

The reason for these names will become apparent as we develop the responses to various types of inputs. Note that in order for Eq. 2-4.4 to be dimensionally consistent, τ must have dimension of time, and K must have dimension of Y over dimension of X.

Any linear first-order differential equation can be transformed into the standard form of Eq. 2-4.4 as long as the dependent variable Y(t) appears in the equation. We can then obtain the transfer function of a first-order system by taking the Laplace transform of ·Eq. 2-4.4. To do this we apply the linearity property, Eq. 2-1.3, and the real differentiation theorem, Eq. 2- 1.4, noting that the initial condition of the deviation variable Y(t) is zero. The result is

$$\tau s Y(s) + Y(s) = K X(s) \tag{2-4.5}$$

Solving for Y(s) yields

$$Y(s) = \left[\frac{K}{\tau s + 1} \right] X(s) \tag{2-4.6}$$

The term in brackets is the transfer function of the first-order system in standard form. What is characteristic of this form is that the second term in the denominator is unity. When the transfer function is in this form, the numerator term is the gain, and the coefficient of s is the time constant.

The root of the denominator of the transfer function is $r = -1/\tau$. From what we learned in the preceding section, we can see that the response of a first-order system is monotonic (one real root) and that it is stable if its time constant is positive. Furthermore, the time required for the transients to be reduced to less than 1% of their initial value-specifically $e^{-5} = 0.0067$, or 0.67%—is $-5/r = 57$, or five times the time constant. The final steady-state change in the output, obtained by letting $s = 0$ in the transfer function, is K times the sustained change in input, which is precisely why K is the gain; the definition of the gain is the steady-state change in output divided by the sustained change in input.

Having established the general characteristics of the response of first-order systems, we next look at the actual responses to three typical input signals.

2-4.1 Step Response

To obtain the step response of magnitude Δx, we let $X(t) = Ax\,u(t)$, where $u(t)$ stands for the unit step function at time zero (see Example 2- 1.1 (a)). From Table 2- 1.1, the transform of the input is $X(s) = \Delta x/s$. Substitute this into Eq. 2-4.6 and expand in partial fractions to obtain

$$Y(s) = \frac{K}{\tau s + 1} \frac{Ax}{s} = -\frac{K \Delta x}{s + \dfrac{1}{\tau}} - \frac{K \Delta x}{s}$$
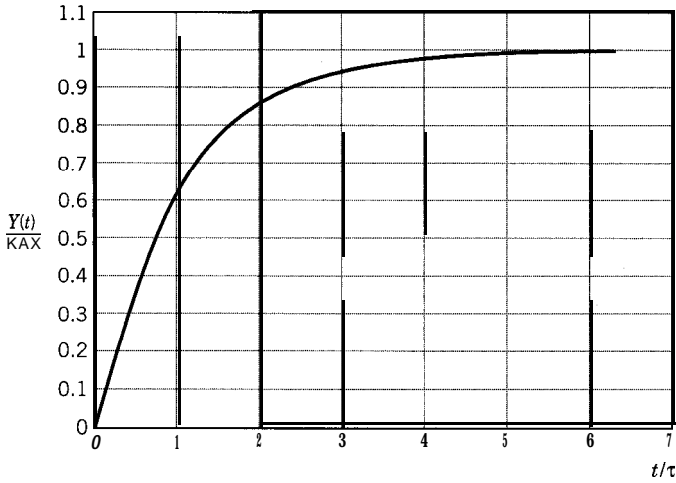
Figure 2-4.1 First-order step response.

Invert by matching entries in Table 2-1.1, with a $= 1/\tau$.

$$Y(t) = K\Delta x[u(t) - e^{-t/\tau}]$$ (2-4.7)

This is a very important result. Figure 2-4.1 gives a graph of the response, and Table 2-4.1 lists values of the normalized response for several values of t/τ. Note that the response starts at maximum rate of change right after the step is applied, and then the rate of change decreases such that the final steady-state value of $K\Delta x$ is approached exponentially. After one time constant the response reaches 63.2% of its final change, and in five time constants it reaches over 99% of the change. In other words, the response is essentially complete after five time constants.

Table 2-4.1 First-Order Step Response

$\dfrac{t}{\tau}$	$\dfrac{Y(t)}{K\Delta x}$
0	**0**
1.0	0.632
2.0	0.865
3.0	0.950
4.0	0.982
5.0	0.993
. . .	. . .
∞	1.000

2-4.2 Ramp Response

A ramp is a linear increase in the input with time starting at time zero. The input function is given by $X(t) = rt$, where r is the slope (or rate) of the ramp. From Table 2-1.1, the Laplace transform is $X(s) = r/s^2$. Substitute into Eq. 2-4.6 and expand in partial fractions.

$$Y(s) = \frac{K}{\tau s + 1} \frac{r}{s^2} = \frac{A_1}{s + \dfrac{1}{\tau}} + \frac{A_2}{s^2} + \frac{A_3}{s}$$

Coefficient A, is obtained from Eq. 2-2.9, and A, and A, from Eq. 2-2.13.

$$A_1 = \lim_{s \to -1/\tau} \left(s + \frac{1}{\tau} \right) \frac{Kr}{(\tau s + 1)s^2} = Kr\tau$$

$$A_2 = \lim_{s \to 0} s^2 \frac{Kr}{(\tau s + 1)s^2} = Kr$$

$$A, = \lim_{s \to 0} \frac{d}{ds} \left[\frac{Kr}{\tau s + 1} \right] = -Kr\tau$$

Substitute into the transform and invert by matching entries in Table 2-1.1.

$$Y(t) = Kr\tau e^{-t/\tau} + (Krt - Kr\tau)u(t)$$
$$= Kr\tau e^{-t/\tau} + Kr(t - \tau)u(t) \tag{2-4.8}$$

The ramp response, after the exponential term dies out in approximately five time constants, becomes a ramp with slope Kr and delayed by one time constant. To illustrate the way the output is delayed by exactly one time constant relative to the input, Fig. 2-4.2 superimposes plots of $X(t)$ and $Y(t)/K$ versus t; the two ramps are parallel because they both have slopes (or rates) of r. It is obvious from the plots that the output ramp "lags" the input ramp by one time constant τ. This is why systems represented by a first-order transfer function are also referred to as **first-order lags.**

2-4.3 Sinusoidal Response

To obtain the response of a first-order system to a sine wave, we let the input function be $X(t) = A \sin \omega t$, where A is the amplitude and ω is the frequency in radians/time. From Table 2-1.1, the Laplace transform is $X(s) = A\omega/(s^2 + \omega^2)$. Substitute into Eq. 2-4.6 and expand in partial fractions.

$$Y(s) = \frac{K}{\tau s + 1} \frac{A\omega}{s^2 + \omega^2} = \frac{A_1}{s + \dfrac{1}{\tau}} + \frac{\Lambda_2}{s - i\omega} + \frac{\Lambda_3}{s + i\omega}$$
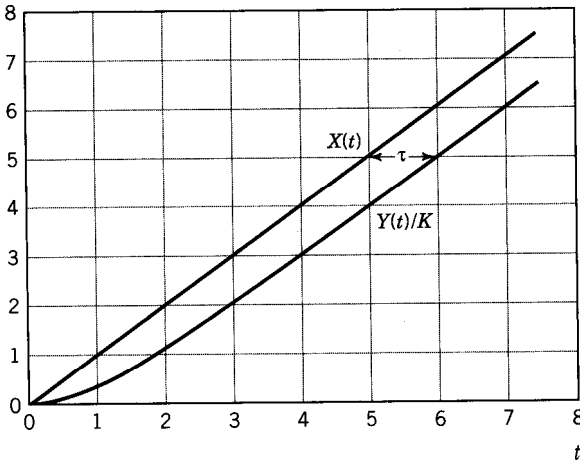
Figure   2-4.2 First-order response to a ramp. The normalized output lags the input by exactly one time constant.

where we have made use of $(s^2 + \omega^2) = (s - i\omega)(s + i\omega)$. The coefficients are obtained by using Eq. (2-2.9):

$$A, = \lim_{s \to -1/\tau} \left(s + \frac{1}{\tau} \right) \frac{KA\omega}{(\tau s + 1)(s^2 + \omega^2)} = \frac{KA\tau\omega}{1 + \tau^2\omega^2}$$

$$A, = \lim_{s \to i\omega} \frac{KA\omega}{(\tau s + 1)(s + i\omega)} = \frac{KA(-\tau\omega - i)}{2(1 + \tau^2\omega^2)}$$

$$A_3 = \lim_{s \to -i\omega} \frac{KA\omega}{(\tau s + 1)(s - i\omega)} = \frac{KA(-\tau\omega + i)}{2(1 + \tau^2\omega^2)}$$

We next substitute into the transform and invert. Then, using Eq. 2-3.11 with $\rho = 0$, after some nontrivial manipulations, we obtain

$$Y(t) = \frac{KA\omega\tau}{1 + \tau^2\omega^2} e^{-t/\tau} + \frac{KA}{\sqrt{1 + \tau^2\omega^2}} \sin(\omega t + \theta) \tag{2-4.9}$$

where $\theta = \tan^{-1}(-\omega\tau)$.

This sinusoidal response of a first-order system is plotted in Fig. 2-4.3. After the exponential term dies out in about five time constants, the response becomes a sine wave of the same frequency ω as the input sine wave. The amplitude of the output sine wave depends on the frequency. At very low frequencies, it is just the product of the steady-state gain and the amplitude of the input, but as the frequency of the input sine wave increases, the amplitude of the output sine wave decreases. There is also a phase shift, a lag, θ, which is a function of frequency. This dependence of the response on
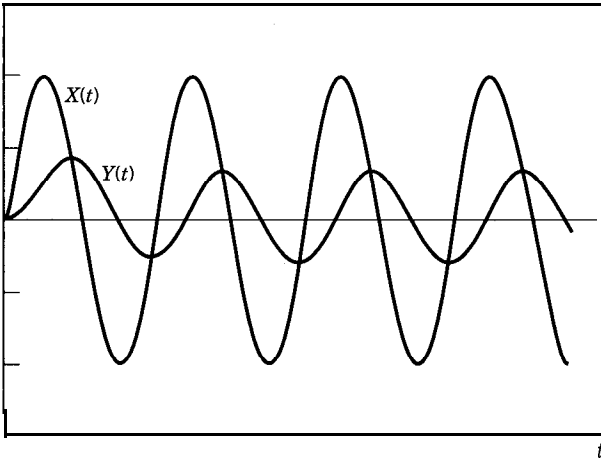
Figure 2-4.3 First-order response to a sine wave. The output
sine wave, Y(t), has the same frequency as the input, X(t).

the frequency of the input sine wave forms the basis for a method for analyzing process
dynamics and control systems known as *frequency response;* Chapter 9 presents this
topic in detail.

2-4.4   Response with Time Delay

As we shall see in Chapter 3, some process responses exhibit time delays (also known
as transportation lag or dead time). By the real translation theorem, the time delay will
modify the standard first-order transfer function of Eq. 2-4.6 as follows:

$$Y(s) = \left[\frac{Ke^{-st_0}}{\tau s + 1} \right] X(S)$$

(2-4.10)

The term in brackets is an important transfer function used to approximate the response
of higher-order processes. We call it a *first-order-plus-dead-time (FOPDT)* transfer
function.

The effect of the time delay on the three responses presented in this section is as
follows:

Step   Response

$$Y(t) = K \, \Delta x \, u(t - t_0)(1 - e^{-(t - t_0)/\tau})$$

(2-4.11)

where the presence of the factor $u(t - t_0)$ shows that the response is zero for $t < t_0$. A
plot of this response is shown in Fig. 2-4.4.

Ramp Response

$$Y(t) = u(t - t_0)[Kr\tau e^{-(t - t_0)/\tau} + Kr(t - t_0 - \tau)]$$

(2-4.12)
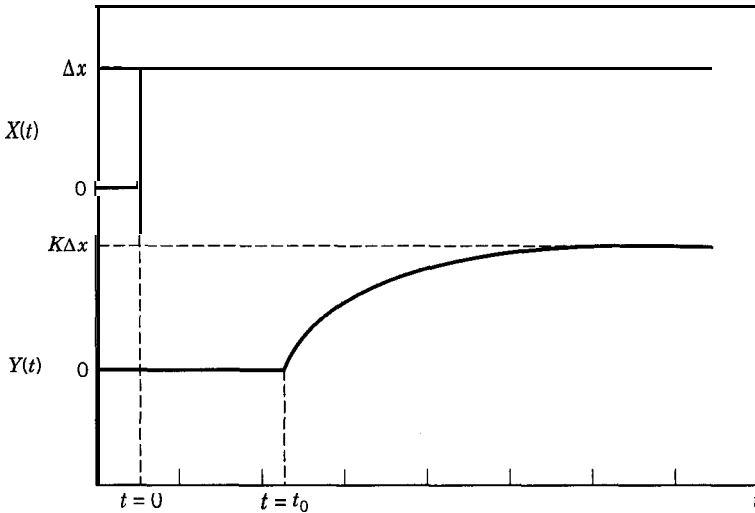
Figure 2-4.4 First-order step response with time delay t_0.

Note that the effect of the time delay in the long-term response is that the output ramp lags the input ramp by the sum of the time delay and the time constant.

Sinusoidal   Response

$$Y(t) = u(t - t_0)\left\{\frac{KA\omega\tau}{1 + \tau^2\omega^2}e^{-(t-t_0)/\tau} + \frac{KA}{\sqrt{1 + \tau^2\omega^2}}\sin[\omega(t - t_0) + \theta]\right\} \qquad \text{(2-4.13)}$$

The only effect of the time delay on the long-term response is to increase the phase lag by ωt_0. This increase in phase lag is proportional to the frequency of the input sine wave. The phase lag θ is the same as in Eq. 2-4.9.

2-4.5   Response of a Lead-Lag Unit

A device that is commonly used for dynamic compensation in feedforward controllers, known as a *lead-lag unit,* has the following transfer function:

$$Y(s) = \left[\frac{\tau_{ld}s + 1}{\tau_{lg}s + 1}\right] X(s) \qquad \text{(2-4.14)}$$

where τ_{ld} is the time constant of the lead term, and τ_{lg} is the time constant of the lag term. Note that a "lead" is a first-order term in the numerator, whereas a "lag," as we saw earlier, is a first-order term in the denominator. The step and ramp responses of lead-lag units are helpful in understanding how to tune them-that is, how to adjust the lead and lag time constants to achieve optimum dynamic compensation.
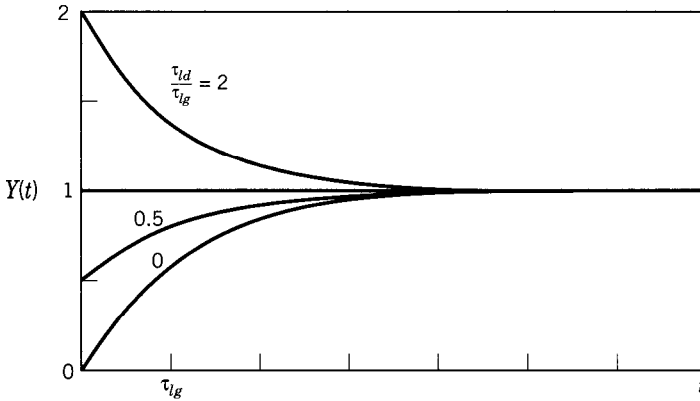
Figure 2-4.5 Lead-lag response to a unit step change.

Step Response

The response to a unit step response, $X(s) = 1/s$, is

$$Y(t) = u(t) + \left(\frac{\tau_{ld}}{\tau_{lg}} - 1\right)e^{-t/\tau_{lg}}$$

(2-4.15)

Figure 2-4.5 shows a plot of this response for various ratios of τ_{ld}/τ_{lg}. Note that the initial change in output is controlled by the ratio of the time constants, whereas the time required for the transient to die out is determined by the lag time constant (about five lag time constants). When the lead-to-lag ratio is greater than unity, the response overshoots its final steady state; when the ratio is less than unity, it undershoots it.

Ramp Response

The response to a ramp of unity rate, $X(s) = 1/s^2$, is

$$Y(t) = (\tau_{lg} - \tau_{ld})e^{-t/\tau_{lg}} + t + \tau_{ld} - \tau_{lg}$$

(2-4.16)

The ramp response is plotted in Fig. 2-4.6 for two cases, one in which the lead is greater than the lag, and the other in which the lag is greater than the lead, along with the input ramp. Note that after the transient term dies out, the response is a ramp that either leads or lags the input ramp by the difference between the lead and the lag, depending on which is longer. It is this response that gives the names *lead* and lug to the numerator and denominator terms of the transfer function.

A physical device cannot have more leads than lags, so in tuning lead-lag units, we must keep in mind that although the lead time constant can be set to zero, the lag time constant cannot be set to zero.

The application of lead-lag units to the dynamic compensation of feedforward controllers is discussed in detail in Section 12-2.3.
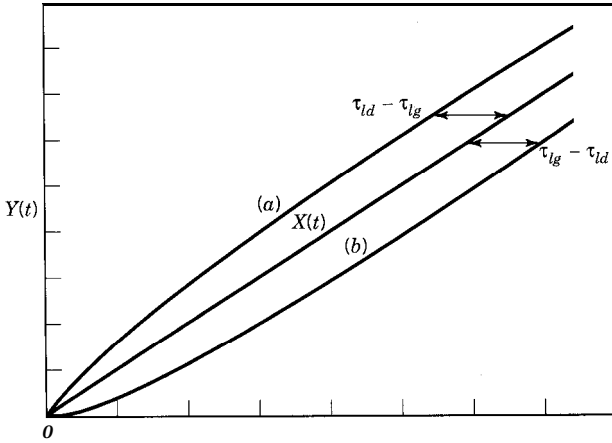
Figure 2-4.6 Lead-lag response to input ramp X(t). (a) Net lead, $\tau_{ld} > \tau_{lg}$. (b) Net lag, $\tau_{lg} > \tau_{ld}$.

Besides the responses presented in this section, there are other responses of interest, such as responses to impulse functions and pulses. These are proposed as exercises in the problems at the end of this chapter. Another interesting problem proposed as an exercise is the response of integrating processes, which are processes that do not contain the term a, in Eq. 2-4.3. Section 4-4.1 contains an example of an integrating process.

2-5   RESPONSE   OF   SECOND-ORDER   SYSTEMS

This section presents the response of linear second-order systems to the same three types of input signals for which the response of first-order processes was presented in the preceding section. We will see that the responses are quite different, depending on whether the roots of the denominator of the transfer function are real or a complex conjugate pair. When both roots are real, the response is said to be *overdumped;* when the roots are complex, the response is said to be *underdumped.* The response of over-damped processes is generalized to systems of order higher than two.

A linear second-order system is one represented by a linear second-order differential equation. A general form of such an equation is

$$a_2 \frac{d^2y(t)}{dt^2} + a_1 \frac{dy(t)}{dt} + a_0y(t) = bx(t) + c \qquad (2\text{-}5.1)$$

where $y(t)$ is the output variable, $x(t)$ is the input variable, and parameters $a_2, a_1, a_0, b,$ and c are constant. Assuming the initial conditions are at steady state, the equation at the initial conditions is

$$a_0y(0) = bx(0) + c \qquad (2\text{-}5.2)$$

Subtract Eq. 2-5.2 from Eq. 2-5.1 to obtain

$$a_2 \frac{d^2Y(t)}{dt^2} + a_1 \frac{dY(t)}{dt} + a_0 Y(t) = bX(t) \tag{2-5.3}$$

where

$$Y(t) = y(t) - y(0)$$
$$X(t) = x(t) - x(0)$$

are the deviation variables. By definition, the initial conditions of the deviation variables are zero. Note that the constant c cancels out.

The four parameters in Eq. 2-5.3 (*a,, a,, a,,* and *b*) can be reduced to three by dividing the entire equation by any of them, provided it is not zero. In process control we obtain what we will call the **standard form** of the second-order equation by dividing by coefficient *a,,* provided it is not zero. The resulting equation in the standard form is

$$\tau^2 \frac{dY(t)}{dt^2} + 2\zeta\tau \frac{dY(t)}{dt} + Y(t) = KX(t) \tag{2-5.4}$$

where

$\tau = \sqrt{\dfrac{a_2}{a_0}}$ is the characteristic time[4]

$\zeta = \dfrac{a_1}{2\tau a_0} = \dfrac{a_1}{2\sqrt{a_0 a_2}}$ is called the **damping ratio**

$K = b/a_0$ is the steady-state gain

The reason why the parameters τ and ζ are defined as they are will become evident momentarily. In the definition of the characteristic time, we assumed that a_2 and *a,* have the same sign; otherwise, the characteristic time would be an imaginary number and would lose its usefulness.

Next we take the Laplace transform of Eq. 2-5.4, apply the linearity property and the real differentiation theorem, and solve for the transform of the output to obtain

$$Y(s) = \left[\frac{K}{\tau^2 s^2 + 2\zeta\tau s + 1} \right] X(s) \tag{2-5.5}$$

where the term in the brackets is the second-order transfer function in standard

---

[4] In some textbooks on control, most notably those written by electrical engineers, the second-order response is characterized by the **natural frequency, ω_n,** which is defined as the reciprocal of the characteristic time τ.

form. To find the roots of the denominator polynomial, we apply the quadratic formula:

$$r_{1,2} = \frac{-\zeta \pm \sqrt{\zeta^2 - 1}}{\tau}$$

(2-5.6)

We see now that the damping ratio determines whether the roots are real or complex. If the absolute value of the damping ratio is 1 or greater than 1, then the roots are real, whereas if the damping ratio is less than unity, then the roots are a pair of complex conjugate numbers. When the damping ratio is unity, the two roots are equal to each other and equal to $-1/\tau$. When the damping ratio is zero, the real part of the complex roots is zero-that is, the roots are-pure imaginary numbers and equal to $\pm i/r$, where i is the square root of -1.

As we learned in Section 2-3, if the roots of the denominator of the transfer function are real numbers, the response is monotonic, whereas complex roots result in an oscillatory response. Furthermore, for the response to be stable, if the roots are real, both must be negative, or if they are complex, the real part must be negative. You are invited to verify that for the second-order transfer function of Eq. 2-5.5, the condition of stability is satisfied if and only if the damping ratio is positive. We can now see that the term **damping ratio** refers to the damping of oscillations. The behavior of the response is summarized as follows:

For	The Response Is
$\zeta \geq 1$	overdamped = monotonic and stable
$0 < \zeta < 1$	underdamped = oscillatory and stable
$\zeta = 0$	undamped = sustained oscillations
$-1 < \zeta < 0$	unstable = growing oscillations
$\zeta \leq -1$	run-away = monotonic unstable

The case of $\zeta = 1$ is sometimes called **critically damped,** but this is only the borderline case. Its response is monotonic and stable, just like the overdamped response.

For our purposes, we need consider only the two cases of real and complex roots, which we will call overdamped and underdamped, respectively. The following sections present the specific response equations for step, ramp, and sinusoidal inputs for both of these cases.

24.1   Overdamped   Responses

When the damping ratio is greater than unity, the roots given by Eq. 2-5.6 are real numbers. In this case it is better to factor the denominator of the transfer function into two first-order terms containing a time constant each, as follows:

$$\tau^2 s^2 + 2\zeta\tau s + 1 = \tau^2(s - r_1)(s - r_2)$$

$$= (\tau_{e1}s + 1)(\tau_{e2}s + 1)$$

(2-5.7)

where τ_{e1} and τ_{e2} are the *effective time constants,* defined as the negative reciprocals of the roots. For the second-order system, from Eq. 2-5.6, the effective time constants are

$$\tau_{e1} = -\frac{1}{r_1} = \frac{\tau}{\zeta - \sqrt{\zeta^2 - 1}}$$

$$\tau_{e2} = -\frac{1}{r_2} = \frac{\tau}{\zeta + \sqrt{\zeta^2 - 1}}$$

(2-5.8)

Next substitute Eq. 2-5.7 into Eq. 2-5.5 to obtain the transfer function in terms of the effective time constants.

$$Y(s) = \left[\frac{K}{(\tau_{e1}s + 1)(\tau_{e2}s + 1)} \right] X(s)$$

(24.9)

This is a more convenient transfer function for representing second-order systems when the roots are real numbers. We will use it to develop the various responses.

Step Response

As in the preceding section, we assume the input is a step change of magnitude Ax. Substitute then $X(s) = \Delta x/s$ into Eq. 2-5.9, and expand in partial fractions.

$$Y(s) = \frac{K}{(\tau_{e1}s + 1)(\tau_{e2}s + 1)} \frac{\Delta x}{s}$$

$$= \frac{A_1}{s + \dfrac{1}{\tau_{e1}}} + \frac{A_2}{s + \dfrac{1}{\tau_{e2}}} + \frac{A_3}{s}$$

Evaluate the coefficients and invert to obtain the output response.

$$Y(t) = K \Delta x \left[u(t) - \frac{\tau_{e1}}{\tau_{e1} - \tau_{e2}} e^{-t/\tau_{e1}} - \frac{\tau_{e2}}{\tau_{e2} - \tau_{e1}} e^{-t/\tau_{e2}} \right]$$

(2-5.10)

For the critically damped case, $\zeta = 1$, the two roots are equal to each other, $\tau_{e1} = \tau_{e2} = \tau$, and the response is given by

$$Y(t) = K \Delta x \left[u(t) - \left(\frac{t}{\tau} + 1 \right) e^{-t/\tau} \right]$$

(2-5.11)

Figure 2-5.1 shows two typical step responses, one overdamped and the other critically damped. Both of them are monotonic (non-oscillatory). Note that the initial rate of
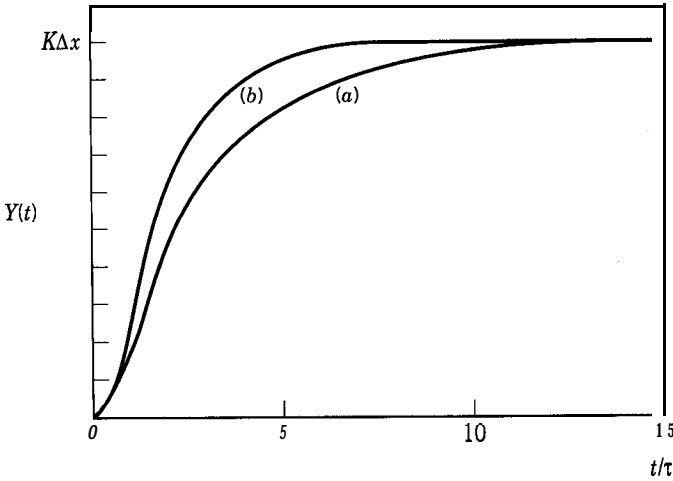
Figure 2-5.1 Second-order step responses. (a) Overdamped ($\zeta =$ 1.5). (b) Critically damped ($\zeta = 1.0$).

change of the response is zero and that it then increases to a maximum and finally decreases to approach exponentially its final steady-state change of $K \Delta x$. This differs from the first-order step response of Fig. 2-4.1, in which the maximum rate of change occurred right after the step change was applied (at time zero).

The S-shaped step responses shown in Fig. 2-5.1 are characteristic of many processes.

Ramp Response

We obtain the response to a ramp of rate r, $X(t) = rt$, by substituting its transform, $X(s) = r/s^2$, into Eq. 2-5.9, expanding in partial fractions, and inverting. When the two time constants are different, the ramp response is

$$Y(t) = Kr\left[\frac{\tau_{e1}^2}{\tau_{e1} - \tau_{e2}} e^{-t/\tau_{e1}} + \frac{\tau_{e2}^2}{\tau_{e2} - \tau_{e1}} e^{-t/\tau_{e2}} + t - (\tau_{e1} + \tau_{e2})\right] \qquad (2\text{-}5.12)$$

When the two time constants are equal to τ, the ramp response is

$$Y(t) = Kr[(t + 2\tau)e^{-t/\tau} + t - 2\tau] \qquad (2\text{-}5.13)$$

The important characteristic that these responses have in common is that after the exponential terms die out, the response becomes a ramp of rate Kr. This output ramp lags the input ramp by the sum of the two timeconstants. We can extend this result to higher-order systems if all the roots of the denominator are negative real numbers. The output ramp for an nth-order system lags the input ramp by the sum of all n effective

time constants, where the effective time constants are defined as the negative reciprocals of the roots.

Sinusoidal   Response

To obtain the response to a sine wave of amplitude A and frequency ω radians/time, $X(t) = A \sin \boldsymbol{wt}$, we substitute its transform, $X(s) = A\omega/(s^2 + \omega^2)$, into Eq. 2-5.6, expand in partial fractions, and invert. Chapter 9 presents a formal procedure for carrying out this operation. We present here the resulting response.

$$Y(t) = A_1 e^{-t/\tau_{e1}} + A_2 e^{-t/\tau_{e2}} + \frac{KA}{\sqrt{1 + \tau_{e1}^2 \omega^2}\, \sqrt{1 + \tau_{e2}^2 \omega^2}} \sin(\omega t + \theta) \qquad \textbf{(2-5.14)}$$

where

$$\theta = \tan^{-1}(-\omega\tau_{e1}) + \tan^{-1}(-\omega\tau_{e2})$$

Because the exponential terms die out and the sinusoidal term doesn't, the coefficients of the exponential terms are not important. For the sinusoidal response, the two important characteristics are that the amplitude of the output sine wave decreases as the frequency of the input sine wave increases and that the phase angle θ becomes more negative as the frequency increases.

It is even more interesting to note that the effect of the two time lags on the amplitude is multiplicative. This means that the reduction in the output amplitude is the product of the reductions that each lag would cause if it were acting alone. Similarly, the effect of the two lags on the phase angle is additive; the effect is the sum of the effects that each individual lag would cause if it were alone. This result can be extended to an nth-order system if all the roots of the denominator of the transfer function are negative real numbers. The reduction in the amplitude of the output wave is the product of the reductions that each of the n lags would cause if acting separately. Similarly, the phase angle is the sum of the phase angles that each of the n lags would separately cause. Chapter 9 presents this concept in more detail.

All of the response equations presented in this section apply for both the damping ratio being greater than or equal to unity and less than or equal to minus unity. The difference is that for positive damping ratios, both effective time constants are positive (both roots are negative) and the response is stable. On the other hand, when the damping ratio is negative, the effective time constants are negative (roots are positive) and the response is monotonically unstable-that is, the output runs away from its initial condition exponentially.

2-5.2 Underdamped Responses

The study of underdamped or oscillatory responses is important because it is the most common response of feedback control systems. Many common devices, such as pendulums, playground swings, yo-yo's, car suspension systems, and doors at department stores, also exhibit oscillatory behavior.

Second-order systems represented by Eq. 2-5.4 are underdamped when the damping ratio is between − 1 and + 1. We can see from Eq. 2-5.6 that the roots of the denominator of the transfer function form a complex conjugate pair.

$$r_{1,2} = \frac{-\zeta \pm \sqrt{-1(1 - \zeta^2)}}{\tau} = -\frac{\zeta}{\tau} \pm i\frac{\sqrt{1 - \zeta^2}}{\tau} \qquad \textbf{(2-5.15)}$$

From what we learned in Section 2-3, we know that these roots result in a response containing a sine wave with the frequency equal to the imaginary part and a decay rate equal to the real part. If the damping ratio is positive, $0 < \zeta < 1$, the amplitude of the oscillations decays with time and the response is stable, whereas for a negative damping ratio, the amplitude increases with time and the response is unstable; for a damping ratio of zero, the oscillations are sustained and the response is said to be **undamped.** Having looked at the generalities of the response, let us next present the specific responses to different input signals.

Step   Response

To obtain the step response, let the input signal be a step of magnitude Δx, $X(t) = \Delta x$ u(t), and substitute its Laplace transform, $X(s) = \Delta x/s$, into Eq. 2-5.5. After expansion in partial fractions, making use of Eq. 2-3.11, obtain the response

$$Y(t) = K\,\Delta x \left[u(t) - \frac{1}{\sqrt{1 - \zeta^2}}\, e^{-(\zeta/\tau)t}\, \sin(\psi t + \phi) \right] \qquad \textbf{(2-5.16)}$$

where

$\psi = \dfrac{\sqrt{1 - \zeta^2}}{\tau}$ is the frequency in radians/time

$\phi = \tan^{-1} \dfrac{\sqrt{1 - \zeta^2}}{\zeta}$ is the phase angle in radians

Figure 2-5.2 presents a plot of this response. Note that just as for the overdamped responses of Fig. 2-5.1, the maximum rate of change does not occur right after the step change is applied, as it does in the first-order response. In fact, the initial rate of change is zero. Unlike the overdamped response, the underdamped response oscillates around its final steady state, $K\Delta x$, which is also the steady state for the other step responses.

The underdamped step response is so important that it has been characterized by several terms. As for any sine wave, the period of oscillation is the time it takes to complete an entire cycle or 2π radians:

$$T = \frac{2\pi}{\psi} = \frac{2\pi\tau}{\sqrt{1 - \zeta^2}} \qquad \textbf{(2-5.17)}$$

As shown in Fig. 2-5.2, the period can be measured in the response by the time between two successive peaks in the same direction. Other term definitions follow.
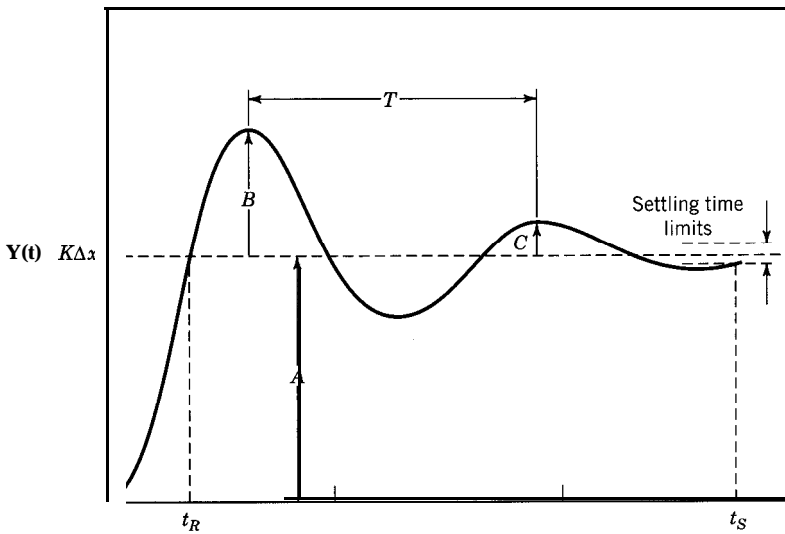
Figure 24.2 Second-order underdamped step response ($\zeta = 0.215$).

Although the SI unit for frequency is the hertz (Hz), which is the reciprocal of the period T in seconds, or the number of cycles in one second, the formulas presented here require that the frequencies be in radians per unit time; they also require that the angles be in radians and not in degrees or other units.

Decay Ratio. The decay ratio is the ratio by which the amplitude of the sine wave is reduced during one complete cycle. It is defined as the ratio of two successive peaks in the same direction, *C/B* in Fig. 2-5.2.

$$\text{Decay ratio} = e^{-(\zeta/\tau)T} = e^{-2\pi\zeta/\sqrt{1-\zeta^2}} \qquad (2\text{-}5.18)$$

The decay ratio is an important term, because it serves as a criterion for establishing satisfactory response of feedback controllers.

Rise Time. This is the time it takes for the response to first reach its final steady-state value, t_R in Fig. 2-5.2. It can be approximated as one-fourth of the period T.

Settling Time. This is the time it takes for the response to come within some prescribed band of the final steady-state value and remain in this band. Typical band limits are $\pm 5\%$, $\pm 3\%$, and $\pm 1\%$ of the total change. The settling time is t_s in Fig. 2-5.2. As discussed in Section 2-3, the real part of the roots of the denominator of the transfer function controls the settling time. For band limits of $\pm 1\%$, it is approximately $5\ \tau/\zeta$.

Overshoot. The overshoot is the fraction (or percent) of the final steady-state change by which the first peak exceeds this change. On the assumption that the first peak occurs

Table 2-5.1 Second-Order Underdamped Step Response

Damping Ratio (ζ)	Decay Ratio	Overshoot, %	$\dfrac{t_R}{\tau}$	$\dfrac{t_S}{\tau}, \pm1\%$
1.0	0	0	—	7.0
0.707	1/500	4.3	2.2	7.1
0.344	1/10	29.3	1.8	14.5
0.215	1/4	50.0	1.6	23.2
0	1/1	100	$\pi/2$	∞

approximately half a cycle from the application of the step change, it is

$$\text{Overshoot} = e^{-(\zeta/\tau)T/2} = e^{-\pi\zeta/\sqrt{1-\zeta^2}} \tag{2-5.19}$$

Figure 2-5.2 shows how to determine the overshoot from a plot of the step response; it is the ratio B/A, where $A = K\Delta x$.

You can see from the number of terms presented so far that the step response of underdamped systems is an important topic. Table 2-5.1 shows the numerical values of some of these terms for several values of the damping ratio. Figure 2-5.3 contains plots of underdamped step responses for the same values of the damping ratio.

Ramp Response

To obtain the response of an underdamped second-order system to a ramp of rate r, $X(t) = rt$, substitute its transform, $X(s) = r/s^2$, into Eq. 2-5.5. Then expand in partial
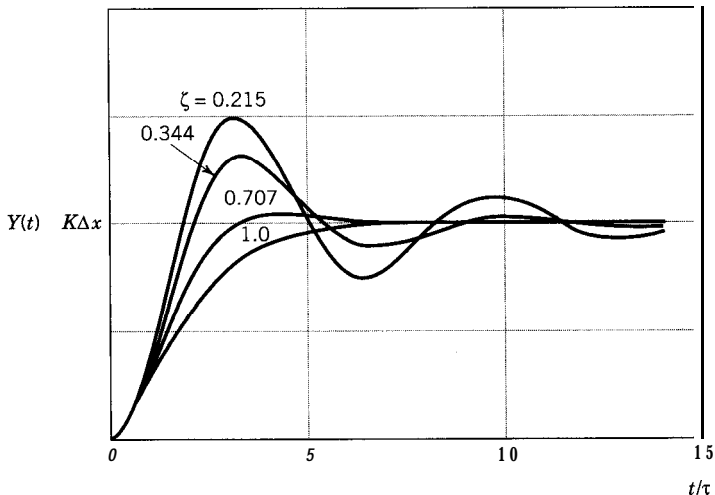
Figure 2-5.3 Effect of damping ratio on the second-order underdamped step response.

fractions, evaluate the coefficients, and invert. The result is

$$Y(t) = Kr\left[\frac{\tau}{\sqrt{1-\zeta^2}}\,e^{-(\zeta/\tau)t}\sin(\psi t + \phi) + t - 2\zeta\tau\right] \qquad (2\text{-}5.20)$$

where

$$\phi = \tan^{-1}\left(\frac{2\zeta\sqrt{1-\zeta^2}}{2\zeta^2 - 1}\right)$$

and ψ is the frequency, which is the same as for the step change. The important characteristic of this response is that after the sinusoidal term dies out, the output becomes a ramp of rate Kr that lags the input ramp by a time that decreases as the damping ratio decreases. For the undamped response, $\zeta = 0$, the output response is a sustained oscillation around the input ramp.

Sinusoidal Response

To obtain the underdamped response of the second-order system to a sine wave of amplitude A and frequency ω, $A \sin \omega t$, substitute its transform, $X(s) = A\omega/(s^2 + \omega^2)$, into Eq. 2-5.5. Then expand into partial fractions, evaluate the coefficients, and invert. The result is

$$Y(t) = KADe^{-(\zeta/\tau)t}\sin(\psi t + \phi) + \frac{KA}{\sqrt{(1 - \omega^2\tau^2)^2 + (2\zeta\tau\omega)^2}}\sin(\omega t + \theta) \qquad (2\text{-}5.21)$$

where

$$\theta = -\tan^{-1}\left(+\frac{2\zeta\tau\omega}{- \omega^2\tau^2}\right)$$

The amplitude D and phase angle ϕ in the first sine term are not important, because this is the term that decays with time. After this first term decays, the output response is a sine wave with frequency equal to the frequency of the input signal. The amplitude and the phase angle of the output are functions of the frequency. An interesting effect in the sinusoidal response of underdamped system is what happens when the input frequency is the system **resonant frequency,** equal to $1/\tau$. According to Eq. 2-5.21, at the resonant frequency the ratio of the amplitude of the output sine wave to that of the input is $K/2\zeta\tau\omega$—that is, it is inversely proportional to the damping ratio. This phenomenon, known as **resonance,** can result in very high output amplitudes when the damping ratio is small. In the 1940s a bridge at Tacoma Narrows, Washington, collapsed when the wind drove it at its resonant frequency.

2-5.3 Higher-Order Responses

The response of systems represented by differential equations of order higher than two can be thought of as a combination of first-order lags and second-order underdamped

responses. When all the roots of the denominator of the transfer function are real, an nth-order system becomes a combination of n first-order lags. We can easily extend the results for the second-order overdamped responses to higher-order overdamped responses. For example, consider the following nth-order overdamped system:

$$Y(s) = \left[\frac{K}{\prod_{k=1}^{n} (\tau_k s + 1)} \right] X(s)$$

(25.22)

where K is the gain and τ_k are the n effective time constants, or negative reciprocals of the n roots of the denominator polynomial. The response of this system to a step change of magnitude Ax, $X(s) = Ax/s$, if all the time constants are different from each other, is given by

$$Y(t) = K\,\Delta x \left[u(t) \quad \sum_{k=1}^{n} \frac{\tau_k^{n-1}}{\prod_{\substack{j=1 \\ j \neq k}}^{n} (\tau_k - \tau_j)} e^{-t/\tau_k} \right]$$

(2-5.23)

Note that Eq. 2-5.10 is a special case of this equation for $n = 2$. When all the n time constants are equal to each other, the step response is

$$Y(t) = K\,\Delta x \left[u(t) - e^{-t/\tau} \sum_{k=1}^{n} \frac{1}{(n-k)!} \left(\frac{t}{\tau} \right)^{n-k} \right]$$

(2-5.24)

where τ is the time constant that is repeated n times. Note that Eq. (2-5.11) is a special case of this equation for $n = 2$.

If the transfer function of the nth-order system contains lead terms,

$$Y(s) = \left[\frac{K \prod_{j=1}^{m} (\tau_{ld_j} s + 1)}{\prod_{k=1}^{n} (\tau_{lg_k} s + 1)} \right] X(s)$$

(2-5.25)

where $n \geq m$, then the step response for all the lag time constants being different from each other is

$$Y(t) = K\,\Delta x \left[u(t) - \sum_{k=1}^{n} \frac{\tau_{lg_k}^{n-m-1} \prod_{j=1}^{m} (\tau_{lg_k} - \tau_{ld_j})}{\prod_{\substack{j=1 \\ j \neq k}}^{n} (\tau_{lg_k} - \tau_{lg_j})} e^{-t/\tau_{lg_k}} \right]$$

(2-5.26)

The effect of the lead terms is to speed up the response if the lead time constants are positive or to slow it down if they are negative.

For higher-order underdamped systems, the second-order step response terms defined in this section also apply. However, the formulas presented to calculate the characteristic terms are valid only for estimating the contribution of individual pairs of complex conjugate roots to the overall response. The accuracy of the estimates of the overshoot, rise time, and decay ratio of the total response depends on how dominant is the pair of complex conjugate roots with respect to the other roots. Recall, from Section 2-3, that the dominant roots are those with the least negative real parts-that is, the terms of the response that take the longest to decay.

2-6 LINEARIZATION

A major difficulty in analyzing the dynamic response of many processes is that they are nonlinear-that is, they cannot be represented by linear differential equations. A linear differential equation consists of a sum of terms each of which contains no more than one variable or derivative, which must appear to the first power. In the preceding sections, we learned that the method of **Laplace** transforms allows us to relate the response characteristics of a wide variety of physical systems to the parameters of their transfer functions. Unfortunately, only linear systems can be analyzed by **Laplace** transforms. There is no comparable technique by which we can analyze the dynamics of a nonlinear system and generalize the results to represent similar physical systems.

This section presents the technique known as *linearization,* which is used to approximate the response of nonlinear systems with linear differential equations that can then be analyzed by **Laplace** transforms. The linear approximation to the nonlinear equations is valid for a region near some base point around which the linearization is made. To facilitate manipulation of the linearized equations, we will select the initial steady state as the base point for linearization and will use *deviation variables,* or *perturbation variables,* as defined in Section 2-3.1.

The following is a list of common nonlinear functions that appear in process dynamic models.

- Enthalpy, *H,* as a function of temperature, T:

$$H[T(t)] = H_0 + a_1T(t) + a_2T^2(t) + a_3T^3(t) + a_4T^4(t) \qquad \text{(2-6.1)}$$

where $H_0, a_1, a_2, a_3,$ and a_4 are constants.
- Antoine equation for the vapor pressure of a pure substance, p^o, as a function of temperature, T:

$$p^o[T(t)] = e^{A - B/[T(t) + C]} \qquad \text{(2-6.2)}$$

where *A, B,* and C, are constants.
- Equilibrium vapor mole fraction, y, as a function of liquid mole fraction, x:

$$y[x(t)] = \frac{\alpha x(t)}{1 + (a - 1)x(t)} \qquad \text{(2-6.3)}$$

where α is the relative volatility, usually assumed constant.

- Fluid flow, f, as a function of pressure drop, Ap:

$$f[\Delta p(t)] = k\sqrt{\Delta p(t)} \tag{2-6.4}$$

where k is a constant conductance coefficient.
- Radiation heat transfer rate, q, as a function of temperature, T:

$$q[T(t)] = \bullet \quad AT4(t) \tag{2-6.5}$$

where ϵ, σ, and A, are constants.
- Arrhenius equation for the dependence of reaction rate coefficient, k, on temperature, T:

$$k[T(t)] = k_0 e^{-E/RT(t)} \tag{2-6.6}$$

where k_0, E, and R, are constants.
- Reaction rate, r, as a function of temperature, T, and reactants concentration, c_A, c_B, . . . :

$$r[T(t), c_A(t), c_B(t), \ldots] = k[T(t)]c_A^a(t)c_B^b(t) \ldots \tag{2-6.7}$$

where $k[T(t)]$ is given by Eq. 2-6.6, and a and b are constant.

All of the foregoing nonlinear functions except the last one are functions of a single variable. Next, we will introduce the linearization procedure for functions of one variable and then extend it to functions of two or more variables.

2-6.1   Linearization of Functions of One Variable

Any function can be expanded in a Taylor series about a base point, as follows:

$$f[x(t)] = f(\bar{x}) + \frac{df}{dx}\bigg|_x [x(t) - X1] + \frac{1}{2!}\frac{d^2f}{dx^2}\bigg|_{\bar{x}} [x(t) \quad \bar{x}]^2 + \ldots \tag{2-6.8}$$

where $\bar{x}$ is the base value of x around which the function is expanded. The linearization of function $f[x(t)]$ consists of approximating it with only the first two terms of the Taylor series expansion:

$$\boxed{f[x(t)] \doteq f(\bar{x}) + \frac{df}{dx}\bigg|_{\bar{x}} [x(t) - \bar{x}]} \tag{2-6.9}$$

This is the basic linearization formula. Because $\bar{x}$ is a constant, the right-hand side of the equation is linear in the variable $x(t)$.

Figure 2-6.1 presents a graphical interpretation of the linearization formula, Eq.
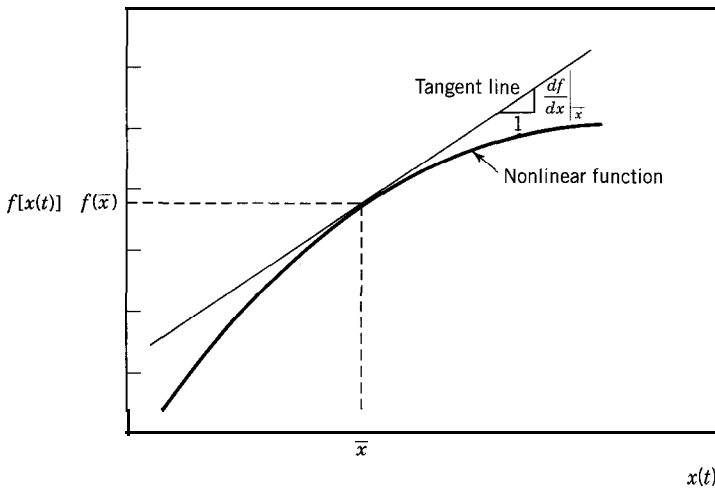
Figure 2-6.1 The linear approximation is the tangent to the nonlinear function at the base point $\bar{x}$.

2-6.9. The linear approximation is a straight line passing through the point $[\bar{x}, f(\bar{x})]$ with slope $df/dx|_{\bar{x}}$. This line is by definition the tangent to $f(x)$ at $\bar{x}$. Note that the difference between the nonlinear function and its linear approximation is small near the base point $\bar{x}$ and becomes larger the farther $x(t)$ is from $\bar{x}$. The width of the range in which the linear approximation is accurate depends on the function. Some functions are more curved than others and thus have a narrower range over which the linear approximation is accurate.

It is important to realize that what affects the parameters of the transfer function of a linearized system is the slope, $df/dx|_{\bar{x}}$, not the value of the function itself, $f(\bar{x})$. This will become obvious when we show how to apply the linearization technique to nonlinear differential equations. The following example illustrates the application of the linearization formula.

EXAMPLE 2-6.1

Linearize the Arrhenius equation, Eq. 2-6.6, for the temperature dependence of chemical reaction rate coefficients. For a reaction with a coefficient $k(\bar{T}) = 100$ s^{-1} and an energy of activation $E = 22,000$ kcal/kmole, estimate the error in the slope of the function in the range $\pm 10°C$ around $\bar{T} = 300°C$ (573 K).

SOLUTION

Apply the linearization formula, Eq. 2-6.9, to Eq. 2-6.6.

$$k[T(t)] \doteq k(\bar{T}) + \frac{dk}{dT}\bigg|_{\bar{T}} [T(t) - \bar{T}]$$

where

$$\frac{dk}{dT}\bigg|_{\bar{T}} = \frac{d}{dT}\left[k_0 e^{-E/RT(t)}\right]_{T=\bar{T}}$$

$$= k_0 e^{-E/R\bar{T}}\frac{E}{R\bar{T}^2} = k(\bar{T})\frac{E}{R\bar{T}^2}$$

For the numerical values given, with $R = 1.987$ kcal/kmole-K (ideal gas law constant), the base value of the slope is

$$\frac{dk}{dT}\bigg|_{300°C} = (100)\frac{22,000}{(1.987)(300 + 273)^2} = 3.37\frac{s^{-1}}{°C}$$

and the linear approximation of the function is

$$3k[T(t)] \doteq 100 + 3.37[T(t) - \bar{T}]$$

In the range 290 to 310°C, the actual function and slope are

At $T = 290°C$, $k(T) = 70.95$ s^{-1}, $dk/dT\,|_{\bar{T}} = 2.48$ s^{-1}/°C
At $T = 310°C$, $k(T) = 139.3$ s^{-1}, $dk/dT\,|_{\bar{T}} = 4.54$ s^{-1}/°C

In comparison, the linear approximation of the function predicts $k(290°C) = 100 + 3.37(290 - 300) = 66.3$ s^{-1}, which is -6.6% in error, and $k(310°C) = 133.7$ s^{-1}, which is -4% in error. As for the slope, it varies from 2.48 to 4.54 s^{-1}/°C, which is from 73.6% to 134.7% of the linear approximation, 3.37 s^{-1}/°C.

This example shows that for the Arrhenius formula, the linear approximation is accurate over a wider range for the function than it is for its slope. Unfortunately, it is the slope that affects the parameters of the transfer function. However, the error of ± 35% in the parameters is usually satisfactory for many control system calculations.

2-6.2   Linearization of Functions of Two or More Variables

We can use the Taylor series expansion to derive the linearization formula for functions of two or more variables, just as we did to derive Eq. 2-6.9. Here we keep the first partial derivative term for each of the variables. The resulting linear approximation is

$$\begin{aligned}f[x_1(t), x_2(t), \ldots] &\doteq f(\bar{x}_1, \bar{x}_2, \ldots)\\[4pt]&+ \frac{\overline{\partial f}}{\partial x_1}[x_1(t) - \bar{x}_1]\\[4pt]&+ \frac{\overline{\partial f}}{\partial x_2}[x_2(t) - \bar{x}_2] + \cdots\end{aligned}$$

(2-6.10)

where $\dfrac{\overline{\partial f}}{\partial x_k} = \dfrac{\partial f}{\partial x_k}\Big|_{(\bar{x}_1, \bar{x}_2, \ldots)}$, and $x_1, x_2, \ldots$, are the base values of each variable.
Recall, from calculus, that the partial derivative is the change with respect to one variable when all other variables are kept constant. The following examples illustrate the use of Eq. 2-6.10, the linearization formula for functions of more than one variable.

EXAMPLE 2-6.2

A nonlinear function that commonly occurs in component and energy balances is the product of two variables (flow and composition, flow and specific enthalpy, and so on). For this function, the linearization is so simple that it is sometimes difficult to grasp. As an example of this simple function, consider the area a of a rectangle as a function of its width, w, and its height, h:

$$a[w(t), h(t)] = w(t)h(t)$$

Linearization, using Eq. 2-6.10, results in

$$a[w(t), h(t)] \doteq a(\bar{w}, \bar{h}) + \frac{\partial a}{\partial w}\Big|_{(\bar{w}, \bar{h})} [w(t) - \bar{w}]$$

$$+ \frac{\partial a}{\partial h}\Big|_{(\bar{w}, \bar{h})} [h(t) - \bar{h}]$$

$$a[w(t), h(t)] \doteq a(\bar{w}, \bar{h}) + \bar{h}[w(t) - \bar{w}] + \bar{w}[h(t) - \bar{h}]$$

Figure 2-6.2 shows a graphical representation of the area. The figure shows that the error in the approximation is the area of the small rectangle in the upper right-hand corner, $[w(t) - \bar{w}][h(t) - h]$. This error is small for small relative increments in the width and height. For example, assume that the base values are $\bar{w} = 2$ m and $\bar{h} = 1$ m
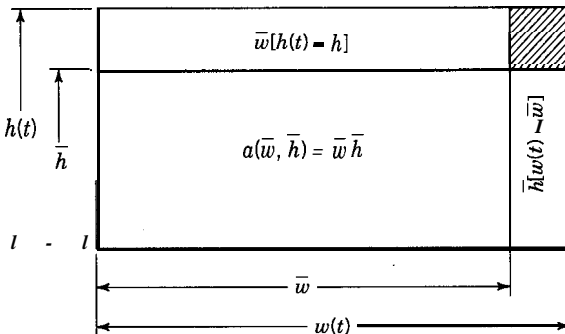
Figure 2-6.2 The cross-hatched area is the error of the linear approximation to the function $a[w(t), h(t)] = w(t)h(t)$.

and that they increment to 2.2 and 1.1 m, respectively. Then the error between the actual area, 2.42 m², and the approximation, $2.0 + 1(0.2) + 2(0.1) = 2.40$ m², is the area of the rectangle, $(0.2)(0.1) = 0.02$ m². Again the accuracy in the function is good (- 0.8% error), but each slope is off by 10%.

EXAMPLE 2-6.3

The density of an ideal gas is a function of pressure and temperature:

$$\rho[p(t), T(t)] = \frac{Mp(t)}{RT(t)}$$

where M is the molecular weight, p(t) is the absolute pressure, T(t) is the absolute temperature, and R is the ideal gas law constant. Obtain the linear approximation and evaluate it for air $(M = 29)$, at 300 K and atmospheric pressure, 101.3 kPa. In these (SI) units, the ideal gas law constant is 8.314 kPa-m³/kmole-K.

SOLUTION

Application of Eq. 2-6.10 results in

$$\rho[p(t), T(t)] \doteq \rho(\bar{p}, \bar{T}) + \overline{\frac{\partial \rho}{\partial p}}[p(t) - \bar{p}] + \overline{\frac{\partial \rho}{\partial T}}[T(t) - \bar{T}]$$

where the partial derivatives are given by

$$\overline{\frac{\partial \rho}{\partial p}} = \frac{\partial}{\partial p}\left[\frac{Mp(t)}{RT(t)}\right]_{(\bar{p},\bar{T})} = \frac{M}{R\bar{T}}$$

$$\overline{\frac{\partial \rho}{\partial T}} = \frac{\partial}{\partial T}\left[\frac{Mp(t)}{RT(t)}\right]_{(\bar{p},\bar{T})} = -\frac{M\bar{p}}{R\bar{T}^2}$$

The linear approximation is then

$$\rho[p(t), T(t)] = \frac{M\bar{p}}{R\bar{T}} + \frac{M}{R\bar{T}}[p(t) - \bar{p}] - \frac{M\bar{p}}{R\bar{T}^2}[T(t) - \bar{T}]$$

Numerically, the values are

$$\rho[p(t), T(t)] = 1.178 + 0.01163[p(t) - \bar{p}] - 0.00393[T(t) - \bar{T}]$$

where ρ is in kilograms per cubic meter (kg/m³), p is in kilopascals (kPa), and T is in kelvins (K).

One thing to observe in the preceding examples is that in the linear approximations, the coefficient of each variable is constant. Thus, although we may sometimes show them as functions of the base values of the variables, the equations would not be linear if these base values were not assumed constant. Whenever we can express the coefficients as functions of the base values, we can calculate them at different base values. Let us next apply our newly acquired skill to the linearization of nonlinear differential equations.

2-6.3   Linearization of Differential Equations

The following procedure to linearize nonlinear differential equations assumes that the equations can be expressed as first-order equations. This is not a significant restriction because, as we shall see in Chapters 3 and 4, creating process dynamic models usually consists of developing a set of first-order differential equations. Only after linearization and **Laplace** transformation should the first-order equations be combined to form higher-order equations. This is because it is much easier to manipulate the algebraic equations relating the transforms than the original differential equations. Most computer simulation programs also require that the differential equations be first-order.

In the following procedure we assume, as we have done so far, that the initial conditions are at steady state. Furthermore, we select the base point for the linearization as the initial steady state, because this greatly simplifies the linearized equations.

Consider the following first-order differential equation with one input:

$$\frac{dy(t)}{dt} = g[x(t),\ y(t)] + b \qquad\qquad (2\text{-}6.11)$$

where $g[x(t),y(t)]$ is a nonlinear function of the input variable, $x(t)$, and the output variable, $y(t)$, and b is a constant. At the initial steady-state conditions, Eq. 2-6.11 can be written as

$$0 = g(x,y) + b \qquad\qquad (2\text{-}6.12)$$

where we have chosen the base point for the linearization to be the initial conditions, $\bar{x} = x(O)$, $\bar{y} = y(O)$. Note that the time derivative is zero because of the initial steady-state assumption. Subtract Eq. 2-6.12 from Eq. 2-6.11 to obtain

$$\frac{dy(t)}{dt} = g[x(t),\ y(t)] - g(\bar{x},\bar{y}) \qquad\qquad (2\text{-}6.13)$$

If we now approximate Eq. 2-6.13 using the formula for the linearization of multivariable functions, Eq. 2-6.10, the result is

$$\frac{dy(t)}{dt} \doteq \frac{\partial g}{\partial x}\bigg|_{(\bar{x},\bar{y})} [x(t) - \bar{x}] + \frac{\partial g}{\partial y}\bigg|_{(\bar{x},\bar{y})} [y(t) - \bar{y}] \qquad\qquad (2\text{-}6.14)$$

The terms in brackets are the deviation variables that were introduced in Section 2-3.1,

because $\bar{x}$ and $\bar{y}$ are the initial conditions. Substitute the deviation variables, $X(t) = x(t) \quad \bar{x}$, $Y(t) = y(t) - \bar{y}$, to obtain

$$\frac{dY(t)}{dt} \doteq a_1 X(t) + a_2 Y(t) \qquad \text{(2-6.15)}$$

where $a_1 = \partial g/\partial x|_{(\bar{x},\bar{y})}$ and $a_2 = \partial g/\partial y|_{(\bar{x},\bar{y})}$.

We can generally apply the preceding procedure to any equation for any number of variables. Equation 2-6.1.5 is the linear approximation of Eq. 2-6.11; compare the two and note that

- The constant b in Eq. 2-6.11 drops out. There should not be any constant terms in the equation relating deviation variables.
- The linearized equation, Eq. 2-6.15, replaces the right-hand side of the original nonlinear equation by a sum of terms each of which consists of a constant times a deviation variable for each variable appearing on the right-hand side of the original differential equation.
- The initial condition of the deviation variable is zero, $Y(O) = y(0) \quad y(0) = 0$.

We will next illustrate the linearization procedure with an example.

EXAMPLE   2-6.4

The following differential equation results from a reactant mass balance in a stirred tank reactor (the complete model of the reactor is developed in Chapter 4).

$$\frac{dc_A(t)}{dt} = \frac{1}{V} f(t)c_{Ai}(t) - \frac{1}{V} f(t)c_A(t) - k[T(t)]c_A(t)$$

where $k[T(t)]$ is the Arrhenius dependence of the reaction rate on temperature, Eq. 2-6.6. We linearized this function in Example 2-6.1. We assume that V, the reactor volume, is constant. The input variables are $f(t)$, the reactants flow, and $c_{Ai}(t)$, the inlet reactant concentration, and the output variables are $T(t)$, the reactor temperature, and $c,(t)$, the reactant concentration. Obtain the linear approximation to the equation and the expressions for the time constant and gains of the transfer function of the reactant concentration.

SOLUTION

Each term of the equation is nonlinear: the first two terms both consist of the product of two variables, and the third term is a product of two variables one of which is the nonlinear Arrhenius function. We could linearize each of these terms in turn, substitute the linear approximations into the equation, and then subtract the initial steady-state equation and express in terms of deviation variables. However, we can simplify the algebraic manipulations by the following procedure:

Define the right-hand side of the equation as a nonlinear function of the four variables that appear in it.

$$\frac{dc_A(t)}{dt} = g[f(t), c_{Ai}(t), T(t), c_A(t)]$$

$$= \frac{1}{V}f(t)c_{Ai}(t) - \frac{1}{V}f(t)c_A(t) - k[T(t)]c_A(t)$$

Then, by comparison with Eq. 2-6.11 and its linear approximation, Eq. 2-6.15, we know the linear approximation is of the form

$$\frac{dC_A(t)}{dt} = a_1 F(t) + a_2 C_{Ai}(t) + a_3 \Gamma(t) + a_4 C_A(t)$$

where $C_A(t) = c_A(t) - \bar{c}_A$, $F(t) = f(t) - \bar{f}$, $C_{Ai}(t) = c_{Ai}(t) - \bar{c}_{Ai}$, and $\Gamma(t) = T(t) - \bar{T}$ are the deviation variables, and the constants are evaluated by taking the partials of function g:

$$a_1 = \frac{\overline{\partial g}}{\partial f} = \frac{\bar{c}_{Ai} - \bar{c}_A}{V} \qquad\qquad a_2 = \frac{\overline{\partial g}}{\partial c_{Ai}} = \frac{\bar{f}}{V}$$

$$a_3 = \frac{\overline{\partial g}}{\partial T} = -k(\bar{T})\frac{E}{R\bar{T}^2}\bar{c}_A \qquad a_4 = \frac{\overline{\partial g}}{\partial c_A} = -\frac{\bar{f}}{V} - k(\bar{T})$$

where the line over the partials is short-hand to indicate that they are evaluated at the base point. We invite you to verify the expressions for the constants by taking the derivatives of the nonlinear function.

The next step is to put the linearized first-order equation in the standard form. To do this, we move the term with $C_A(t)$ to the left side of the equal sign and divide by its coefficient, $- a_4$.

$$\tau \frac{dC_A(t)}{dt} + CA(t) = K_1 F(t) + K_2 C_{Ai}(t) + K_3 \Gamma(t)$$

where the parameters are

$$\tau = -\frac{1}{a_4} = \frac{V}{\bar{f} + Vk(\bar{T})} \qquad K_1 = -\frac{a_1}{a_4} = \frac{\bar{c}_{Ai} - \bar{c}_A}{\bar{f} + Vk(\bar{T})}$$

$$K_2 = -\frac{a_2}{a_4} = \frac{\bar{f}}{\bar{f} + Vk(\bar{T})} \qquad K_3 = -\frac{a_3}{a_4} = \frac{Vk(\bar{T})E\bar{c}_A}{R\bar{T}^2[\bar{f} + Vk(\bar{T})]}$$

Finally, Laplace transform and solve for $C_A(s)$.

$$C_A(s) = \frac{K_1}{\tau s + 1} F(s) + \frac{K_2}{\tau s + 1} C_{Ai}(s) + \frac{K_3}{\tau s + 1} \Gamma(s)$$

This model is not complete. It requires another equation for the temperature, r(t), which is not an independent input. We will present the complete model of the reactor in Chapter 4.

The preceding example shows that the parameters of the transfer function of a linearized equation depend on the values of the variables at the base point. Note that, as was pointed out earlier, the parameters depend on the partials of the nonlinear function (a,, a_2, a_3, and a_4) rather than on the value of the function itself.

This section has shown how to linearize nonlinear differential equations so that the powerful technique of **Laplace** transforms can be applied. Once the transfer function of the linearized equations is developed, the response characteristics can be related to its parameters by the methods discussed in previous sections of this chapter. The important characteristic of nonlinear systems is that their response depends on the operating point. It is convenient to think of the parameters of the linearized system as being valid at the base point rather than in a region of parameter values. In most situations, the gains and time constants do not vary enough to affect the performance of control systems significantly, but we must always keep in mind that the parameters do vary and should make allowances for their variation in the design of control systems for nonlinear   systems.

2-7 REVIEW OF COMPLEX-NUMBER ALGEBRA

The preceding sections have shown that the **Laplace** transform is a powerful tool for analyzing process and control systems dynamics. Manipulation of **Laplace** transforms requires some familiarity with the algebra of complex numbers. This section reviews some of the fundamental operations of complex numbers. Its objective is to provide a ready reference for those readers who might not feel comfortable with complex numbers.

2-7.1 Complex Numbers

A number is said to be complex when it cannot be represented as a pure real number or a pure imaginary number; **an imaginary number** is one that contains the square root of negative unity, defined as the unit of imaginary numbers, **i.** One way to write a complex number is in the **Cartesian form.**

$$c = a + i\,b \qquad\qquad (2\text{-}7.1)$$

where

   a = the **real part** of complex number c
   b = the **imaginary part,** itself a real number

Complex Plane

A complex number can be represented graphically in a plane by plotting the real part on the horizontal or **real** *axis* and the imaginary part on the vertical or **imaginary axis.** Such a plane is known as the **complex plane** and is represented in Fig. 2-7.1. Each point
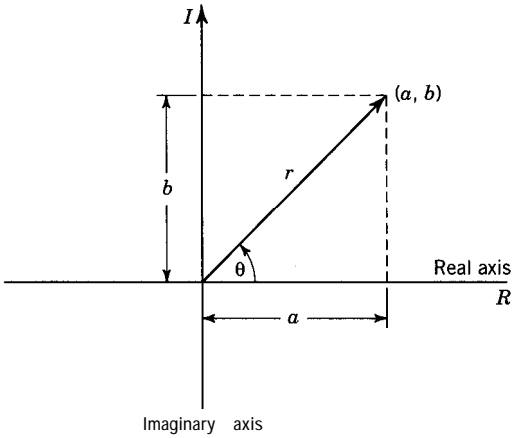
Figure 2-7.1 Complex plane showing the number
a + ib.

on this plane represents a number that can be real if on the real axis, imaginary if on the imaginary axis, and complex if anywhere else.

Polar Notation

An alternative way to represent a complex number is the **polar notation,** in which the complex number is represented by its **magnitude** and its **argument.** The magnitude is the distance from the origin to the complex number (**r** in Fig. 2-7. 1), and the argument is the angle in radians that the line from the origin to the complex number makes with the positive real axis (θ in Fig. 2-7.1). From inspection of Fig. 2-7.1, we can develop the formulas for calculating the magnitude and argument of a complex number from its real and imaginary parts:

$$r = |c| = \sqrt{a^2 + b^2}$$

$$\theta = \sphericalangle c = \tan^{-1}\frac{b}{a}$$

(2-7.2)

Similarly, by inspection of Fig. 2-7.1, we can determine that the formulas for converting from polar to Cartesian notation are

$$a = r \cos \theta$$

$$b = r \sin \theta$$

(2-7.3)

To obtain the **polar form** of complex number c, substitute Eqs. 2-7.3 into Eq. 2-7.1, and factor the magnitude, **r:**

$$c = r(\cos \theta + i \sin \theta) = re^{i\theta}$$

(2-7.4)

where we have made use of the trigonometric identity

$$e^{i\theta} \equiv cos\ \theta + i\ sin\ \theta$$

Equation 2-7.4 is in trigonometric and exponential notation.

Conjugate

The conjugate of a complex number is a complex number that has the same real part and an imaginary part equal in magnitude but opposite in sign:

$$\text{conj. } (a + ib) = a - ib \tag{2-7.5}$$

Having reviewed the basic representations of complex numbers, let us look at some common operations with complex numbers.

2-7.2   Operations with Complex Numbers

Let us consider the following two complex numbers.

$$c = a + ib = re^{i\theta}$$

$$p = v + iw = qe^{i\beta}$$

Addition/Subtraction

The sum, or difference, of two complex numbers is

$$c \pm p = (a \pm v) + i(b \pm w) \tag{2-7.6}$$

Addition and subtraction require the numbers to be in Cartesian form.

Multiplication

In Cartesian form, the product is

$$cp = (a + ib)(v + iw)$$
$$= av + i^2bw + ibv + iaw$$
$$= (av - bw) + i(bv + aw) \tag{2-7.7}$$

where we have substituted $i^2 = -1$. Multiplication is easier to do in polar form:

$$cp = (re^{i\theta})(qe^{i\beta}) = rqe^{i(\theta + \beta)} \tag{2-7.8}$$

The product of a number and its conjugate is a real number equal to the square of the magnitude of the complex number.

$$(a + ib)(a - ib) = a^2 + b^2 = r^2 \tag{2-7.9}$$

where we have made use of Eq. 2-7.2.

Division

Just like addition, subtraction, and multiplication, division of complex numbers follows the rules of algebra, except that to clear the complex number from the denominator, we must multiply its conjugate by the numerator and the denominator.

$$\begin{aligned} \frac{c}{P} &= \frac{a + ib}{v + iw} \frac{v - iw}{v \quad iw} \\ &= \frac{(av + bw) + i(bv \quad aw)}{v^2 + w^2} \\ &= \left(\frac{av + bw}{v^2 + w^2}\right) + i\left(\frac{bv - aw}{v^2 + w^2}\right) \end{aligned} \tag{2-7.10}$$

Like multiplication, division is easier to carry out in polar form:

$$\frac{c}{P} = \frac{re^{i\theta}}{qe^{i\beta}} = \frac{r}{q}e^{i(\theta - \beta)} \tag{2-7.11}$$

Raising to a Power

Raising to a power is also easier to do in polar form:

$$c^n = r^n e^{in\theta} \tag{2-7.12}$$

Roots

When all the real and complex roots of a number are considered, a number-even a real number-has n nth roots. Again, the operation is easier to carry out in polar form:

$$\sqrt[n]{c} = \sqrt[n]{re^{i\theta}} = \sqrt[n]{r}e^{i(\theta + 2k\pi)/n} \tag{2-7.13}$$

where $k = 0, \pm 1, \pm 2, \ldots$, until n distinct roots are obtained.

These are the fundamental operations of complex numbers. Some computer programming languages, FORTRAN for one, have functions that allow the programmer to carry out complex-number operations. Packages such as Mathcad and MatLab also can carry out complex-number operations. The following example illustrates these operations.

EXAMPLE 2-7.1

Given the complex numbers

$$a = 3 + i4 \qquad b = 8 - i6 \qquad c = -1 + i$$

Convert them to polar form.

$$|a| = 5 \qquad\qquad |b| = 10 \qquad\qquad |C| = 1.414$$

$$\angle a = \tan^{-1}\frac{4}{3} \qquad \angle b = \tan^{-1}\frac{-6}{8} \qquad \angle c = \tan^{-}\frac{1}{-1}$$

$$= 0.927 \text{ rad} \qquad = - 0.643 \text{ rad} \qquad = \frac{3\pi}{4} \text{ rad}$$

$$a = 5e^{i0.927} \qquad b = 10e^{-i0.643} \qquad c = 1.414e^{i(3\pi/4)}$$

The complex numbers are plotted in the complex plane in Fig. 2-7.2. Note that **b** is in the fourth quadrant and c is in the second quadrant. The arguments are in radians.
   The following are examples of addition and subtraction:

$$a + b = (3 + 8) + i(4 - 6) = 11 - i2$$

$$a - b = (3 - 8) + i(4 + 6) = -5 + i10$$
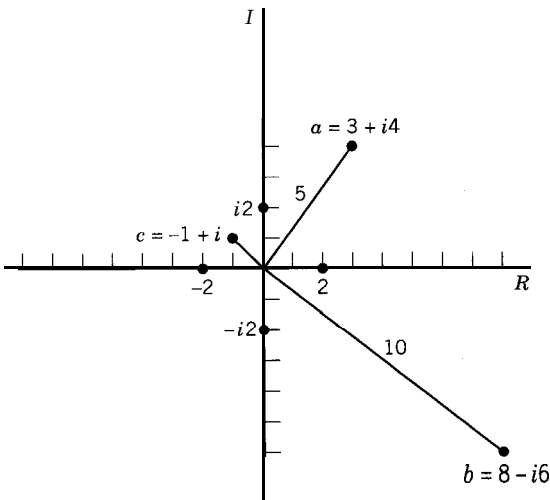
Figure 2-7.2 The complex numbers in Example 2-7.1.

The following are examples of multiplication in Cartesian coordinates:

$$ac = (-3 - 4) + i(3 - 4) = -7 - i$$

$$bc = (-8 + 6) + i(8 + 6) = -2 + i14$$

We can obtain the same answers in polar form:

$$ac = 5e^{i0.927}1.414e^{i(3\pi/4)} = 7.07e^{i3.2834}$$
$$= 7.07 \cos 3.2834 + i7.07 \sin 3.2834 = -7 - i$$

The following illustrates the distributive property of multiplication:

$$(a + b)c = (11 - i2)(-1 + i) \qquad = -9 + i13$$

$$ac + bc = (-7 - i) + (-2 + i14) \qquad = -9 + i13$$

The following exemplifies division:

$$\frac{a}{b} = \frac{3 + i4}{8 - i6}\frac{8 + i6}{8 + i6} = \frac{(24 \quad 24) + i(18 + 32)}{64 + 36} = i0 \quad 5$$

We can obtain answer in polar form:

$$\frac{a}{b} = \frac{5e^{i0.927}}{10e^{-i0.643}} = 0,5e^{i1.570} = 0.5(0 + i) = io.5$$

Finally, let us find the fourth roots of $16 = 16e^{i0}$.

$$x = \sqrt[4]{16e^{i0}} = \sqrt[4]{16}e^{i(0 + 2k\pi/4)} = 2e^{i(k\pi/2)}$$

.The roots are

For	The Roots Are
$k = O$	$x = 2e^{i0} = 2$
$k = 1$	$x = 2e^{i\pi/2} = 2(0 + i) = i2$
$k = -1$	$x = 2e^{-i\pi/2} = 2(0 - i) = -i2$
$k = 2$	$x = 2e^{i\pi} = 2(-1 + i0) = -2$

These roots are also plotted in Fig. 2-7.2.

2-8 SUMMARY

This chapter presented the techniques of **Laplace** transforms and linearization in the way they are applied for analyzing the dynamic response of processes and their control systems. The characteristics of the process response to input signals were related to the roots of the denominator of the process transfer function, and the responses and transfer functions of first- and second-order systems were presented. In the chapters that follow, the transfer functions of specific processes will be related to the physical process parameters through the application of the fundamental laws of conservation. In later chapters, we will use **Laplace** transforms for designing and analyzing process control systems.

PROBLEMS

2-1. Using the definition of the **Laplace** transform, derive the transforms F(s) of the following functions.

(a) $f(t) = t$
(b) $f(t) = e^{-at}$, where a is a constant
(c) $f(t) = \cos \omega t$, where ω is a constant
(d) $f(t) = e^{-at} \cos \omega t$, where a and ω are constant

Note: In parts (c) and (d) you will need the trigonometric identity

$$\cos x = \frac{e^{ix} + e^{-ix}}{2}$$

Check your answers against the entries in Table 2-1.1.

2-2. Using a table of **Laplace** transforms and the properties of the transform, find the transforms $F(s)$ of the following functions.

(a) $f(t) = u(t) + 2t + 3t^2$
(b) $f(t) = e^{-2t}[u(t) + 2t + 3t^2]$
(c) $f(t) = u(t) + e^{-2t} - 2e^{-t}$
(d) $f(t) = u(t) - e^{-t} + te^{-t}$
(e) $f(t) = u(t - 2)[1 - e^{-2(t-2)} \sin(t - 2)]$

2-3. Check the validity of your results to Problem 2-2 by applying the initial and final value theorems. Do these theorems apply in all of the cases?

2-4. In Example 2-1.1 (b), the **Laplace** transform of a pulse was obtained by application of the definition of the transform. Show that the same transform can be obtained by application of the real translation theorem. Note that the pulse is the difference between two identical step changes of size H with the second one delayed by the duration of the pulse, T:

$$f(t) = Hu(t) - Hu(t - T)$$

2-5. In the statement of the real translation theorem, we pointed out that for the theorem to apply, the delayed function has to be zero for all times less than the

delay time. Show this by calculating the **Laplace** transform of the function

$$f(t) = e^{-(t-t_0)/\tau}$$

where t_0 and τ are constants.

(a) Assuming that it holds for all times greater than zero-that is, that it can be rearranged as

$$f(t) = e^{t_0/\tau}e^{-t/\tau}$$

(b) If it is zero for $t \le t_0$—that is, that it should be properly written as

$$f(t) = u(t \quad t_0)e^{-(t-t_0)/\tau}$$

Sketch the graph of the two functions. Are the two answers the same? Which one agrees with the result of the real translation theorem?

2-6. Obtain the solution $Y(t)$, as a deviation from its initial steady-state condition y(O), of the following differential equations. Use the method of **Laplace** transforms and partial fractions expansion. The forcing function is the unit step function, $X(t) = u(t)$.

(a) $\dfrac{dy(t)}{dt} + 2y(t) = 5x(t) + 3$

(b) $9 \dfrac{d^2y(t)}{dt^2} + 18 \dfrac{dy(t)}{dt} + 4y(t) = 8x(t) - 4$

(c) $9 \dfrac{d^2y(t)}{dt^2} + 9 \dfrac{dy(t)}{dt} + 4y(t) = 8x(t) - 4$

(d) $9 \dfrac{d^2y(t)}{dt^2} + 12 \dfrac{dy(t)}{dt} + 4y(t) = 8x(t) - 4$

(e) $2 \dfrac{d^3y(t)}{dt^3} + 7 \dfrac{d^2y(t)}{dt^2} + 21 \dfrac{dy(t)}{dt} + 9y(t) = 3x(t)$

2-7. Repeat Problem 2-6(d) using as the forcing function
   (a) $X(t) = e^{-t/3}$ (b) $X(t) = u(t - 1)e^{-(t-1)/3}$

2-8. For the differential equations given in Problem 2-6, determine whether the response is stable or unstable, oscillatory or monotonic. Find also the dominant root, the period of the oscillations and the decay ratio if the response is oscillatory; the time required for the slowest term in the response, or the amplitude of the oscillations, to decay to within less than 1% (0.67%) of its initial value; and the final steady-state value of the output. *Note:* It is not necessary to completely solve Problem 2-6 to answer the questions in this problem.

2-9. *Second-Order Response: Bird Mobile.* The bird mobile shown in Fig. P2-1 has a mass of 50 g, and the spring that holds it extends 27 cm when the weight of the bird is applied to it. Neglecting resistance of the air to the motion of the bird,
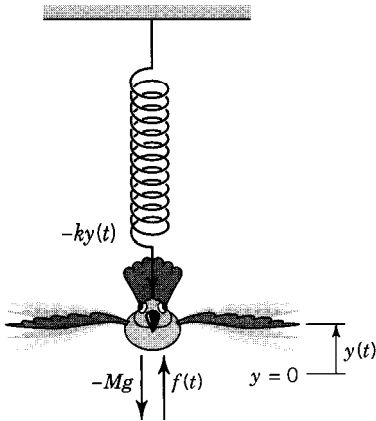
Figure **P2-1** Bird mobile for Problem 2-9.

we can derive the following equation by writing a dynamic force balance on the bird:

$$M \frac{d^2 y(t)}{dt^2} = -Mg - ky(t) + f(t)$$

where $y(t)$ is the vertical position of the bird in m, $f(t)$ is the force required to start the bird in motion in N, M is the mass of the bird in kg, k is the spring constant in N/m, and g is the local acceleration of gravity, 9.8 m/s². Find the period of oscillation of the bird. Does the solution predict that the bird will oscillate forever? What term must be added to the model equation to reflect the actual motion of the bird more accurately? What is the physical significance of this added term?

2-10. For the general first-order differential equation, Eq. 2-4.4, obtain the response to
   (a) An impulse, $X(t) = S(t)$.
   (b) The pulse sketched in Fig. 2-1.1(b).
   Sketch the graph of the response, $Y(t)$, for each case.

2-11. *Response of an Integrating Process.* The response of the liquid level in a tank is given by the first-order differential equation

$$A \frac{dh(t)}{dt} = f(t)$$

where $h(t)$ is the level in the tank in m, $f(t)$ is the flow of liquid into the tank in cubic meters per second (m³/s), and A is the constant area of the tank in square meters. Obtain the transfer function for the tank and the response of the level to a unit step in flow, $F(t) = u(t)$. Sketch the graph of the level response, $H(t)$. Why do you think we call this result the response of an *integrating process?*

2-12. For the second-order differential equations given in Problem 2-6, find the characteristic time and damping ratio, and classify them as overdamped or under- ' damped. For the overdamped equations, figure out the effective time constants,

and for the underdamped equations, find the frequency and period of oscillation, the decay ratio, and the percent overshoot, rise time, and settling time on a step input.

2-13. Evaluate the coefficients of the partial fractions expansion that leads to Eqs. 2-5.10, 2-5.11, 2-5.12, and 2-5.13.

2-14. Derive Eq. 2-5.23 from Eq. 2-5.22.

2-15. A common transfer function that models many second-order interacting systems (see Chapter 4), is

$$Y(s) = \frac{k_1}{[(\tau_1 s + 1)(\tau_2 s + 1)] - k_2} X(s)$$

where τ_1, τ_2, k_1, and k_2 are constants. Relate the gain, characteristic time, and damping ratio of the second-order transfer function to the four constants appearing in the transfer function. Assuming that all four constants are positive real numbers,

(a) Show that the response is overdamped.
(b) Show that the response is stable if $k_2 < 1$.
(c) Relate the two effective time constants to the four constants of the transfer function.

2-16. The transfer function of a feedback control loop is given by

$$C(s) = \frac{K_c}{(3s + 1)(s + 1) + K_c} R(s)$$

where K_c is the controller gain. Relate the gain, characteristic time, and damping ratio of the second-order transfer function to the controller gain. Find the ranges of the controller gain for which the response is (i) overdamped, (ii) underdamped, and (iii) undamped. Can the response be unstable for any positive value of the controller gain?

2-17. Linearize the following nonlinear functions, and express your results in terms of deviations from the base point.

(a) The equation for enthalpy as a function of temperature, Eq. 2-6.1.
(b) The Antoine equation for the vapor pressure, Eq. 2-6.2.
(c) The equation for vapor mole fraction at equilibrium as a function of liquid mole fraction, Eq. 2-6.3.
(d) The equation for fluid flow as a function of pressure drop, Eq. 2-6.4.
(e) The equation for radiation heat transfer rate as a function of temperature, Eq. 2-6.5.

2-18. As pointed out in the text, the error of the linear approximation usually increases as the variable deviates from its base value. The error in the slope is the one that is important. For the rate of radiation heat transfer as a function of temperature, Eq. 2-6.5, find the range of temperature for which the slope of the function, $dq/dT|_T$, remains within $\pm 5\%$ of its base value. Calculate also the temperature range for which the linear approximation of the heat transfer rate, q, is within $\pm 5\%$ of its true value. Consider two base values, $\overline{T} = 400$ K and $\overline{T} = 600$ K. Discuss briefly how the range of applicability of the linear approximation, based on its

ability to match the slope of the function, varies with the base value of the temperature.

2-19. Repeat Problem 2-18 for the formula of the equilibrium vapor mole fraction as a function of liquid mole fraction, Eq. 2-6.3. Calculate the range of values of the liquid mole fraction x for which the slope, $dy/dx|_x$, is within $\pm 5\%$ of its value at the base point, $\bar{x}$. Calculate also the range of values of x for which the linear approximation to the equilibrium mole fraction y remains within $\pm 5\%$ of its true value. Consider the following cases:

(a) $\alpha = 1.10$ $\quad \bar{x} = 0.10$ $\qquad$ (c) $\alpha = 5.0$ $\quad \bar{x} = 0.10$
(b) $\alpha = 1.10$ $\quad \bar{x} = 0.90$ $\qquad$ (d) $\alpha = 5.0$ $\quad \bar{x} = 0.90$

Discuss briefly why you think the range of applicability of the linear approximation, based on its ability to match the slope of the function, varies with both the parameter and the base value for this function.

2-20. The rate of a chemical reaction is given by the expression

$$r[c_A(t), c_B(t)] = kc_A^2(t)c_B(t)$$

where $k = 0.5$ m^6/(kmole^2-h) is a constant (isothermal operation). Obtain the linear approximation of this function at $\bar{c}_A = 2$ kmole/m^3, $\bar{c}_B = 1$ kmole/m^3, and find the error in the parameters of the approximation (the partial derivatives) when each of the concentrations changes, independently, by 1 kmole/m^3. Express the linear approximation in terms of deviation variables.

2-21. Raoult's law gives the vapor mole fraction y(t) at equilibrium as a function of the temperature, T(t), pressure p(t), and liquid mole fraction, x(t):

$$y[T(t), p(t), x(t)] = \frac{p^o[T(t)]}{p(t)} x(t)$$

where $p^o[T(t)]$ is the vapor pressure of the pure component, given by Antoine's equation, Eq. 2-6.2. Obtain the linear approximation for the vapor mole fraction and express it in terms of deviation variables. Evaluate the parameters of the approximation for benzene at atmospheric pressure (760 mm Hg), 95°C, and a liquid mole fraction of 50%. The Antoine constants for Benzene are $A = 15.9008$, $B = 2788.51°C$, and C = 220.80°C, for the vapor pressure in mm Hg.

2-22. Evaluate the parameters of the transform obtained in Example 2-6.4, using the parameters of the Arrhenius formula given in Example 2-6.1 and the following reactor parameters:

$$V = 2.6 \text{ m}^3 \qquad \bar{f} = 0.002 \text{ m}^3/\text{s} \qquad \bar{c}_{Ai} = 12 \text{ kmoles/m}^3$$

Note that the initial value of the reactant concentration, $\bar{c}_A$, can be found from the condition that the initial condition is, at steady state, $\bar{g} = 0$.

2-23. A compressed air tank at a gas station is punctured by a stray bullet fired by a careless robber. The mass balance of air in the tank is

$$V\frac{d\rho(t)}{dt} = w_i(t) - A_o\sqrt{2\rho(t)[p(t) - p_o]}$$

where

$$\rho(t) = \frac{M}{RT} p(t)$$

w(t) kg/s is the inlet flow from the air compressor, $V = 1.5$ m^3 is the volume of the tank, $A, = 0.785$ cm^2 is the area of the bullet hole, $M = 29$ kg/kmole is the molecular weight of air, $R = 8.314$ kPa-m^3/kmole-K is the ideal gas law constant, $p_o = 101.3$ kPa, and the temperature T is assumed constant at 70°C.

Obtain a linear approximation to the differential equation around the initial pressure of 500 kPa gauge. Obtain also the Laplace transform for the pressure in the tank, and evaluate the time constant and gain of the transfer function.

2-24. The temperature of a turkey in an oven, assumed uniform throughout the bird, and neglecting the heat absorbed by the cooking reactions, is given by the following differential equation:

$$Mc_v \frac{dT(t)}{dt} = \sigma \epsilon A[T_s^4(t) - T^4(t)]$$

where M is the mass of the turkey in pounds, c_v is the specific heat in Btu/lb-°R, $T(t)$ is the temperature of the turkey in °R, $\sigma = 0.17$ 14 . 10^{-8} Btu/h-ft^2-°R^4 is the Stephan-Boltzmann constant, ϵ is the emissivity of the skin of the turkey, A is the area of the turkey in square feet, and $T,(t)$ is the temperature of the oven in "R.

Obtain a linear approximation to the differential equation. Also obtain the Laplace transform for the temperature of the turkey, and write the expressions for the time constant and the gain of the transfer function. What is the input variable for this problem?

2-25. The temperature of a slab being heated by an electric heater is given by the differential equation

$$C \frac{dT(t)}{dt} = q(t) - \alpha[T^4(t) - T_s^4]$$

where $T(t)$ is the temperature of the slab in "R, assumed uniform, $q(t)$ is the rate of heat input in Btu/h, $C = 180$ $Btu/°R$ is the heat capacity of the slab, $T_s = 540°R$ is the surrounding temperature (constant), and $\alpha = 5$. 10^{-8} Btu/h-°R^4 is the coefficient of heat radiation. Obtain the linear approximation of the differential equation around the initial steady-state temperature of 700°R. Obtain also the Laplace transform of the temperature of the slab, and find the gain and time constant of the transfer function.

Chapter **3**

First-Order Dynamic Systems

As briefly presented in Chapter 1, the dynamic response of processes is of prime consideration in the design, analysis, and implementation of process control systems. An interesting and important characteristic of chemical processes is that their dynamics change from one process to another. For instance, the response of temperature is different from the response of level. Further, the response of temperature in a heat exchanger is different from the response of temperature in a furnace. The principal objective of this chapter is to show how to describe the dynamic response of simple processes by using mathematical models, transfer functions, and block diagrams. Though simple, these processes are taken from actual industrial applications. Chapter 4 presents more complex processes.

The mathematical models will be developed starting from first principles. From the models come the transfer functions that reveal the terms that describe the process response: gain, time constant, and dead time (transportation delay or time delay). Most of the mathematics reviewed in Chapter 2 is extensively used in this chapter.

The modeling of industrial processes usually starts with a balance of a conserved quantity: mass or energy. The balance can be written as

$$
\begin{array}{ccccc}
\text{Rate of mass/energy} & & \text{Rate of mass/energy} & & \text{Rate of accumulation} \\
\text{into} \quad \text{control} & - & \text{out of control} & = & \text{of mass/energy} \\
\text{volume} & & \text{volume} & & \text{in control volume}
\end{array}
$$

In processes where chemical reactions are not present, the moles are also conserved. Thus in these processes, we may substitute the term *moles* for *mass* in the balance equation. Section 3-6 discusses processes where chemical reactions are present.

In writing these balances, and all other auxiliary equations, we call on our knowledge of many areas of process engineering, such as thermodynamics, heat transfer, fluid flow, mass transfer, and reaction engineering. This makes the modeling of industrial processes most interesting and challenging!

3-1 PROCESSES AND THE IMPORTANCE OF PROCESS CHARACTERISTICS

It is important to start this chapter, and indeed to launch the entire subject of process control, by explaining what a "process" is and describing its characteristics from a process control point of view. To do this, let us consider the heat exchanger of Chapter 1, shown again in Fig. 3-1.1a.

The controller's job is to control the process. In the example at hand, the controller is to take action that keeps the outlet temperature, T(t), at a specified value, its set point. However, the controller receives a signal from the transmitter. It is through the transmitter that the controller "sees" the controlled variable. Thus, realistically, as far as the controller is concerned, the controlled variable is the transmitter output (TO). The relation between the transmitter's output and the physical variable to control, T(t), is given by the transmitter calibration as presented in Chapter 5.

In this example, the controller is to manipulate the steam valve position to maintain the controlled variable at set point. Note, however, that the way the controller manipulates the valve position is by changing its output signal to the valve. The controller does not manipulate the valve position directly; it manipulates only its output signal. Thus, as far as the controller is concerned, the manipulated variable is its own output (CO).

We can now define the *process* as *anything between the controller output (CO) and the controller input.* Most often, the controller input is provided by the transmitter output. There are some instances, however, in which this may not be the case, such as when a mathematical manipulation, such as filtering, is done on the signal from the transmitter before it is received by the controller. Because the transmitter usually provides the input to the controller directly, we can say that *controller input is equal to the transmitter output (TO).* We will point out those examples in which this is not the case. In Fig. 3-1.1a, the process is anything within the area delineated by the solid curve. The process includes the I/P transducer, the valve, the heat exchanger with associated piping, the sensor, and the transmitter.
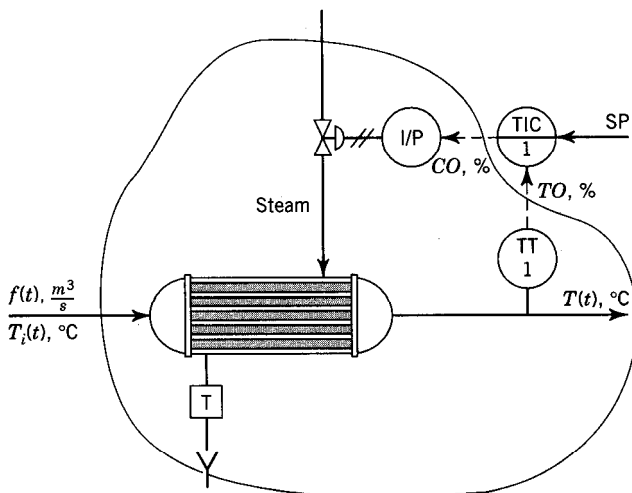
Figure 3-1.1a Heat exchanger control system.
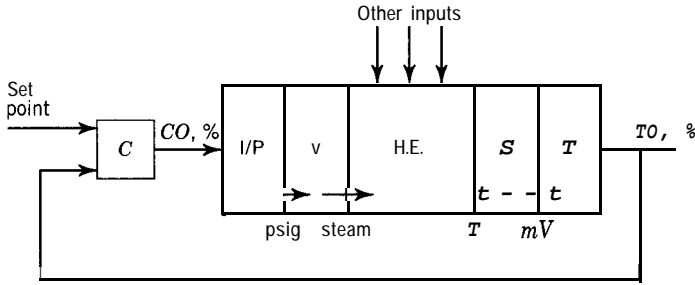
Figure 3-1.1b Heat exchanger temperature control.

For further understanding of what we have just discussed, consider Fig. 3-1. lb. The diagram shows all the parts of the process and how they are related. It shows that the output signal from the controller (CO) enters the I/P transducer, producing a pneumatic signal. This signal then goes to the valve (V), producing a steam flow. This flow enters the heat exchanger (H.E.) and, along with other process inputs, produces an output temperature T(t). This temperature is measured by a sensor (S), and the output signal from the sensor, maybe in millivolts, is received by the transmitter (T), which produces a signal (TO) to the controller. Thus the diagram clearly shows that as far as the controller is concerned, the controlled variable is the transmitter's output (TO) and the manipulated variable is the controller's output (CO).

Why is it important to understand the characteristics of the process to be controlled? As we noted in Chapter 1, the control performance provided by the controller depends on the adjustment or specification of different terms in the controller. Setting these terms is referred to as tuning the controller. The optimum controller tuning depends on the process to be controlled and on the tuning criterion. Every controller must be tuned specifically for the process it controls. Consequently, to tune a controller, we must first understand the characteristics, or behavior, of the process to be controlled.

Another way to explain the need to understand the characteristics of the process is to realize that in tuning the controller, what we are doing is "adapting" the controller to the process. Thus it makes sense first to obtain the process characteristics and then to tune the controller, or adapt the "controller characteristics," to that of the process. If this is done correctly, the complete closed-loop control system, process plus controller, will perform as required.

The present chapter and Chapter 4 discuss processes and their characteristics. Chapter 5 briefly presents some terms related to transmitters and also discusses control valves and controllers and their characteristics. Finally, Chapter 6 puts everything together; it "closes the loop." Chapter 7 shows how to tune the feedback controller once the process characteristics are known. Herein, then, lies the importance of knowing, understanding, and obtaining the process characteristics. *We can tune the controller only after the steady state and dynamic characteristics of the process are known.*

3-2   THERMAL   PROCESS   EXAMPLE

Consider the well-stirred tank shown in Fig. 3-2.1. In this process, constant and equal inlet and outlet volumetric flows, liquid densities, and heat capacities are assumed; all of these properties are known. The liquid in the tank is assumed to be well mixed, and
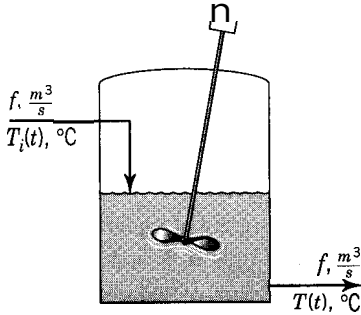
Figure 3-2.1 Thermal process.

the tank is well insulated-that is, there are negligible heat losses to the surroundings. Finally, the energy input by the stirrer is assumed negligible.

We are interested in developing the mathematical model and transfer function that describe how the outlet temperature, T(t), responds to changes in inlet temperature, $T_i(t)$. An unsteady-state energy balance on the contents of the tank, the control volume, gives us the desired relation between the inlet and outlet temperatures. That is,

$$\begin{matrix} \text{Rate of energy} \\ \text{into tank} \end{matrix} - \begin{matrix} \text{Rate of energy} \\ \text{out of tank} \end{matrix} = \begin{matrix} \text{Rate of accumulation} \\ \text{of energy in tank} \end{matrix}$$

or, in terms of an equation,

$$f\rho_i h_i(t) - f\rho h(t) = \frac{d[V\rho u(t)]}{dt}$$

where

$$f = \text{volumetric flow, } m^3/s$$
$$\rho_i, \rho = \text{inlet and outlet liquid densities, respectively, } kg/m^3$$
$$V = \text{volume of liquid in tank, } m^3$$
$$h_i(t), h(t) = \text{inlet and outlet liquid enthalpies, respectively, } J/kg$$
$$u(t) = \text{internal energy of liquid in tank, } J/kg$$

In terms of temperatures, using as reference state for u(t) and h(t) the pure component in the liquid state at 0°F and the pressure of the system, the foregoing equation can be written as

$$f\rho_i C_{pi} T_i(t) - f\rho C_p T(t) = \frac{d[V\rho C_v T(t)]}{dt} \qquad (3\text{-}2.1)$$

where

$$C_{pi}, C_p = \text{inlet and outlet liquid heat capacities at constant pressure,}$$
$$\text{respectively, } J/kg\text{-}°C$$
$$C_v = \text{liquid heat capacity at constant volume, } J/kg\text{-}°C$$
$$T_i(t), T(t) = \text{inlet and outlet temperatures, respectively, } °C$$

Because the densities and the heat capacities are assumed constant over the operating temperature range, the last equation can be written as

$$f\rho C_p T_i(t) \quad f\rho C_p T(t) = V\rho C_v \frac{dT(t)}{dt}$$

(3-2.2)

This equation is a first-order linear ordinary differential equation that provides the relationship between the inlet and outlet temperatures. It is important to note that in this equation there is only one unknown, T(t). The inlet temperature, $T_i(t)$, is an input variable and the one that forces the outlet temperature to change. In this example, we want to study how T,(t) affects T(t), so it is up to us to decide how this inlet temperature will change. Thus it is not considered an unknown. In this chapter and the following one, input variables are not considered unknowns because we have the freedom to change them as we wish.

To show that there is one equation with one unknown, we explicitly write

$$f\rho C_p T_i(t) - f\rho C_p T(t) = V\rho C_v \frac{dT(t)}{dt}$$

(3-2.3)

1 eq., 1 unk. [T(t)]

Equation 3-2.3 is the mathematical model for this process. The solution of this differential equation yields the response of the outlet temperature as a function of time. As just mentioned, the inlet temperature is the input variable, which is sometimes referred to as the forcing function because it is the variable that forces the outlet temperature to change. The outlet temperature is the output variable, which is sometimes referred to as the responding variable because it is the variable that responds to changes in the forcing function, or input variable.

As stated in the beginning of this example, we are interested in obtaining the transfer function relating T(t) to $T_i(t)$. To do so, we follow a series of steps that yield the desired transfer function; after this example, we will formalize the procedure. We begin by making a variable change that simplifies development of the required transfer function.

Write a steady-state energy balance on the contents of the tank at the initial conditions.

$$f\rho C_p \bar{T}_i - f\rho C_p \bar{T} = 0$$

(3-2.4)

Subtracting this equation from Eq. 3-2.3 yields

$$f\rho C_p [T_i(t) - \bar{T}_i] - f\rho C_p [T(t) - \bar{T}] = V\rho C_v \frac{d[T(t) - \bar{\bar{T}}]}{dt}$$

(3-2.5)

Note that the derivative of the temperature is also equal to

$$\frac{d[T(t) - \bar{T}]}{dt} = \frac{dT(t)}{dt} - \frac{d\bar{T}}{dt} 0 \frac{dT(t)}{dt}$$

which is the result of subtracting the right-hand side of Eqs. 3-2.3 and 3-2.4. This is only a trick that proves helpful in the definition of deviation variables and the development of transfer functions.

As presented in Chapter 2, we now define the following deviation variables:

$$\Gamma(t) = \mathbf{T(t)} - \bar{T} \tag{3-2.6}$$

$$\Gamma_i(t) = T_i(t) - \bar{T}_i \tag{3-2.7}$$

where

$\bar{T}, \bar{T}_i$ = initial steady-state values of outlet and inlet temperatures, respectively, °C

$\Gamma(t), \Gamma_i(t)$ = deviation variables of outlet and inlet temperatures, respectively, °C

Substituting Eqs. 3-2.6 and (3-2.7) into (3-2.5) yields

$$f\rho C_p \Gamma_i(t) - f\rho C_p \Gamma(t) = V\rho C_v \frac{d\Gamma(t)}{dt} \tag{3-2.8}$$

Eq. 3-2.8 is the same as Eq. 3-2.3 except that it is written in terms of deviation variables. The solution of this equation yields the deviation variable $\Gamma(t)$ versus time for a certain input $\Gamma_i(t)$. If the actual outlet temperature, $T(t)$, is desired, the steady-state value $\bar{T}$ must be added to $\Gamma(t)$ in accordance with Eq. 3-2.6.

Deviation variables are used almost exclusively throughout control theory. Thus the meaning and importance of deviation variables in the analysis and design of process control systems must be well understood. As explained in Chapter 2, their value indicates the degree of deviation from some initial steady-state value. In practice, this steady-state value may be the desired value of the variable. Another advantage in the use of these variables is that their initial value, assuming we start from the initial steady state, is zero, which simplifies the solution of differential equations such as Eq 3-2.8 by the **Laplace** transform.

Equation 3-2.8 can now be rearranged as follows:

$$\frac{V\rho C_v}{f\rho C_p} \frac{d\Gamma(t)}{dt} + \Gamma(t) = \Gamma_i(t)$$

and we let

$$\tau = \frac{V\rho C_v}{f\rho C_p} \tag{3-2.9}$$

so

$$\tau \frac{d\Gamma(t)}{dt} + \Gamma(t) = \Gamma_i(t) \tag{3-2.10}$$

The units of τ are those of time. From Eq. 3-2.9 we see that for this example,

$$\tau = \frac{[m^3][kg/m^3][J/kg\text{-}°C]}{[m^3/s][kg/m^3][J/kg\text{-}°C]} = \text{seconds}$$

Because Eq. (3-2.10) is a linear differential equation, the use of **Laplace** transform yields

$$\tau s \Gamma(s) - \tau \Gamma(0) + \Gamma(s) = \Gamma_i(s)$$

But the initial value of the temperature, T(O), is at $\overline{T}$, so $\Gamma(0) = 0$. Performing some simple algebraic manipulations gives

$$\Gamma(s) = \frac{1}{\tau s + 1} \Gamma_i(s) \qquad (3\text{-}2.\ 11)$$

or

$$\frac{\Gamma(s)}{\Gamma_i(s)} = \frac{1}{\tau s + 1} \qquad (3\text{-}2.12)$$

Equation 3-2.12 is the desired transfer function. It is a first-order transfer function because it is developed from a first-order differential equation. As we saw in Chapter 2, processes described by this type of transfer function **are called first-order processes** or *first-order lugs.* Equation 2-4.6 presented the general form of this type of transfer function. In the present example the term, **K,** is unity.

The term **transfer function** arises from the fact that the solution of the equation translates, or transfers, the input, $\Gamma_i(t)$, to the output, $\Gamma(t)$. Transfer functions are further discussed in Section 3-3.

As a brief review of Chapter 2, let us assume that the inlet temperature, $T_i(t)$, to the tank increases by M "C. That is, the inlet temperature experiences a step change of M degrees in magnitude. Mathematically, this is written as

$$T_i(t) = \overline{T}_i \qquad\qquad t < 0$$

$$T_i(t) = \overline{T}_i + M \qquad t \geq 0$$

or, in terms of deviation variables,

$$\Gamma_i(t) = Mu(t)$$

where $u(t)$, as shown in Chapter 2, represents a step change of unit magnitude.

Taking the Laplace transform, we obtain

$$\Gamma_i(s) = \frac{M}{s}$$

Substituting this expression for $\Gamma_i(s)$ into Eq. 3-2.11 results in

$$\Gamma(s) = \frac{M}{s(\tau s + 1)}$$

Using the method of partial fractions presented in Chapter 2 yields

$$\Gamma(s) = \frac{M}{s(\tau s + 1)} = \frac{A}{s} + \frac{B}{\tau s + 1}$$

Obtaining the values of A and B by partial fraction and inverting back to the time domain, we get

$$\Gamma(t) = M(1 - e^{-t/\tau}) \qquad \text{(3-2.13)}$$

or

$$T(t) = \overline{T} + M(1 - e^{-t/\tau}) \qquad \text{(3-2.14)}$$

The solutions of Eqs. 3-2.13 and 3-2.14 are shown graphically in Fig. 3-2.2. The steepest slope of the response curve occurs at the beginning of the response; this is the typical response of first-order systems to a step change in input. Note also that because K is unity, the total change in the output is M "C.
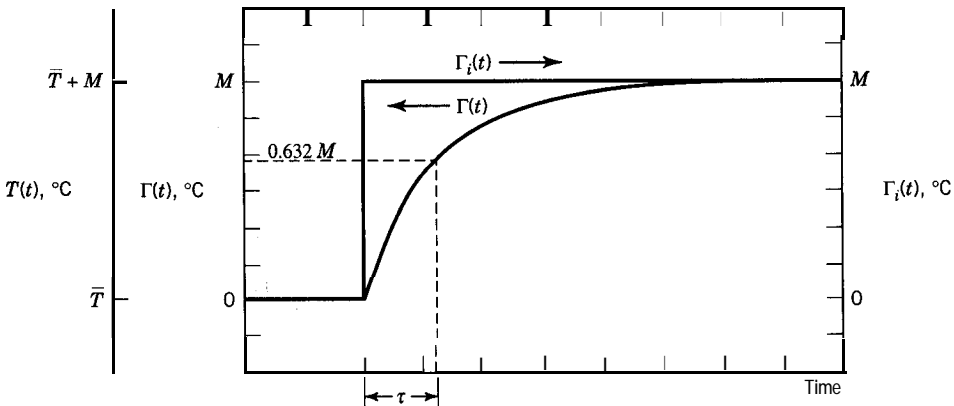
Figure 3-2.2 Response of a first-order process to a step change in input variable.

Section 2-4 presented the significance *ofprocess time constant, τ.* However, to review again, let $t = \tau$ in Eq. 3-2.13, which yields

$$r(r) = M(1 - e^{-\tau/\tau}) = M(1 - e^{-1})$$

$$r(r) = 0.632M$$

That is, *for a step change in input variable, the time constant indicates the time it takes the output variable to reach 63.2% of its total change;* this is shown graphically in Fig. 3-2.2. In five time constants, 5τ, the process reaches 99.7% of its total change; essentially the response is completed. Therefore, *the time constant is related to the speed of response of the process.* The slower a process responds to an input, the larger the value of τ. The faster the process responds to an input, the smaller the value of τ.

It is important to realize that the time constant is composed of the different physical properties and operating parameters of the process, as shown by Eq. 3-2.9. That is, the time constant depends on the volume of liquid in the tank *(V),* the heat capacities $(C_P$ and C,), and the process flow (f). If any of these characteristics changes, the behavior of the process also changes and this change is reflected in the speed of response of the process, or the time constant.

Up to now, the tank has been assumed to be well insulated-that is, there are negligible heat losses to the surroundings. Consequently, there is not a heat loss term in the energy balance. Let us remove this assumption and develop the mathematical model and the transfer functions that relate the outlet temperature, $T(t)$, to the inlet temperature, $T_i(t)$, and to the surrounding temperature, $T_s(t)$.

As before, using the same reference state for enthalpies and internal energy, we start with an unsteady-state energy balance:

$$f\rho C_p T_i(t) - q(t) - f\rho C_p T(t) = V\rho C_v \frac{dT(t)}{dt}$$

or

$$f\rho C_p T_i(t) - UA[T(t) - T_s(t)] - f\rho C_p T(t) = V\rho C_v \frac{dT(t)}{dt} \qquad \text{(3-2. 15)}$$

1 eq., 1 unk. [T(t)]

where

$q(t)$ = heat transfer rate to the surroundings, J/s
U = overall heat transfer coefficient, $\text{J/m}^2\text{-K-s}$
A = heat transfer area, m^2
$T_s(t)$ = temperature of surroundings, "C, an input variable

The overall heat transfer coefficient, $U,$ is a function of several things, one of them being temperature. However, in this particular example, it is assumed to be constant. Because the mass of liquid in the tank and its density are also assumed to be constant, the height of liquid is constant, and consequently the heat transfer area, *A,* is also constant.

Equation 3-2.15 provides the mathematical model of the process. To obtain the transfer functions, we start by introducing the deviation variables. This is done by first writing a steady-state energy balance for this process at the initial conditions.

$$f\rho C_p \bar{T}_i - UA[\bar{T} - \bar{T}_s] - f\rho C_p \bar{T} = 0 \tag{3-2.16}$$

Subtracting Eq. 3-2.16 from Eq. 3-2.15 yields

$$f\rho C_p[T_i(t) - \bar{T}_i] - UA[(T(t) - \bar{T}) - (T_s(t) - \bar{T}_s)]$$

$$- f\rho C_p[T(t) - \bar{T}] = v\rho c_v \frac{d[T(t) - \bar{T}]}{dt} \tag{3-2.17}$$

Please note that the trick with the differential term (accumulation) has been done again. Define a new deviation variable as

$$\Gamma_s(t) = T_s(t) - \bar{T}_s \tag{3-2.18}$$

Substituting Eqs. 3-2.6, 3-2.7, and 3-2.18 into Eq. 3-2.17 yields

$$f\rho C_p \Gamma_i(t) - UA[\Gamma(t) - \Gamma_s(t)] - f\rho C_p \Gamma(t) = V\rho C_v \frac{d\Gamma(t)}{dt} \tag{3-2.19}$$

Equation 3-2.19 is the same as Eq. 3-2.15 except that it is written in terms of deviation variables. This equation is also a first-order linear ordinary differential equation. In this case, there is still one equation with one unknown, $\Gamma(t)$. The new variable is the surrounding temperature $\Gamma_s(t)$, which is another input. As this temperature changes, it affects the heat losses and consequently the process liquid temperature.

Equation 3-2.19 can be arranged as follows:

$$\frac{V\rho C_v}{f\rho C_p + UA} \frac{d\Gamma(t)}{dt} + \Gamma(t) = \frac{f\rho C_p}{f\rho C_p + UA} \Gamma_i(t) + \frac{UA}{f\rho C_p + UA} \Gamma_s(t)$$

or

$$\tau \frac{d\Gamma(t)}{dt} + \Gamma(t) = K_1 \Gamma_i(t) + K_2 \Gamma_s(t) \tag{3-2.20}$$

where

$$\tau = \frac{V\rho C_v}{f\rho C_p + UA}, \text{ seconds} \tag{3-2.21}$$

$$K_1 = \frac{f\rho C_p}{f\rho C_p + UA}, \text{ dimensionless} \tag{3-2.22}$$

$$K_2 = \frac{UA}{f\rho C_p + UA}, \quad \text{dimensionless} \tag{3-2.23}$$

The right-hand side of Eq. 3-2.20 shows the two input variables, $\Gamma_i(t)$ *and* $r_s(t)$, acting on the output variable, $r(t)$.

Taking the Laplace transform of Eq. 3-2.20 gives

$$\tau s\Gamma(s) - \tau\Gamma(0) + \Gamma(s) = K_1\Gamma_i(s) + K_2\Gamma_s(s)$$

But the initial value of the temperature, $T(O)$, is at $\overline{T}$, so $\Gamma(0) = 0$. Rearranging this equation yields

$$\Gamma(s) = \frac{K_1}{\tau s + 1}\Gamma_i(s) + \frac{K_2}{\tau s + 1}\Gamma_s(s) \tag{3-2.24}$$

If the surrounding temperature remains constant, $T_s(t) = \overline{T}_s$ then $\Gamma_s(t) = 0$, and the transfer function relating the process temperature to the inlet temperature is

$$\frac{\Gamma(s)}{\Gamma_i(s)} = \frac{K_1}{\tau s + 1} \tag{3-2.25}$$

If the inlet liquid temperature remains constant, $T_i(t) = \overline{T}_i$ then $\Gamma_i(t) = 0$, and the transfer function relating the process temperature to the surrounding temperature is

$$\frac{\Gamma(s)}{\Gamma_s(s)} = \frac{K_2}{\tau s + 1} \tag{3-2.26}$$

If both the inlet liquid temperature and the surrounding temperature change, then Eq. 3-2.24 provides the complete relationship.

Equations 3-2.25 and 3-2.26 are the typical first-order transfer functions presented in Section 2-4. In this case, however, the **steady-state gains** (sometimes also called **process gains), K,** and K_2, are not unity, as was the case in Eq. 3-2.12. To review briefly the significance of the steady-state gains, let us assume that the inlet temperature to the tank increases, in a step fashion, by M °C, that is,

$$\Gamma_i(s) = \frac{M}{s}$$

The response of the temperature to this forcing function is given by

$$\Gamma(s) = \frac{K_1 M}{s(\tau s + 1)}$$

from   which

$$\Gamma(t) = K_1 M (1 - e^{-t/\tau}) \tag{3-2.27}$$

or

$$T(t) = \bar{T} + K_1 M (1 - e^{-t/\tau}) \tag{3-2.28}$$

The output response is shown graphically in Fig. 3-2.3. The total amount of change in T(t) is given by $K_1 M$, the gain times the change in input. Thus the gain *tells us how much the output changes per unit change in input,* or *how much the input affects the output.* That is, the gain defines *the sensitivity* relating the output and input variables! It can also be defined mathematically as follows:

$$K = \frac{\Delta O}{\Delta I} = \frac{A\ \text{output variable}}{A\ \text{input variable}} \tag{3-2.29}$$

The gain is another parameter that describes the characteristics of the process. Consequently, it depends on the physical properties and operating parameters of the process, as shown by Eqs. 3-2.22 and 3-2.23. The gains in this process depend on the flow, density, and heat capacity of the process liquid (f, p, and C_p), on the overall heat transfer coefficient (U), and on the heat transfer area *(A)*. If any of these changes, the behavior of the process changes and is reflected in the gain.

There are two gains in this example. The first one, K, , relates the outlet temperature to the inlet temperature. The other gain, K_2, relates the outlet temperature to the surrounding temperature. The units of the gain term must be the units of the output variable divided by the units of the input variable; this is obvious from Eq. 3-2.29.

Note that the gain clearly indicates the process characteristics. In the first part of this example, the tank was assumed to be well insulated, and the gain, given by Eq. 3-2.12, was unity. That is, in the steady state, all the energy entering with the inlet stream exits with the outlet stream, and the inlet and outlet temperatures are the same. This is not
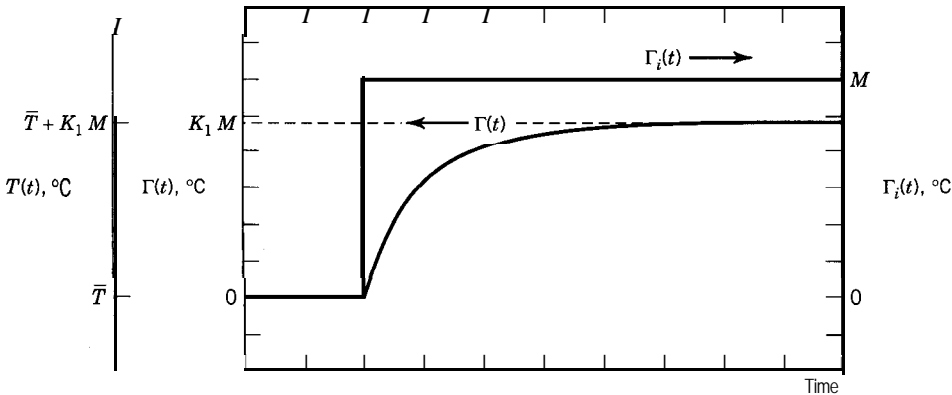
Figure   3-2.3 Response   of   a   first-order   process   to   a   step   change   in   input   variable.

the case when the assumption of good insulation is removed and the tank is permitted to transfer energy with the surroundings. Note that K_1 and K_2, given by Eqs. 3-2.22 and 3-2.23, are less than unity, which indicates that when the inlet temperature increases by M degrees, the outlet temperature does not increase by that much. That is, if the energy in the inlet stream increases, then the energy in the outlet stream does not increase as much because there is some energy transfer to the surroundings; this of course makes sense. It also makes sense that if $UA \ll f\rho C_p$, then the inlet temperature will have a greater effect on the outlet temperature than will the surrounding temperature, that is, $K_2 \ll K_1$.

Equation 3-2.24 shows that there is only one time constant in this process. That is, the time it takes the outlet temperature to reach a certain percentage of its total change due to a change in inlet temperature is equal to the time it takes to reach the same percentage of the total change when the surrounding temperature changes.

It is always important, during the analysis of any process, to stop at some point to check the development for possible errors. After the development of Eq. 3-2.20 is usually a convenient point. A quick check can be made by examining the signs of the equation to see whether they make sense in the real world. In Eq. 3-2.20, both gains are positive. The equation indicates that if the inlet temperature increases, then the outlet temperature also increases; which makes sense for this process. The equation also shows that if the surrounding temperature increases, then the outlet temperature increases. This makes sense because when the surrounding temperature increases, the rate of heat losses from the tank decreases, thereby increasing the temperature of the contents of the tank. Another check consists of examining the units of τ and K. We know what each of them should be, and the defining equations, Eqs. 3-2.21 through 3-2.23 in this example, should confirm these expectations. This quick check builds our confidence and permits us to proceed with the analysis with a renewed hope of success.

Before finishing with this section, let us summarize the procedure we followed to develop the transfer functions.

1. Write the set of unsteady-state equations that describes the process. This is called modeling.
2. Write the steady-state equations at the initial conditions.
3. Subtract the two sets of equations, and define the deviation variables.
4. Obtain the Laplace transforms of the model in deviation variables.
5. Obtain the transfer functions by solving the Laplace transform explicitly for the transformed output variable(s).

We followed these five steps in our thermal example. They constitute an organized procedure that yields the transfer functions.

3-3 DEAD TIME

Consider the process shown in Fig. 3-3.1. This is essentially the same process as the one shown in Fig. 3-2.1. In this case, however, we are interested in knowing how T,(t) responds to changes in inlet and surrounding temperatures.

Let us make the following two assumptions about the exit pipe between the tank and point 1. First, the pipe is well insulated. Second, the flow of liquid through the pipe is ideal plug flow (highly turbulent) with no energy diffusion or dispersion so that there is essentially no backmixing of the liquid in the pipe.
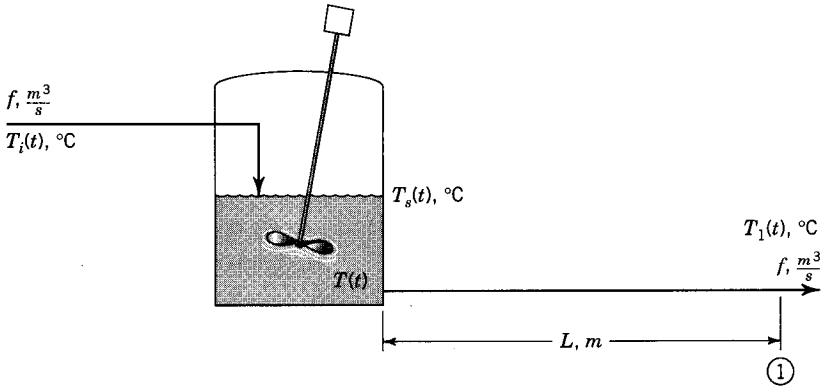
Figure 3-3.1 Thermal process.

Under these assumptions, the response of $T_1(t)$ to the disturbances $T_i(t)$ and $T_s(t)$ will be the same as $T(t)$ except that it will be delayed by some amount of time. That is, there will be a finite amount of time between the initial response of $T(t)$ and the change of $T_1(t)$; this delay is shown graphically in Fig. 3-3.2. This finite amount of time has developed because of the time it takes the liquid to move from the exit of the tank to point 1 and is called a *pure dead time, transportation delay,* or *time delay.* It is represented by t_0 and in this case can be easily estimated from

$$t_0 = \frac{\text{distance}}{\text{velocity}} = \frac{L}{f/A_p} = \frac{A_p L}{f} \tag{3-3.1}$$

where

f = volumetric flow, m³/s
A_{p} = cross-sectional area of pipe, m²
L = length of pipe, m
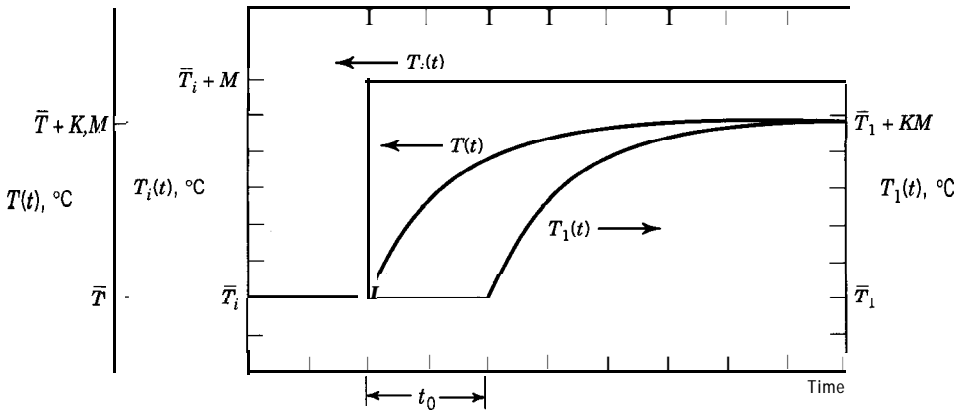
Figure 3-3.2 Response of a thermal process to a step change in inlet temperature.

Different physical variables travel at different velocities:

- Electric voltage and current travel at the speed of light: 300,000 km/s, or 984,106 ft/s.
- Liquid flow and pressure travel at the speed of sound in the fluid: 340 m/s, or 1100 ft/s.
- Temperature, composition, and other fluid properties travel at the velocity of the fluid: typically, up to 5 m/s (15 ft/s) for liquids and 60 m/s (200 ft/s) for gases.
- Solid properties travel at the velocity of the solid, such as coal in a conveyor, cake in a filter bed, and paper in a paper machine.

From this information, we can see that for the distances typical of industrial process control systems, pure dead time is significant only for temperature, composition, and other fluid and solid properties that are propagated through space by the moving fluid or solid.

Even when pure dead time (dead time due to transportation) is negligible relative to the process time constant, the response of many processes may appear to exhibit dead time due to the combination of several first-order processes in series, as we shall see in Chapters 4 and 6. This pseudo-dead time cannot be easily evaluated from fundamental principles and must be obtained empirically by approximation of the process response. Methods to carry out such empirical evaluation will be presented in Chapter 7.

Because dead time is an integral part of processes, it must be accounted for in the transfer functions. Equation 2-1.8 indicates that the **Laplace** transform of a delayed function is equal to the **Laplace** transform of the nondelayed function times the term $e^{-t_0 s}$; the term $e^{-t_0 s}$ is the **Laplace** transform of dead time. Thus, if the transfer functions relating $T_i(t)$ to $T_i(t)$ and $T_i(t)$ are required, using the assumptions stated at the beginning of the section, the transfer functions given by Eqs. 3-2.25 and 3-2.26 are multiplied by $e^{-t_0 s}$ or

$$\frac{\Gamma_1(s)}{\Gamma_i(s)} = \frac{K_1 e^{-t_0 s}}{\tau s + 1} \tag{3-3.2}$$

and

$$\frac{\Gamma_1(s)}{\Gamma_s(s)} = \frac{K_2 e^{-t_0 s}}{\tau s + 1} \tag{3-3.3}$$

At this point, it must be recognized that the dead time is another parameter that helps define the characteristics of the process. Equation 3-3.1 shows that t_0 depends on some physical properties and operating characteristics of the process, similar to K and τ. If any condition of the process changes, then this change may be reflected in a change in t_0.

Before concluding this section, we must stress that one *of the worst things that can happen to a feedback control loop is a significant amount of dead time in the loop. The* performance of feedback control loops is severely affected by dead time, as we will see in Chapters 6, 8, and 9. Thus processes and control systems should be designed to keep the dead time to a minimum. Some steps we can take to minimize dead time include putting the measurements as close to the equipment as possible, selecting rapidly re-

sponding sensors and final control elements, and using electronic instead of pneumatic instrumentation for processes with short time constants.

3-4 TRANSFER FUNCTIONS AND BLOCK DIAGRAMS

3-4.1 Transfer Functions

Chapter 2 presented the concept of transfer functions. This concept is so fundamental to the study of process dynamics and automatic process control that at this time we briefly consider, once more, some of its important properties and characteristics.

We have already defined a transfer function as the ratio of the Laplace-transformed output variable to the Laplace-transformed input variable. Transfer functions are usually represented by

$$G(s) = \frac{Y(s)}{X(s)} = \frac{K(a_m s^m + a_{m-1} s^{m-1} + \ldots + a_1 s + 1) e^{-t_0 s}}{(b_n s^n + b_{n-1} s^{n-1} + \ldots + b_1 s + 1)} \qquad \textbf{(3-4.1)}$$

where

$G(s)$ = general representation of a transfer function
$Y(s)$ = Laplace transform of the output variable
$X(s)$ = Laplace transform of the input variable
$K, a's, b's$ = constants
t_0 = dead time

Equation 3-4.1 shows the most general and best way to write a transfer function. When it is written in this way, K represents the gain of the system and has as units the units of $Y(t)$ over the units of $X(t)$. The other constants, a's and $b's$, have as units (time)^i, where i is the power of the Laplace variable, s, multiplied by the particular constant; this will render a dimensionless term inside the parentheses because the unit of s is 1/time. Notice that the coefficient of s^0 is 1.

Note: In general, the unit of s is the reciprocal of the unit of the independent variable used in the definition of Laplace transform, Eq. 2-1.1. In process dynamics and control, the independent variable is time, so the unit of s is 1/time.

The transfer function completely defines the steady-state and dynamic characteristics, or the total response, of a system described by a linear differential equation. It is characteristic of the system, and its terms determine whether the system is stable or unstable and whether its response to a non-oscillatory input is oscillatory. The system, or process, is said to be stable when its output remains bound (finite) for all times for a bound input. Chapter 2 presented some discussions on stability and how it is related to terms in the transfer function. Chapters 6, 8, and 9 treat in more detail the subject of stability of process systems.

The following are some important properties of transfer functions.

1. In the transfer functions of real physical systems, the highest power of s in the numerator is never higher than that in the denominator. In other words, $n \geq m$.
2. The transfer function relates the transforms of the deviation of the input and output variables from some initial steady state. Otherwise, the nonzero initial conditions would contribute additional terms to the transform of the output variable.

3. For stable systems, the steady-state relationship between the change in output variable and the change in input variable can be obtained by

$$\lim_{s \to 0} G(s)$$

This stems from the final value theorem, presented in Chapter 2.

$$\lim_{t \to \infty} Y(t) = \lim_{s \to 0} sY(s)$$

$$\lim_{t \to \infty} Y(t) = \lim_{s \to 0} sG(s)X(s)$$

$$\lim_{t \to \infty} Y(t) = [\lim_{s \to 0} G(s)][\lim_{s \to 0} sX(s)]$$

$$\lim_{t \to \infty} Y(t) = [\lim_{s \to 0} G(s)][\lim_{t \to \infty} X(t)]$$

This means that the change in the output variable after a very long time, if bound, can be obtained by multiplying the transfer function evaluated at $s = 0$ times the final value of the change in input.

3-4.2 Block Diagrams

A very useful tool in process control is the graphical representation of transfer functions by means of block diagrams. This section offers an introduction to block diagrams and block diagram algebra.

All block diagrams are formed by a combination of four basic elements: arrows, summing points, branch points, and blocks; Fig. 3-4.1 shows these elements. The *arrows* in general indicate flow of information; they represent process variables or control signals. Each arrowhead indicates the direction of the flow of information. The *summing points* represent the algebraic summation of the input arrows, $E(s) = R(s) - C(s)$. A *branch point* is the position on an arrow at which the information
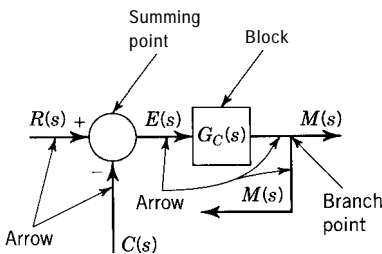
Figure 3-4.1 Elements of a block diagram.
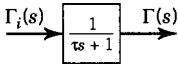
Figure 3-4.2 Block diagram of Eq. 3-2.12.

branches out and goes concurrently to other summing points or blocks. The **blocks** represent the mathematical operation, in transfer function form such as G(s), which is performed on the input to produce the output. The arrows and block shown in Fig. 3-4.1 represent the mathematical expression

$$M(s) = G_c(s)E(s) = G_c(s)[R(s) - C(s)]$$

Any block diagram can be handled, or manipulated, algebraically. Table 3-4.1 shows some rules of block diagram algebra. These rules are important any time a complicated block diagram is simplified. Let us look at some examples of block diagram algebra.

EXAMPLE 3-4.1

Draw the block diagram depicting Eqs. 3-2.12 and 3-2.24.
    Equation 3-2.12 is shown in Fig. 3-4.2. Equation 3-2.24 may be drawn in at least two different ways, as shown in Fig. 3-4.3. Often the diagram with fewer blocks is preferred because it is simpler.

    The block diagrams of Eq. 3-2.24 show graphically that the total response of the system is obtained by algebraically adding the response due to a change in inlet temperature to the response due to a change in surrounding temperature. This algebraic addition of responses due to several inputs to obtain the final response is a property of linear systems and is called the **principle of superposition.** This principle also serves as the basis for defining linear systems. That is, we say that a system is linear if it obeys the principle of superposition.
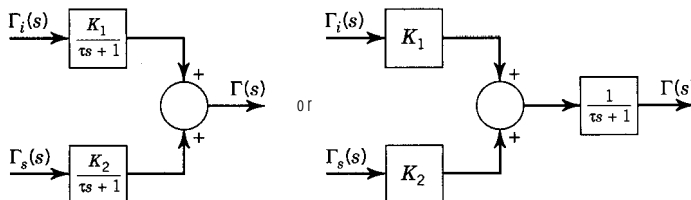
Figure 3-4.3 Block diagram of Eq. 3-2.24.

Table 3-4.1 Rules for Block Diagram Algebra

1. $Y(S) = X_1(S) \quad X_2(s) - X_3(s)$
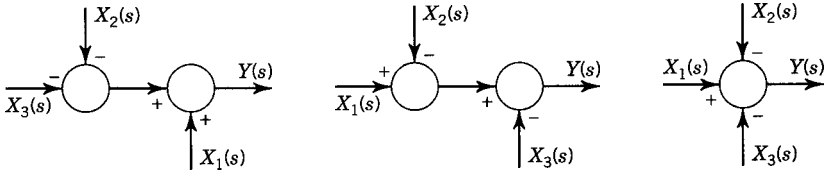
2. Associative and Commutative Properties:
$$Y(s) = G_1(s) \, G_2(s) \, X(s) = G_2(s) \, G_1(s) \, X(s)$$
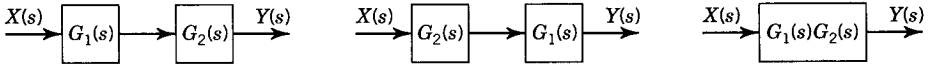
3. Distributive Property:
$$Y(s) = G_1(s) \, [X_1(s) - X_2(s)] = G_1(s) \, X_1(s) - G_1(s) \, X_2(s)$$
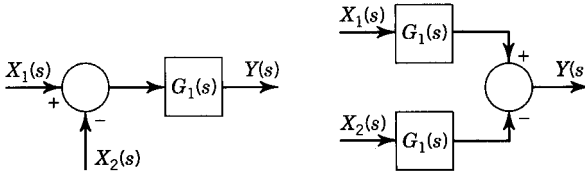
4. Blocks in Parallel:
$$Y(s) = [G_1(s) + G_2(s)] \, X(s) = G_1(s) \, X(s) + G_2(s) \, X(s)$$
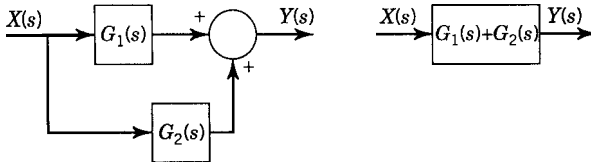
5. Positive Feedback Loop:
$$Y(s) = G_1(s) \, [X(s) + G_2(s) \, Y(s)] = \frac{G_1(s)}{1 - G_1(s) \, G_2(s)} X(s)$$
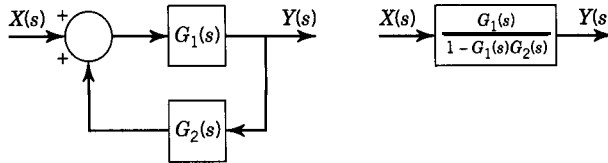
6. Negative Feedback Loop:
$$Y(s) = G_1(s) \, [X(s) + G_2(s) \, Y(s)] = \frac{G_1(s)}{1 + G_1(s) \, G_2(s)} X(s)$$
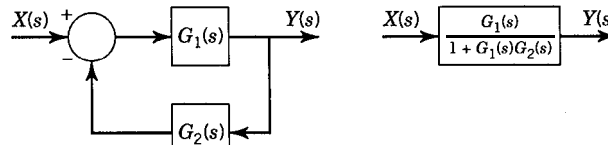
EXAMPLE 3-4.2

Determine the transfer functions relating Y(s) to $X_1(s)$ and $X_2(s)$ from the block diagram shown in Fig. 3-4.4a. That is, obtain

$$\frac{Y(s)}{X_1(s)} \qquad \text{and} \qquad \frac{Y(s)}{X_2(s)}$$

Using rule 4 of Table 3-4.1, the block diagram shown in Fig. 3-4.4a can be reduced to that of Fig. 3-4.4b (please note that reduction is used in this context to mean

(a)

(b)

(c)
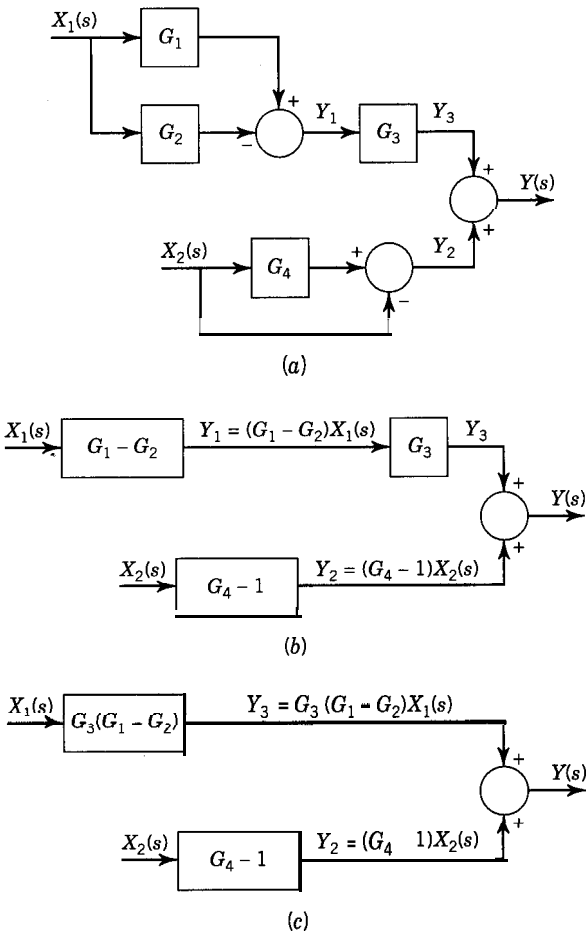
Figure 3-4.4 Block diagram for Example 3-4.2.

simplification and that it consists of reducing the number of blocks). Using rule 2, Fig. 3-4.4b can be further reduced to Fig. 3-4.4c. Then

$$Y(s) = G_3(G_1 - G_2)X_1(s) + (G_4 - 1)X_2(s)$$

from which the two desired transfer functions can be determined. They are

$$\frac{Y(s)}{X_1(s)} = G_3(G_1 - G_2)$$

and

$$\frac{Y(s)}{X_2(s)} = G_4 - 1$$

Example 3-4.2 has shown a procedure to reduce a block diagram to a transfer function. This reduction of block diagrams is necessary in the study of process control, as will be clear in later chapters. In these chapters, numerous examples of block diagrams of feedback, cascade, feedforward, and multivariable control systems are developed. Let us look at the reduction to transfer functions of some of these block diagrams.

Figure 3-4.5 shows the block diagram of a typical feedback control system. From this diagram,   determine

$$\frac{C(s)}{L \ o} \quad \text{and} \quad \frac{C(s)}{C^{\text{set}}(s)}$$

Figures 3-4.5b through d show the different reduction steps using the rules in Table 3-4.1. Finally, from Fig. 3-4.5d, we obtain the transfer functions as

$$\frac{C(s)}{C^{\text{set}}(s)} = \frac{G_1 G_c G_2 G_3 G_4}{1 + G_c G_2 G_3 G_4 G_6} \tag{3-4.2}$$

and

$$\frac{C(s)}{L(s)} = \frac{G_5 G_4}{1 + G_c G_2 G_3 G_4 G_6} \tag{3-4.3}$$
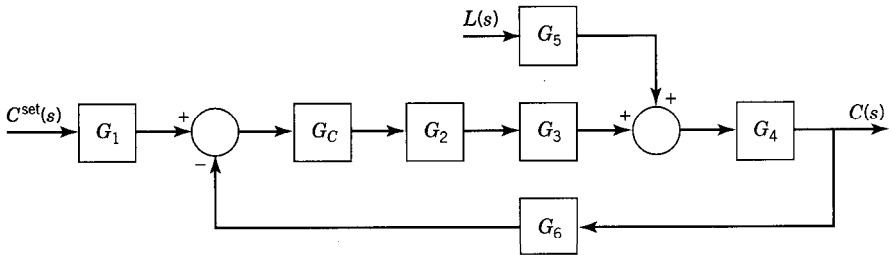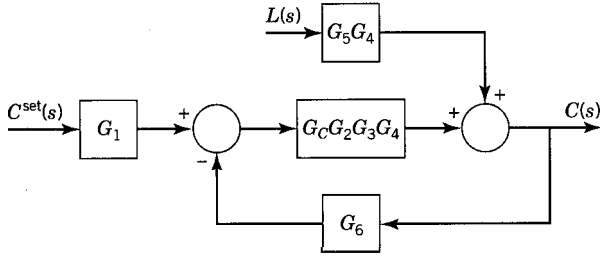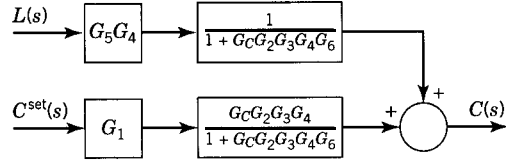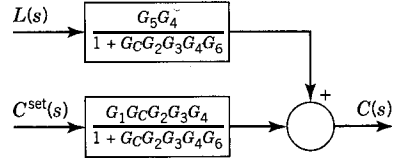
Figure 3-4.5 Block diagram of a feedback control system.

Example 3-4.3 shows how to reduce a simple feedback control system block diagram to transfer functions. These types of block diagrams and transfer functions will become useful in Chapters 6, 7, 8, and 9, when feedback control is discussed.

The transfer functions given by Eqs. 3-4.2 and 3-4.3 are referred to as **closed-loop transfer functions.** The reason for this term will become evident in Chapter 6. Looking at Eq. 3-4.2, note that the numerator is the product of all of the transfer functions in the forward path between the two variables related by the transfer function, $C(s)$ and $C^{set}(s)$. The denominator of this equation is 1 plus the product of all the transfer functions in the control loop shown in Fig. 3-4.5a. Inspection of Eq. 3-4.3 shows that the

numerator is again the product of the transfer functions in the forward path between $L(s)$ and C(s). The denominator is the same as that of Eq. 3-4.2. If there had been more than one forward path between input and output, the development would have shown the numerator to be the algebraic summation of the product of the transfer functions in each forward path.

EXAMPLE 3-4.4

Consider another typical block diagram as shown in Fig. 3-4.6a. Chapter 10 shows that this block diagram depicts a cascade control system. For now, simply determine the following transfer functions:

$$\frac{C(s)}{R(s)} \quad \text{and} \quad \frac{C(s)}{L(s)}$$

The block diagram of Fig. 3-4.6a can be thought of as being composed of two closed-loop systems, one inside the other (in practice, this is exactly what it is). Figure 3-4.6b and c show the steps to reduce the block diagram of Fig. 3-4.6a; rule 6 is applied twice. From Fig. 3-4.6c, the following transfer functions are obtained:

$$\frac{C(s)}{R(s)} = \frac{G_{c_1}G_{c_2}G_1G_4}{1 + G_{c_2}G_1G_3 + G_{c_1}G_{c_2}G_1G_4G_5} \tag{3-4.4}$$

and

$$\frac{C(s)}{L(s)} = \frac{G_2G_4}{1 + G_{c_2}G_1G_3 + G_{c_1}G_{c_2}G_1G_4G_5} \tag{3-4.5}$$

We have learned how to develop several transfer functions (Eqs. 3-4.2, 3-4.3, 3-4.4, and 3-4.5) from block diagrams. We have not intended, however, to give their significance; this will be done in the chapters where control systems are presented.

A useful recommendation is to write, next to each arrow, the units of the process variable or control signal that the arrow represents. This makes it fairly simple to recognize the units of the gain of a block, which are the units of the output arrow over the units of the input arrow. This procedure also helps avoid the algebraic summation of arrows with different units. It is extensively illustrated in Chapters 6, 8, 9, 10, 11, 12, and 13.

As mentioned at the beginning of this section, block diagrams are a very helpful tool in process control. They show the flow of information in a graphical way, identify the input and output signals (or variables) in a system, and show the occurrence of loops and parallel paths. We will learn more about the logic of block diagrams and get more practice drawing them as we continue our study of process dynamics and control. Later chapters make use of block diagrams to help analyze and design control systems.
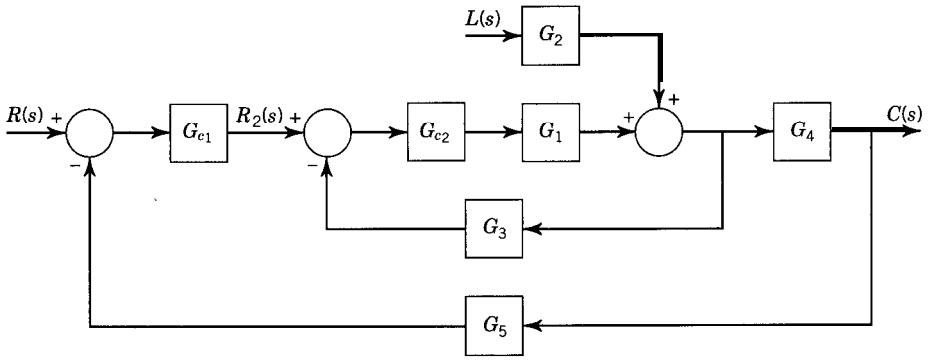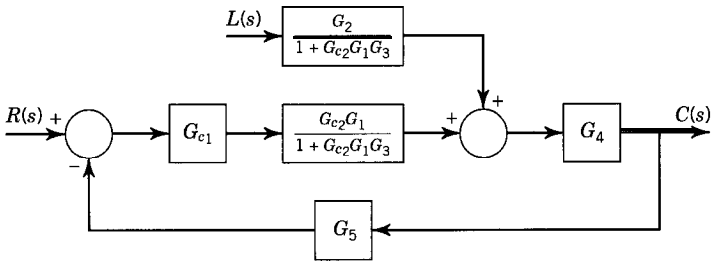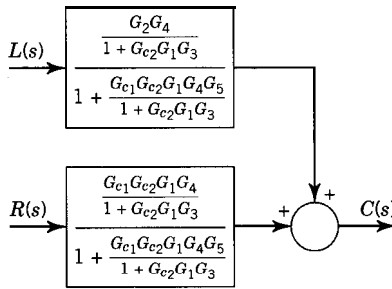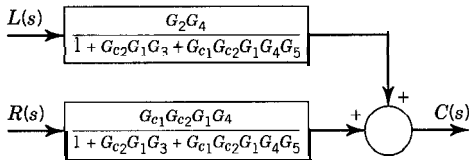
Figure 3-4.6 Block diagram of a cascade control system.
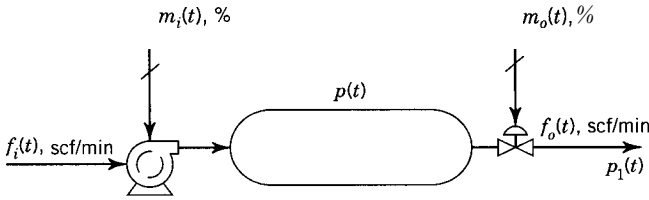
Figure 3-5.1 Gas vessel.

3-5 GAS PROCESS EXAMPLE

Consider the gas vessel shown in Fig. 3-5.1. A fan blows air into a tank, and from the tank the air flows out through a valve. For purposes of this example, let us suppose that the air flow delivered by the fan is given by

$$f.(t) = 0.16 m_i(t)$$

where

$f_i(t)$ = gas flow in scf/min, where scf is cubic feet at standard conditions
of 60°F and 1 atm
$m,(t)$ = signal to fan, %

The flow through the valve is expressed by

$$f_o(t) = 0.00506 m_o(t) \sqrt{p(t)[p(t) - p_1(t)]}$$

where

$f_o(t)$ = gas flow, scf/min
$m,(t)$ = signal to valve, %
$p(t)$ = pressure in tank, psia
$p_1(t)$ = downstream pressure from valve, psia

The volume of the tank is 20 ft³, and it can be assumed that the process occurs iso-thermally at 60°F. The initial steady-state conditions are

$$\bar{f}_i = \bar{f}_o = 8 \text{ scfm} \qquad \bar{p} = 40 \text{ psia} \qquad \bar{p}_1 = 1 \text{ atm} \qquad \bar{m}_i = \bar{m}_0 = 50\%$$

We want to develop the mathematical model, transfer functions, and block diagram that relate the pressure in the tank to changes in the signal to the fan, $m,(t)$; in the signal to the valve, $m,(t)$; and in the downstream pressure, $p_1(t)$.

We must first develop the mathematical model for this process. An unsteady-state mole balance around the control volume, defined as the fan, tank, and outlet valve, provides the starting relation. That is

$$\begin{array}{ccc} \text{Rate of moles into} & \text{Rate of moles out} & \text{Rate of accumulation} \\ \text{control volume} & - & \text{of control volume} & = & \text{of moles in control volume} \end{array}$$

or, in equation form,

$$\bar{\rho}f_i(t) - \bar{\rho}f_o(t) = \frac{dn(t)}{dt} \tag{3-5.1}$$

1 eq., 3 unk. $[f_i(t), f_o(t), n(t)]$

where

$\bar{\rho}$ = molal density of gas at standard conditions, 0.00263 lbmoles/scf
$n(t)$ = moles of gas in tank, lbmoles

The fan provides another equation:

$$f_i(t) = 0.16m_i(t) \tag{3-5.2}$$

2 eq., 3 unk.

Note that because $m_i(t)$ is an input variable, it is up to us to decide how it will change. Thus it is not considered an unknown.
    The valve provides still another equation:

$$f_o(t) = 0.00506m_o(t) \sqrt{p(t)[p(t) - p_1(t)]} \tag{3-5.3}$$

3 eq., 4 unk. $[p(t)]$

The signal $m,(t)$ and downstream pressure $p_1(t)$ are other input variables and thus are not considered unknowns.
    Because the pressure in the tank is low, the ideal gas equation of state can be used to relate the moles in the tank to the pressure.

$$p(t)V = n(t)RT \tag{3-5.4}$$

4 eq., 4 unk.

    The set of Eqs. 3-5.1 through 3-5.4 constitutes the mathematical model for this process. The solution of this set of equations describes, considering the assumptions taken, how the pressure in the tank (the output) responds to changes in $m,(t)$, $m,(t)$, and $p_1(t)$ (the inputs).
    So far we have completed the first step of the procedure, outlined at the end of Section 3-2. Before proceeding to the second step, we must realize that the expression for $f_o(t)$, Eq. 3-5.3, is a nonlinear equation. The Laplace transformation can be applied only to linear equations. Thus, before continuing to the second step, we must linearize all the nonlinear terms. This linearization is done using Taylor Series expansion as presented in Chapter 2.

Because $f_o(t) = f_o[m_o(t), p(t), p_1(t)]$, its linearization is done with respect to $m_1(t)$, $p(t)$, and $p_1(t)$ about their steady-state values $\overline{m}_o, \overline{p}, \text{and } \overline{p}_1$.

$$f_o(t) \approx \overline{f}_o + \left.\frac{\partial f_o(t)}{\partial m_o(t)}\right|_{ss} [m_o(t) - \overline{m}_o] + \left.\frac{\partial f_o(t)}{\partial p(t)}\right|_{ss} [p(t) - \overline{p}] + \left.\frac{\partial f_o(t)}{\partial p_1(t)}\right|_{ss} [p_1(t) - \overline{p}_1]$$

or

$$f_o(t) \approx \overline{f}_o + C_1[m_o(t) - \overline{m}_o] + C_2[p(t) - \overline{p}] + C_3[p_1(t) - \overline{p}_1] \qquad (3\text{-}5.5)$$

where

$$C_1 \frac{\partial f_o(t)}{\partial m_o(t)}\bigg|_{ss} = 0.00506 \sqrt{\overline{p}(\overline{p} - \overline{p}_1)} \qquad (3\text{-}5.6)$$

$$C_2 = \frac{\partial f_o(t)}{\partial p(t)}\bigg|_{ss} = 0.00506\overline{m}_o(1/2)[\overline{p}(\overline{p} - \overline{p}_1)]^{-1/2}(2\overline{p} - \overline{p}_1) \qquad (3\text{-}5.7)$$

$$C_3 = \frac{\partial f_o(t)}{\partial p_1(t)}\bigg|_{ss} = 0.00506\overline{m}_o(1/2)[\overline{p}(\overline{p} - \overline{p}_1)]^{-1/2}(-\overline{p}) \qquad (3\text{-}5.8)$$

and

$$\overline{f}_o = f_o[\overline{m}_o, \overline{p}, \overline{p}_1] \qquad (3\text{-}5.9)$$

Now there is a set of linear equations (Eqs. 3-5.1, 3-5.2, 3-5.4, and 3-5.5) that describes the process around the linearization values of $\overline{m}_o$, $\overline{p}$, and $\overline{p}_1$.

To simplify this set somewhat, solve for $n(t)$ in Eq. 3-5.4 and substitute it in Eq. 3-5.1.

$$\overline{\rho}f_i(t) - \overline{\rho}f_o(t) = \frac{V}{RT}\frac{dp(t)}{dt} \qquad (3\text{-}5.10)$$

With this simple substitution, the set of equation is reduced to three equations, Eqs. 3-5.10, 3-5.2, and 3-5.5, with three unknowns, $f_i(t), f_o(t)$, and $p(t)$.

We can now proceed with the next two steps of the procedure, which call for writing the steady-state equations, subtracting them from their respective counterparts, and defining the required deviation variables.

First we write a steady-state mole balance around the tank.

$$\overline{\rho}\overline{f}_i - \overline{\rho}\overline{f}_o = 0$$

Subtracting this equation from Eq. 3-5.10 gives

$$\bar{\rho}[f_i(t) - \bar{f}_i] - \bar{\rho}[f_o(t) - \bar{f}_o] = \frac{V}{RT}\frac{d[p(t) - \bar{p}]}{dt} \tag{3-5.11}$$

Defining the following deviation variables

$$F_i(t) = f_i(t) - \bar{f}_i \qquad F_o(t) = f_o(t) - \bar{f}_o \qquad P(t) = p(t) - \bar{p}$$

and substituting these variables into Eq. 3-5.11 yield

$$\bar{\rho}F_i(t) - \bar{\rho}F_o(t) = \frac{V}{RT}\frac{dP(t)}{dt} \tag{3-5.12}$$

Writing the steady-state equation for the fan and subtracting it from Eq. 3-5.2 give

$$F_i(t) = 0.16M_i(t) \tag{3-5.13}$$

where $M_i(t) = m_i(t) - \bar{m}_i$.
   From Eq. 3-5.5, after subtraction off, from both sides of the equation,

$$F_o(t) = C_1 M_o(t) + C_2 P(t) + C_3 P_1(t) \tag{3-5.14}$$

where

$$M_o(t) = m_o(t) - \bar{m}_o$$
$$P_1(t) = p_1(t) - \bar{p}_1$$

Recapping what has been done, there are now three equations, Eqs. 3-5.12 through 3-5.14, and three unknowns, $F_i(t)$, $F_o(t)$, and $P(t)$. All of these equations and variables are in deviation form.
   We now proceed with the last two steps of the procedure. Substituting Eqs. 3-5.13, and 3-5.14 into Eq. 3-5.12, taking the Laplace transform and rearranging, yield

$$P(s) = \frac{K_1}{\tau s + 1}M_i(s) - \frac{K_2}{\tau s + 1}M_o(s) - \frac{K_3}{\tau s + 1}P_1(s) \tag{3-5.15}$$

where

$$K_1 = \frac{0.16}{C_2}\frac{\text{psi}}{\%} \qquad K_2 = \frac{C_1}{C_2}\frac{\text{psi}}{\%} \qquad K_3 = \frac{C_3}{C_2}\frac{\text{psi}}{\text{psi}} \qquad \tau = \frac{V}{RT\bar{\rho}C_2}, \quad \text{min} \tag{3-5.16}$$

The desired transfer functions can now be obtained.

$$\frac{P(s)}{M_i(s)} = \frac{K_1}{\tau s + 1} \tag{3-5.17}$$

$$\frac{P(s)}{M_o(s)} = \frac{-K_2}{\tau s + 1} \tag{3-5.18}$$

$$\frac{P(s)}{P_1(s)} = \frac{-K_3}{\tau s + 1} \tag{3-5.19}$$

Because the steady-state values and other process information are known, all gains and the time constant can be evaluated as

$$K_1 = 0.615 \text{ psi/\%} \qquad K_2 = 0.619 \text{ psi/\%} \qquad K_3 = -0.611 \qquad \tau = 5.242 \text{ min}$$

All of the transfer functions are of first order. Fig. 3-5.2 shows the block diagram for this process.

After considering the presentation in Chapter 2 about transfer functions and their response to inputs, and after what has been presented in this chapter, we should have a good feeling for the complete response of any first-order system. We know by analyzing Eq. 3-5.17 that if the signal to the fan increases by 10%, then the pressure in the tank will ultimately change by $+ (10)(K_1)$ psi. We also know that 63.2% of the change, or $0.63(10)(K_1)$, will occur in one time constant. This response is shown graphically in Fig. 3-5.3. Remember that K, is the gain that $M_i(t)$ has on $P(t)$ and that τ gives how fast $P(t)$ responds to a change in $M_i(t)$.

Equation 3-5.18 indicates that if the signal to the valve increases by 5%, then the pressure in the tank will decrease by $(5)K_2$ psi. The negative sign in front of the gain indicates this type of response. Certainly it makes sense that if the signal to the valve increases, opening the valve and thus extracting more gas from the tank, then the pressure in the tank should fall.

Equation 3-5.19 indicates that if the downstream pressure from the valve increases by 3 psi, then the pressure in the tank will decrease by $(3)K_3$ psi. That is, if $P_1(t)$ changes by $+ 3$ psi, $P(t)$ will change by $(3)K_3$. From a physical point of view, however, this does not make any sense. If the downstream pressure increases, then the flow through the valve decreases, increasing the pressure in the tank. Where is the discrepancy? Reviewing the definition of K_3, we see that it depends on C_3, and on the basis of Eq. 3-5.7, it is obvious that C_3 is negative. Thus K_3 is negative ($K_3 = -0.611$), and consequently, the pressure in the tank actually increases.
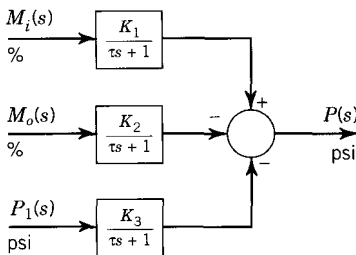
Figure 3-5.2 Block diagram for gas process.
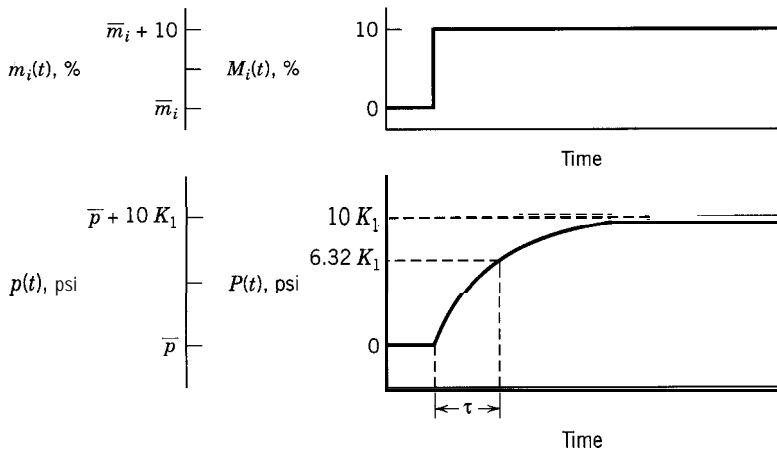
Figure 3-5.3 Response of pressure to signal to fan.

At this point, we should reformulate the procedure for obtaining the transfer functions. This is necessary because we now realize that linearization of nonlinear terms is an important step in the procedure.

1. Write the set of unsteady-state equations that describes the process. This is called modeling.
2. Linearize the model if necessary.
3. Write the steady-state equations at the initial conditions.
4. Subtract the two sets of equations, and define the deviation variables.
5. Obtain the Laplace transform of the linear model in deviation variables.
6. Obtain the transfer functions by solving the Laplace transform explicitly for the transformed output variable(s).

3-6 CHEMICAL REACTORS

3-6.1 Introductory Remarks

The example presented in this section involves a chemical reaction. Because the stoichiometries of the reactions are given in moles, the balances done in chemical reactors are usually mole balances, either on a specific component i or on total moles. The problem however, is that mole balances cannot be written using the equations presented in the introduction to this chapter. That is, taking the reactor as control volume,

$$
\begin{array}{ccc}
\text{Rate of} & \text{Rate of} & \text{Rate of accumulation} \\
\text{component i} - \text{component i} \neq & \text{of component i} \\
\text{into reactor} & \text{out of reactor} & \text{in reactor}
\end{array}
$$

Moles are not necessarily conserved in chemical reactions. Consider, for example, the reaction $2A + B \rightarrow 3S + P$. Under steady-state operation, the moles of reactant A exiting the reactor are not the same as those entering (reactant A is consumed!). Sim-

ilarly, 3 moles of reactants are used, whereas 4 moles of products are formed, so the total moles are not conserved, either. Remember, however, that the **total mass is always conserved.**

Therefore, the mole balance equations must account for the production or depletion of moles due to reaction. The unsteady-state component mole balance that accounts for this production or depletion is written as

$$\begin{matrix} \text{Rate of} & \text{Rate of} & \text{Rate of} & \text{Rate of accumulation} \\ \text{component i} - \text{component i} + \text{production of} = & \text{of component i} \\ \text{into reactor} & \text{out of reactor} & \text{component i} & \text{in reactor} \end{matrix}$$

The rate of production of component i in the reactor is usually given by

$$\text{Rate of production of component } i = \nu_i r_k V, \text{ moles of component i/time}$$

where

ν_i = the stoichiometric coefficient of component i in the reaction
V = volume of reacting mixture
r_k = rate of reaction of the key component in the reaction. This rate (always positive) is usually given in $\dfrac{\text{moles of key component formed/reacted}}{\text{(volume of reacting mixture)(time)}}$.

An important term in these definitions is **key component.** The key component may be any component-reactant or product-in the reaction. The stoichiometric coefficient, ν, of the chosen component is made equal to 1. A positive ν_i indicates production of component i; a negative ν_i indicates depletion of component i. Thus the rate of change of any component i is expressed as a multiple of the rate of reaction of the key component, the volume of the reacting mixture, and the number of moles of component i changing per mole of key component reacting.

To demonstrate further the application of this component mole balance, consider the reaction previously given. Assume that for this reaction, the rate is experimentally determined to be $r_B = kc_A(t)c_B(t), \dfrac{\text{moles of B}}{\text{(volume) (time)}}$, where B is the key component.

Therefore, $\nu_A = -2$, $\nu_B = -1$, $\nu_S = 3$, and $\nu_p = 1$.

An unsteady-state mole balance on component A is written as

$$\begin{matrix} \text{Rate of moles of} \\ \text{A into reactor} \end{matrix} - \begin{matrix} \text{Rate of moles of} \\ \text{A out of reactor} \end{matrix} + (-2)r_B V = \dfrac{dn_A(t)}{dt}$$

where $n_A(t)$ is the moles of A accumulated in the reactor.

An unsteady-state mole balance on component S is written as

$$\begin{matrix} \text{Rate of moles of} \\ \text{S into reactor} \end{matrix} \quad \begin{matrix} \text{Rate of moles of} \\ \text{S out of reactor} \end{matrix} + 3r_B V = \dfrac{dn_S(t)}{dt}$$

where $n_S(t)$ is the moles of S accumulated in the reactor.

The unsteady-state total mole balance is written as

<table>
<tr><td>Rate of total
moles into
reactor</td><td>−</td><td>Rate of total
moles out
of reactor</td><td>+</td><td>Rate of
production
of total moles</td><td>=</td><td>Rate of accumulation
of total moles
in reactor</td></tr>
</table>

and

$$\text{Rate of production of total moles} = \nu_T r_k V, \text{ moles/time}$$

where $\nu_T = \sum_i \nu_i$. For the particular reaction at hand, $\nu_T = \nu_S + \nu_P + \nu_A + \nu_B = 3 + 1 - 2 - 1 = 1$.

Similarly, the energy balance must also account for the energy given off or taken in by the reaction. The energy balance equation is usually written as

<table>
<tr><td>Rate of energy
into reactor</td><td>−</td><td>Rate of energy
out of reactor</td><td>−</td><td>Rate of energy
associated with
the reaction</td><td>=</td><td>Rate of accumulation
of energy in reactor</td></tr>
</table>

A usual reference state for the enthalpies and internal energy is the pure components in the phase (liquid, gas, or solid) in which the reaction takes place, a temperature of 25°C, and the pressure of the system. Using this reference state, we can write

$$\text{Rate of energy associated with reaction} = V r_k \Delta H_r, \text{ energy/time}$$

where **AH,** is the enthalpy of reaction evaluated at 25°C in energy/mole of key component.

3-6.2 Chemical Reactor Example

Consider the chemical reactor system shown in Fig. 3-6.1. The reactor is a vessel where the "well-known" reaction A → B occurs. Let us assume that the reaction occurs at constant volume and temperature. In addition, let's assume constant physical properties and that the reactor is well mixed. The rate of reaction is given by the expression

$$r_A(t) = k c_A^2(t)$$
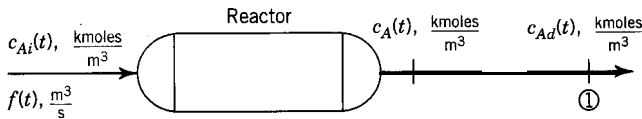
Figure 3-6.1 Isothermal well-mixed chemical reactor.

where

$r_A(t)$ = rate of reaction of component A, kmoles of A/m^3-s
k = constant of reaction, $m^3/kmoles$-s
$c_A(t)$ = concentration of component A in reactor, kmoles of A/m^3

The objective is to develop the mathematical model, find the transfer functions, and draw the block diagram, relating $c_A(t)$ and $c_{Ad}(t)$ to the inputs $f(t)$ and $c_{Ai}(t)$.

Our procedure calls for first developing the mathematical model. Remember, in our way of doing things, those input variables, $f(t)$ and $c_{Ai}(t)$ in this case, are not considered unknowns. The control volume includes the valve and reactor. For this process, an unsteady-state mole balance on component A, of the type presented at the beginning of this section, provides the first equation:

Rate of moles of	Rate of moles of	Rate of change of	Rate of accumulation of
component A into	component A out	component A in	moles of component A
control volume	of control volume	control volume	in control volume

Rate of moles of component A into control volume − Rate of moles of component A out of control volume + Rate of change of component A in control volume = Rate of accumulation of moles of component A in control volume

or, in equation form,

$$f(t)c_{Ai}(t) - f(t)c_A(t) + (-1)Vr_A(t) = V \frac{dc_A(t)}{dt} \tag{3-6.1}$$

1 eq., 2 unk. $[c_A(t), r_A(t)]$

The rate-of-reaction expression provides another equation:

$$r_A(t) = kc_A^2(t) \tag{3-6.2}$$

2 eq., 2 unk.

Equations 3-6.1 and 3-6.2 constitute the mathematical model for this process. Writing this model is the first step in our procedure. The second step calls for linearizing the nonlinear terms in the model.

Linearizing the first two terms of Eq. 3-6.1 and Eq. 3-6.2 around the initial steady-state values off, $\bar{c}_{Ai}$, and $\bar{c}_A$ yields

$$f(t)c_{Ai}(t) \approx \bar{f}\bar{c}_{Ai} + \bar{c}_{Ai}(f(t) - \bar{f}) + \bar{f}(c_{Ai}(t) - \bar{c}_{Ai}) \tag{3-6.3}$$

$$f(t)c_A(t) \approx \bar{f}\bar{c}_A + \bar{c}_A(f(t) - \bar{f}) + \bar{f}(c_A(t) - \bar{c}_A) \tag{3-6.4}$$

$$r_A(t) \approx \bar{r}_A + 2k\bar{c}_A(c_A(t) - \bar{c}_A) \tag{3-6.5}$$

Substituting Eqs. 3-6.3, 3-6.4, and 3-6.5 into Eq. 3-6.1 yields

$$\bar{f}\bar{c}_{Ai} + \bar{c}_{Ai}(f(t) - \bar{f}) + \bar{f}(c_{Ai}(t) - \bar{c}_{Ai}) - \bar{f}\bar{c}_A - \bar{c}_A(f(t) - \bar{f})$$
$$- \bar{f}(c_A(t) - \bar{c}_A) - V\bar{r}_A - 2Vk\bar{c}_A(c_A(t) - \bar{c}_A) = v \frac{dc_A(t)}{dt} \tag{3-6.6}$$

Equation 3-6.6 is the equation that describes the process around the linearization values. We can now proceed to obtain the transfer functions. Writing a mole balance at the initial steady state and subtracting it from Eq. 3-6.6 yield

$$\bar{c}_{Ai}F(t) + \bar{f}C_{Ai}(t) - \bar{c}_AF(t) - \bar{f}C_A(t) - 2k\bar{c}_AVC_A(t) = V\frac{dC_A(t)}{dt} \qquad (3\text{-}6.7)$$

where $F(t) = f(t) - \bar{f}$, $C_{Ai}(t) = c_{Ai}(t) \quad \bar{c}_{Ai}$, and $C_A(t) = c_A(t) - \bar{c}_A$.
From Eq. 3-6.7,

$$C_A(s) = \frac{K_1}{\tau s + 1}F(s) + \frac{K_2}{\tau s + 1}C_{Ai}(s) \qquad (3\text{-}6.8)$$

where

$$K_1 = \frac{\bar{c}_{Ai} - \bar{c}_A}{\bar{f} + 2k\bar{c}_AV}, \text{kmoles/s}$$

$$K_2 = \frac{\bar{f}}{\bar{f} + 2k\bar{c}_AV}$$

$$\tau = \frac{V}{\bar{f} + 2k\bar{c}_AV}, \text{seconds}$$

From Eq. 3-6.8, the desired transfer functions can be obtained. They are

$$\frac{C_A(s)}{F(s)} = \frac{K_1}{\tau s + 1} \qquad (3\text{-}6.9)$$

and

$$\frac{C_A(s)}{C_{Ai}(s)} = \frac{K_2}{\tau s + 1} \qquad (3\text{-}6.10)$$

To obtain the relationships for $c_{Ad}(t)$, assuming ideal plug flow and no reaction occurring in the outlet pipe, we can state

$$c_{Ad}(t) = c_A(t - t_0)$$

or, in terms of deviation variables,

$$C_{Ad}(t) = C_A(t - t_0) \qquad (3\text{-}6.11)$$

and

$$t_0 = \frac{LA_p}{\bar{f}} \qquad (3\text{-}6.12)$$

(a)

(b)
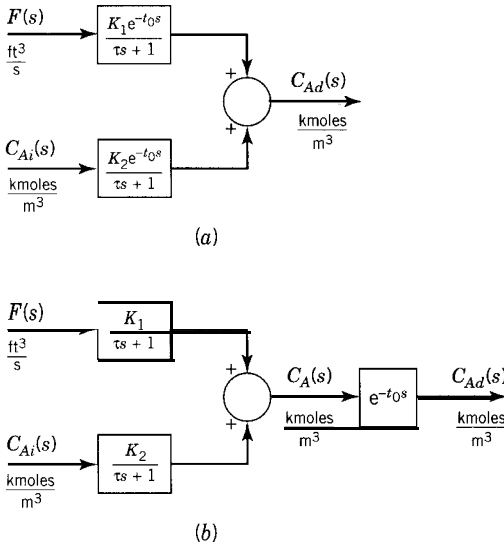
Figure 3-6.2 Block diagram for well-mixed isothermal chemical reactor.

where

t_0 = dead time between the reactor outlet and point 1, seconds
L = distance between the reactor outlet and point 1, m
$A,$ = cross-sectional area of pipe, m^2

The Laplace transform of Eq. 3-6.11 gives

$$C_{Ad}(s) = e^{-t_0s}C_A(s) \tag{3-6.13}$$

Thus, from Eqs. 3-6.9, 3-6.10, and 3-6.13, the final desired transfer functions are

$$\frac{C_{Ad}(s)}{F(s)} = \frac{K_1 e^{-t_0s}}{\tau s + 1} \tag{3-6.14}$$

and

$$\frac{C_{Ad}(s)}{C_{Ai}(s)} = \frac{K_2 e^{-t_0s}}{\tau s + 1} \tag{3-6.15}$$

Figure 3-6.2 shows two different ways to draw the block diagram for this reactor.

3-7 EFFECTS OF PROCESS NONLINEARITIES

A most important characteristic of processes is their linear or nonlinear behavior. To understand what these terms mean and appreciate their significance, consider the ther-

mal process presented in Section 3-2. In this particular process, because the flowf is considered constant, the gains, K_1 and K_2, are constants over the complete operating range. That is, their numerical values (given by Eqs. 3-2.22 and 3-2.23) do not ever change, no matter what the process operating condition. The value of the time constant τ, Eq. 3-2.21, is also constant for this system. The fact that the parameters that describe the characteristics of this process are constants means that the behavior of the process is also constant. That is, the process will behave in the same manner, sensitivity and speed of response, at any operating condition. Processes that exhibit this characteristic are called **linear processes.**

In Section 3-1, we noted that the controller must be tuned, or adapted, to the process to obtain adequate control performance. Because the behavior of a linear process is the same over the complete operating range, if the controller is optimally tuned at one operating condition, it is also optimum at any other operating condition. This is certainly an ideal operation and the one we could hope for.

However, consider now the gas process presented in Section 3-5. In this process the gains $K,$, K_2, and K_3 as given by Eq. 3-5.16 depend on C, , C,, and C_3 and the numerical values of these terms depend on the values of p , p , , and $\bar{m}_o$ around which the lineari-zation of the nonlinear function $f_o(t)$ was done. Therefore, the numerical values of $K,$, K_2, and K_3 also depend on where the linearization was performed. The numerical value of the time constant, as also given in Eq. 3-5.16, also depends on C_2. This means that the values of the terms that describe the process characteristics, and thus the process behavior itself, depend on the operating condition. *The process behavior changes as the operating conditions change!* Processes that exhibit these characteristics are called **nonlinear processes.** Nonlinearity is a characteristic of most chemical processes.

To demonstrate graphically the effect of the process nonlinearities, two different cases are shown. In the first case, the pressure in the tank was allowed to vary between 25 psia and 70 psia, while keeping the process flow constant and maintaining other process conditions at their steady state. The signal to the outlet valve was allowed to vary to keep the process flow constant. This is the case when it is desired to run the process at different pressures even though the process flow has not changed. Figure 3-7.1 shows how K_1, K_2, K_3, and τ vary as the pressure in the tank varies. Figure 3-7.1a shows that $K,$ varies by a factor of 4, and Fig. 3-7.1b shows that K_2 *varies* by a factor of 10. Similarly, Fig. 3-7.1d shows that τ varies by a factor of 5. K_3 is not affected so much, as shown in Fig. 3-7.1c.

Another interesting case occurs when the signal to the fan varies, thus varying the process flow through the tank, while the pressure is kept constant. This could happen when a pressure control system is having to react to upsets, the upset being the signal to the fan in this case. The signal to the outlet valve was allowed to change to match the outlet gas flow to the inlet flow provided by the fan and to keep the pressure in the tank constant. Figure 3-7.2 shows how K_1, K_2, K_3, and τ vary as the signal to the fan varies. All of these figures show the nonlinear characteristics of this simple process.

The nonlinear behavior of processes is very detrimental to their control. As the process behavior changes with operating conditions, the controller should be re-tuned, or re-adapted, to maintain optimum control performance. Often, the best we can do is tune the controller so that its performance is best at the design operating point and acceptable over the expected range of operating conditions; tuning methods are presented in Chapter 6. Techniques have been developed to permit the controller to re-tune itself, auto-
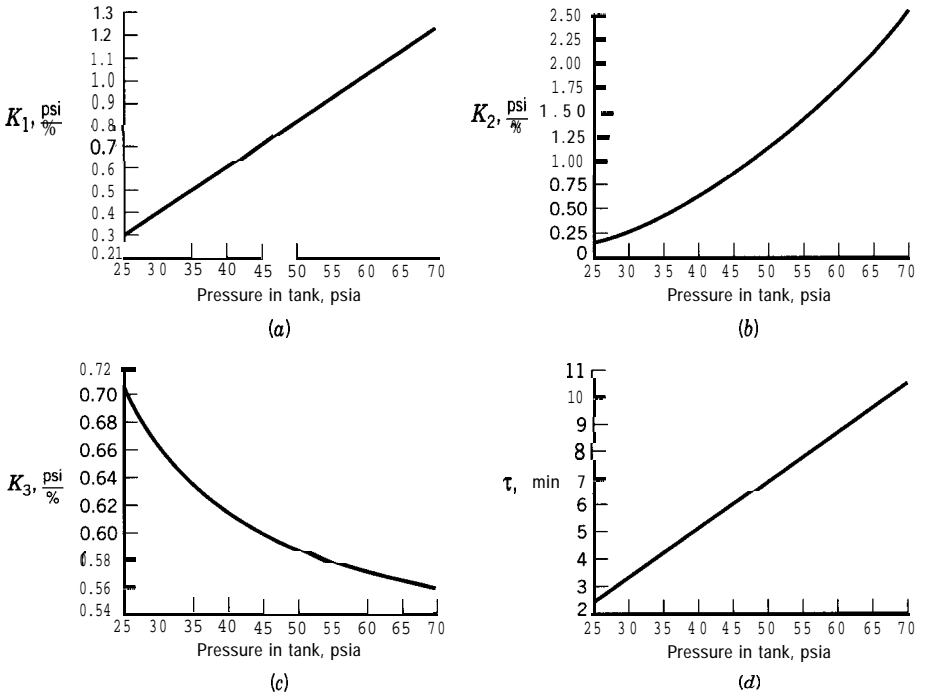
Figure 3-7.1 Gains and time constant as a function of pressure in tank.
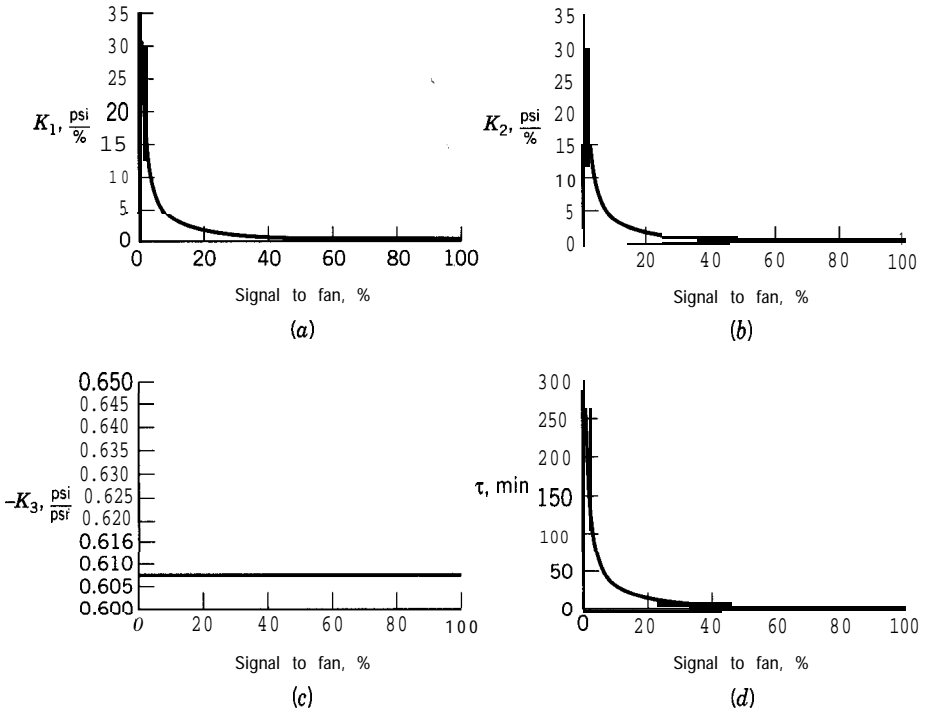
Figure 3-7.2 Gains and time constant as a function of signal to fan.

matically, as the process characteristics change. These techniques are referred to as self-tuning, or adaptive tuning, and are presented in Chapter 15. Computer control systems provide the necessary computing power for the realistic application of the technique.

Although it is not presented in Sections 3-5 and 3-6, the dead time also depends on the operating conditions. Equation 3-3.1 shows that if the process flow varies, then the dead time will also vary. Thus all the terms that describe the process behavior are functions of the operating conditions.

Process nonlinearities are certainly not a desirable characteristic, but they are unfortunately a realistic and very common one.

3-8 ADDITIONAL COMMENTS

It is now important to analyze what we have done from a more "general" point of view. If we look at the form of the transfer functions that have been developed in the different examples (Eqs. 3-2.12, 3-2.25, 3-2.26, 3-5.17, 3-5.18, 3-5.19, 3-6.9, and 3-6.10), we see that they are all of the form

$$\frac{Y(s)}{X(s)} = \frac{K}{\tau s + 1} \tag{3-8.1}$$

where

$Y(s)$ = **Laplace** transform of the output variable
$X(s)$ = **Laplace** transform of the input variable

In Section 2-4, this Eq. 3-8.1 was defined as the standard form of the transfer function for a first-order system. The distinguishing characteristic of this form is that the second term of the denominator is unity. This is the form of all first-order systems regardless of whether they are thermal, fluid, reacting, mechanical, or electrical systems. This is important because it says that the behavior of any system, no matter what type, described by Eq. 3-8.1 is the same; they all respond the same way to forcing functions. The meaning of gain, **K,** and time constant, τ, is the same for all of them.

Sometimes dead time is present, and in this case the transfer function becomes

$$\frac{Y(s)}{X(s)} = \frac{Ke^{-t_0 s}}{\tau s + 1} \tag{3-8.2}$$

Equation 3-8.2 is more general than Eq. 3-8.1.

One of the most important terms in the study of automatic control is the time constant, τ. We have developed several expressions for τ: Eqs. 3-2.9, 3-2.21, and 3-5.16. These equations are all analogous, that is, they are all of the form

$$\tau = \frac{\text{capacitance}}{\text{conductance}} \tag{3-8.3}$$

The capacitance is a measure of the ability of the process to accumulate the quantity conserved (mass or energy). The conductance is a measure of the ability of the process to regulate itself.

For example, for the thermal system of Section 3-2, the time constant, Eq. 3-2.9, is

$$\tau = \frac{V\rho C_v}{f\rho C_p} = \frac{\text{capacitance}}{\text{conductance}}, \frac{\frac{J}{°C}}{\frac{J/s}{°C}}$$

We can also write this expression for τ, assuming that $C_v = C_p$, which is a good assumption for liquids, as

$$\tau = \frac{V}{f}$$

This expression clearly shows the accumulation, V, and flow, f, terms. Table 3-8.1 presents the analogous expressions for the processes shown in this chapter. A process that has not been presented here, but is given as an exercise in Problem 3-1, is that of mixing, or blending. This process is similar to a reacting process in which no reaction occurs. It is expressed by assigning the reaction constant, k, a value of zero in the time constant in Section 3-6. In this case,

$$\tau = \frac{V}{\bar{f} + 2k\overline{C}_A V} = \frac{V}{\bar{f}}, \frac{m^3}{m^3/s}$$

Another comment we wish to make concerns the method used to obtain the desired transfer functions and block diagrams. As you have undoubtedly noticed, the procedure first of all requires a good knowledge of process engineering. The steps followed to obtain the transfer functions were outlined in Section 3-5.

You must also have noted that most of the time, the developed equations that describe the process are nonlinear. We have linearized them to be able to obtain the desired transfer functions. These transfer functions describe the process in a region close to the

Table 3-8.1 Time Constant Analogy for Different Processes

Process	Variable	Time Constant	Capacitance	Conductance
Thermal	Temperature	$\dfrac{V\rho C_v}{f\rho C_p}$ (Eq. 3-2.9)	$V\rho C_v, \dfrac{J}{°C}$	$f\rho C_p, \dfrac{J/s}{°C}$
Thermal	Temperature	$\dfrac{V\rho C_v}{f\rho C_p + UA}$ (Eq. 3-2.21)	$V\rho C_v, \dfrac{J}{°C}$	$f\rho C_p + UA, \dfrac{J/s}{°C}$
Gas	Gas pressure, flow	$\dfrac{V}{RT\bar{\rho} C_2}$ (Eq. 3-5.16)	$\dfrac{V}{RT}, \dfrac{\text{lbmole}}{\text{psia}}$	$\bar{\rho} C_2, \dfrac{\text{lbmole}}{\text{min-psia}}$
Reacting	Concentration	$\dfrac{V}{\bar{f} + 2k\bar{c}_A V}$ (Eq. 3-6.16)	V, m^3	$\bar{f} + 2k\bar{c}_A V, m^3/s$

linearization values. Outside this region, the linearization will "break down" and give erroneous results. The size of the region where the transfer functions are valid depends on the degree of nonlinearity of the process. For a very nonlinear process, the valid region is very close to the linearization values. The region "opens" as the degree of nonlinearity of the process lessens. The only way to obtain an accurate solution from the set of equations, *the mathematical model,* over the complete operating range is by numerical methods, or computer solution. However, this technique does not allow a general analysis of the process dynamics.

Finally, a comment about the response of first-order systems to different types of forcing functions; the responses to step functions, ramp functions, and sinusoidal functions are presented in Section 2-4. The response to a step function is particularly important in process control studies, and thus it has also been shown in this chapter. It is clear in Fig. 3-2.2 that the steepest slope occurs at the beginning of the response. This characteristic is typical of first-order systems.

All the responses shown in this chapter reach a new operating value. That is, the responses to a bounded input are also bounded; the system "regulates" itself to a new value. The majority of processes are of this type and are sometimes referred to as self-*regulating processes.* There are some processes, however, that do not regulate themselves to a new value before they reach an extreme operating condition, These processes are referred to as *non-self-regulating processes,* and examples of them are given in Chapter 4.

3-9 SUMMARY

This chapter began by explaining that from a controls point of view, a "process" is everything except the controller. That is, the process consists of the sensor, transmitter, process unit, valve, and transducer, if present. We noted that as far as the controller is concerned, its controlled variable is the signal it receives from the transmitter (TO). Its manipulated variable is its own output signal (CO), that is, the signal the controller sends out to the final control element. A discussion of why it is necessary to study the process characteristic was included.

The chapter presented the development of mathematical models, transfer functions, and block diagrams for simple processes, not including sensors, transmitters, and valves. All of the processes studied in the chapter are described by first-order ordinary differential equations. The starting point is usually a balance equation. In order to develop the set of equations, we must end up with the same number of independent equations as unknowns. That is why we have stressed the unknowns next to each equation. This should help us keep track of the equations needed to describe the process and develop the model. We shall call this method of writing equations and unknowns the *description method.* [1]

Several other concepts were reviewed and further explained in this chapter. Transfer functions were defined as the ratio of the Laplace-transformed output variable to the Laplace-transformed input variable. The meaning of *transfer functions* was explained: they *fully describe the steady-state and dynamic behavior of the system.* Transfer func-

---

[1] This term was developed by Dr. J. C. Busot at the University of South Florida in his thermodynamics course.

tions indicate how much and how fast processes change. The variables used in the transfer functions are in deviation form.

The transfer functions developed in this chapter are of the general form

$$\frac{Y(s)}{X(s)} = \frac{Ke^{-t_0s}}{\tau s + 1}$$

(3-9.1)

These transfer functions are called first-order-plus-dead-time (FOPDT) transfer functions, or first-order lags or single capacitances. This transfer function contains three parameters: process gain, K; process time constant, τ; and process dead time, t_0. Understanding these parameters is fundamental to the study of process control. The process gain, K, specifies the amount of change of the output variable per unit change in the input variable; it is defined mathematically as follows:

$$K = \frac{\Delta Y}{\Delta X} = \frac{\text{A output variable}}{\text{A input variable}}$$

(3-9.2)

The process time constant, τ, is related to the speed of response of the process once *the process starts to respond* to an input. The time constant was shown, in Chapter 2 and in Eq. 3-2.13, to be the time required for the output variable to reach 63.2% of the total change, AY, when the input variable changes in a step fashion. The slower a process is to respond to an input, the larger its value of τ. The process dead time, t_0, is the time interval between the change in input variable and the time the output variable starts to respond. Therefore, the process time constant and the process dead time are the terms that describe the dynamics of the system; the process gain describes the steady-state characteristics of the system.

We must remember that these three parameters, K, τ, and t_0, are functions of the physical parameters of the process. It was shown that for a linear system, these parameters are constant over the complete operating range of the process. For a nonlinear system, the parameters were shown to be functions of the operating conditions and, consequently, not to be constant over the operating range. This nonlinear characteristic of processes is a most important consideration for their control. This chapter showed how to obtain the aforementioned parameters starting from balance equations. Chapter 7 shows how to evaluate them from process data.

The next chapter shows the development of transfer functions and block diagrams for more complex processes.

REFERENCES

1. Perry, R. H., and D. W. Green. 1984. Chemical Engineering Handbook, 6th ed. New York: McGraw-Hill, pp. 5– 19.

PROBLEMS

3-1. Consider the mixing process shown in Fig. P3-1. You may assume that the density of the input streams and that of the output stream are very similar and that
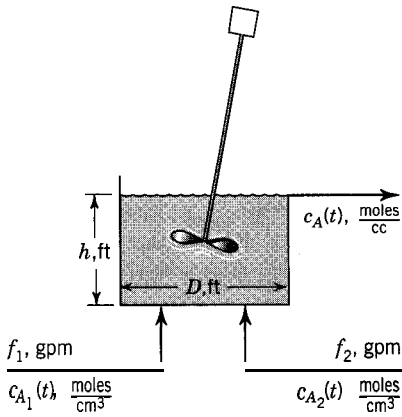
Figure P3-1 Sketch for Problem 3-1.

the flow rates f_1 and f_2 are constant. It is desired to understand how each inlet concentration affects the outlet concentration. Develop the mathematical model, determine the transfer functions, and draw the block diagram for this mixing process. Show the units of all gains and time constants.

3-2. Consider the isothermal reactor shown in Fig. **P3-2.** The rate of reaction is given by

$$r_A(t) = kc_A(t), \text{ moles of } A/(ft^3\text{-min})$$

where k is constant. You may assume that products and reactants are similar in density and all other physical properties. You may also assume that the flow regime between points 2 and 3 is very turbulent (plug flow), minimizing back-mixing. Develop the mathematical model, and obtain the transfer functions relating
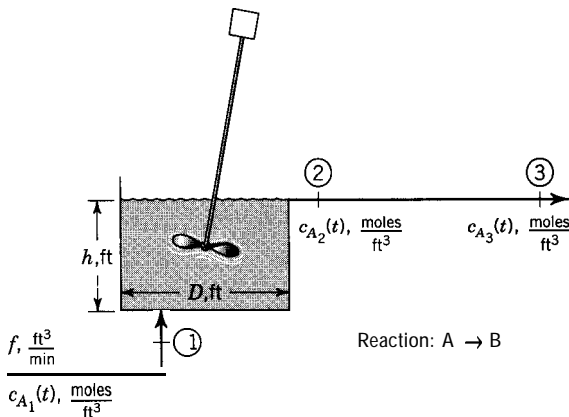
Figure P3-2 Sketch for Problem 3-2.

(a) The concentration of A at point 2 to the concentration of A at point 1.
(b) The concentration of A at point 3 to the concentration of A at point 2.
(c) The concentration of A at point 3 to the concentration of A at point 1.

3-3. A storage tank has a diameter of 20 ft and a height of 10 ft. The output volumetric flow from this tank is given by

$$f_{out}(t) = 2h(t)$$

where h(t) is the height of liquid in the tank. At a particular time, the tank is at steady state with an input flow of 10 ft³/min.

(a) What is the steady-state liquid height in the tank?
(b) If the input flow is ramped up at the rate of 0.1 ft³/min, how many minutes will it take for the tank to overflow?

(Copyright 1992 by the American Institute of Chemical Engineers; reproduced by permission of Center for Chemical Process Safety of AIChE.)

3-4. Consider the temperature sensor sketched in Fig. P3-3. The bulb and its surrounding thermowell are at a uniform temperature, $T_b(t)$, °C, and the surroundings are also at a uniform temperature, T,(t). The exchange of heat between the surroundings and the bulb is given by

$$q(t) = hA[T_s(t) - T_b(t)]$$

where

$$q(t) = \text{heat transfer rate, J/s}$$
$$h = \text{film coefficient of heat transfer, J/s-m}^2\text{-°C}$$
$$A = \text{contact area between the bulb and its surroundings, m}^2$$

Let *M,* kg, be the mass of the bulb and thermowell, and let C_v, J/kg-"C, be its heat capacity. Obtain the transfer function that represents the response of the temperature of the bulb when the surrounding temperature changes. List all assumptions and draw the block diagram for the bulb. Express the time constant and the gain in terms of the bulb parameters. *Note:* The transfer function derived here generally represents the dynamic response of most temperature sensors, regardless of their type.

3-5. Hot water at a rate of 2 liters/min (constant) and temperature $T_h(t)$ is mixed with cold water at a constant rate of 3 liters/min and a constant temperature of 20°C.

$q(t)$

$T_s(t)$, °C

$T_b(t)$
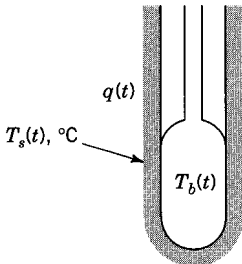
Figure P3-3 Sketch for Problem 3-4.

Both streams flow into a bathtub, but because of carelessness, the water is over-flowing and keeping the bathtub full of water. The volume of the bathtub is 100 liters. Assuming the water in the bathtub is perfectly mixed, derive the differential equation relating the temperature in the bathtub, T(t), to the temperature of the hot water, $T_h(t)$. Obtain the transfer function $\dfrac{T(s)}{T_h(t)}$, and calculate its gain and time constant.

3-6. Process waste water (density = 1000 kg/m³) flows at 500,000 kg/h into a hold-ing pond with a volume of 5000 m³ and then flows from the pond to a river. In-itially, the pond is at steady state with a negligible concentration of pollutants $[x(0) = 0]$. Because of a malfunction in the waste water treating process, the concentration of pollutants in the inlet stream suddenly increases to 500 mass ppm (kg of pollutant per million kg of water) and stays constant at that value (step change).

(a) Assuming a perfectly mixed pond, obtain the transfer function of the pol-lutant concentration in the outlet stream to the concentration of the inlet stream, and determine for how long the process malfunction can go unde-tected before the outlet concentration of pollutants exceeds the regulated maximum value of 350 ppm.

(b) Repeat part (a), assuming that the water flows in plug flow (without mixing) through the pond. Note that this means the pond behaves as a pipe and the response of the concentration is a pure transportation lag.

(c) In both parts (a) and (b), it is assumed that the entire volume of the pond is active. How would your answers be affected if portions of the pond were stagnant and were not affected by the flow of water in and out?

3-7. In Dr. Corripio's home, the hot water line between the water heater and his shower is 1/2 copper tubing (cross-sectional area = 0.00101 ft²) and about 30 ft long. On a cold Baton Rouge morning, Dr. Corripio turned the hot water valve on the shower fully open and got a flow of 2 gal per minute. How long did he have to wait for the hot water to reach the shower (and probably bum him)? Write the transfer function $T_s(s)/T_h(s)$ for the hot water line, where $T_s(t)$ is the temperature at the shower, and $T_h(t)$ is the temperature in the hot water heater, when the hot water valve is opened. Draw the block diagram for the hot water line. What is the transfer function when the hot water valve is closed? Could you predict this from your previous answer?

3-8. Brine from a pond is pumped at 100 ft³/min to a process through a line that has two different diameters, before and after the pump. The inside diameters and lengths of the pipes are as follows:

	Before the Pump	After the Pump
Inside diameter, in.	6.00	5.25
Length, ft	1000	2000

You may assume that the brine does not mix in the pipe. When the concentration in the pond changes, how long does it take for the concentration of the stream entering the process to change? Write the transfer function for the concentration out of the pipe to the concentration in the pond.

3-9. It is desired to model the response of the temperature, T(t), °C, in a fish tank to changes in the heat input from the electric heater, $q(t)$, W; ambient temperature, T,(t), °C; and ambient partial pressure of water in the air, p,(t), Pa, under the following assumptions:

(a) The water in the tank is perfectly mixed.

(b) Transfer of heat and mass to the surroundings is only from the free surface of the water (transfer of heat through the glass sides is negligible).

(c) The overall heat transfer coefficient to the surroundings, U, W/m²-°C, and the overall mass transfer coefficient of water vapor, K_y, kg/s-m²-Pa, are constant.

(d) The physical properties of water (specific heat, c_p, J/kg-°C, and latent heat, λ, J/kg) are constant.

(e) The rate of vaporization of water from the tank is proportional to the difference in partial pressures.

$$w = K_y A[p^o(T) - p_s(t)], \text{kg/s}$$

where $p^o(T)$, Pa, is the vapor pressure of water and is given by Antoine's equation. A, m², is the area of the free surface of the water.

(f) The rate of vaporization is so small that the total mass of water in the tank, M, kg, may be assumed constant.

Obtain the transfer functions that represent the response of the tank temperature when the heat input from the electric heater, the surrounding temperature, and the surrounding water partial pressure change. Draw the block diagram for this system.

3-10. Water is poured at a rate $f_i(t)$, cm³/s, into a cup measuring 6.5 cm in diameter and 10 cm high. The cup has a circular hole in the bottom measuring 0.2 cm in diameter. The velocity of the water through the hole is given, from Bernoulli's equation, by

$$v(t) = \sqrt{2gh(t)}$$

where g is the local acceleration of gravity, 980 cm/s²; and $h(t)$, cm, is the level of the water in the cup. Obtain the transfer function between the level of the water in the cup, H(s), and the inlet flow $F_i(s)$, when the cup is half full of water (h = 5 cm).

3-11. Consider the flash drum shown in Fig. P3-4. Here $z(t)$, $x(t)$, and $y(t)$ are the mole fractions of the more volatile component in the feed, liquid, and vapor streams, respectively. The total mass of liquid and vapor accumulated in the drum, the temperature, and the pressure can all be assumed constant. If equilibrium between the vapor and liquid phases leaving the drum is assumed, then the following relationship between $y(t)$ and $x(t)$ can be established

$$y(t) = \frac{\alpha x(t)}{1 + (a - 1)x(t)}.$$
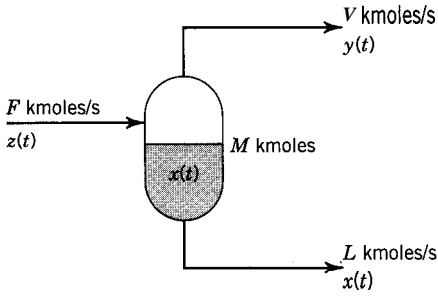
Figure P3-4 Sketch for Problem 3-11.

The steady-state and other process information is $M = 500$ kmoles, $F = 10$ kmoles/s, $L = 5$ kmoles/s, $\alpha = 2.5$, and x(O) $= 0.4$. Obtain the transfer function that relates the outlet liquid composition, x(t), to the feed composition, z(t). Determine also the numerical value of all the terms in the transfer function.

3-12. Figure P3-5 shows a tray of a distillation column. The flow from the tray is given by the Francis weir formula (adapted from Perry, 1984):

$$f_o(t) = 0.415wh^{1.5}(t) \sqrt{2g}$$

where

    $h(t)$ = liquid level on the tray above the top of the weir, ft
    w = width of the weir over which the liquid overflows, ft
    g = local acceleration of gravity (32.2 ft/s^2)

The steady-state inlet flow and process parameters are as follows: tray cross-sectional area $= 11.2$ ft^2, $w = 3.0$ ft, and $f_i(0) = 30$ ft^3/min.

Obtain the transfer functions that relate the height of the water above the weir and the flow from the tray to the inlet flow to the tray. State all assumptions, and calculate the numerical values of the tray time constant and gain. Also draw the complete block diagram relating the variables.
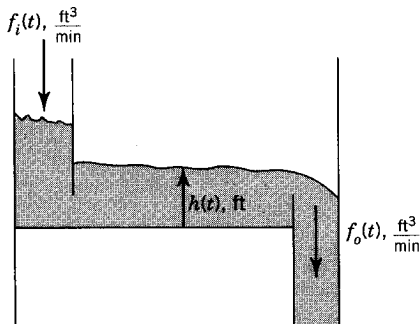
Figure P3-5 Sketch for Problem 3-12.

3-13. Consider an adiabatic, exothermic, perfectly mixed (what else?) chemical reactor where the reaction $A + B \rightarrow C$ (what else?) takes place. Let

ρ = density of reactants and product (constant), $kmoles/m^3$
f = flow of inlet and outlet streams (constant), m^3/s
$T_i(t)$ = inlet temperature, K
$T(t)$ = temperature in reactor, K
AH, = heat of reaction (constant and negative), $J/kmole$
c_p, c_v = heat capacities, $J/kmole\text{-}K$
V = volume of liquid in tank (constant), m^3

The kinetics for the reaction is expressed by the following zeroth-order expression

$$r_A = -k_o e^{-E/RT(t)}$$

where

k_o = frequency factor, $kmoles/m^3\text{-}s$
E = activation energy, J/kmole
R = ideal gas constant, $J/kmole\text{-}K$

Determine the transfer function $\Gamma(s)/\Gamma_i(s)$ for the reactor. Express the time constant and gain in terms of the physical parameters. Under what conditions can the time constant be negative? What would be the consequences of a negative time constant?

3-14. Consider the process shown in Fig. P3-6. The tank is spherical with a radius of 4 ft. The nominal mass flow into and out of the tank is 30,000 lb/hr, the density
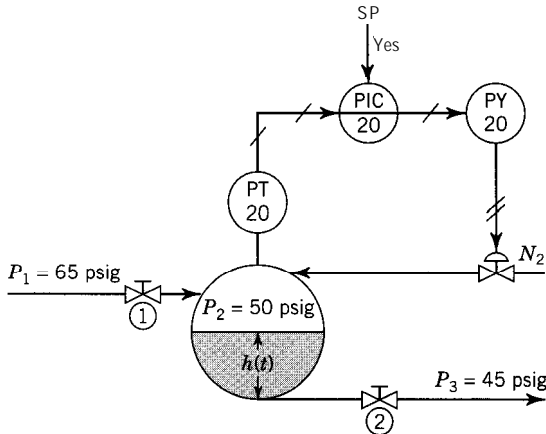
Figure P3-6 Sketch for Problem 3-14.

of the liquid is 70 lb/ft³, and the steady-state level is 5 ft. The volume of a sphere is given by $4\pi r^3/3$. The relation between volume and height is given by

$$V(t) = V_T\left[\frac{h^2(t)[3r - h(t)]}{4r^3}\right]$$

and the flows through the valves are given by

$$w(t) = 500 \ C_v vp(t)\sqrt{G_f\Delta P(t)}$$

where

r = radius of sphere, ft
$V(t)$ = volume of liquid in tank, ft³
V_T = total volume of tank, ft³
$h(t)$ = height of liquid in tank, ft
$w(t)$ = mass flow rate, lb/hr
C_v = valve coefficient, gpm/(psi$^{1/2}$)
C_{v_1} = 20.2 gpm/(psi$^{1/2}$) and C_{v_2} = 28.0 gpm/(psi$^{1/2}$)
$\Delta P(t)$ = pressure drop across valve, psi
G_f = specific gravity of fluid
$vp(t)$ = valve position, a fraction of valve opening

The pressure above the liquid level is maintained constant at a value of 50 psig. Obtain the transfer functions that relate the level of liquid in the tank to changes in the positions of valves 1 and 2. Also, plot the gains and time constants versus different operating levels while keeping the valve positions constant.

3-15. Consider the heating tank shown in Fig. P3-7. A process fluid is being heated in the tank by an electrical heater. The rate of heat transfer, $q(t)$, to the process fluid is related to the signal, $m(t)$, by

$$q(t) = am(t)$$

You may assume that the heating tank is well insulated, that the fluid is well

$f(t), \frac{m^3}{s}$
$\rho\frac{kg}{m^3}$
$T(t), °C$
$C_p, \frac{J}{kg\text{-}°C}$

$f(t), \frac{m^3}{s}$
$\rho\frac{kg}{m^3}$
$T_i(t), °C$
$C_p, \frac{J}{kg\text{-}°C}$

Heater

Tank

$m(t), \%$
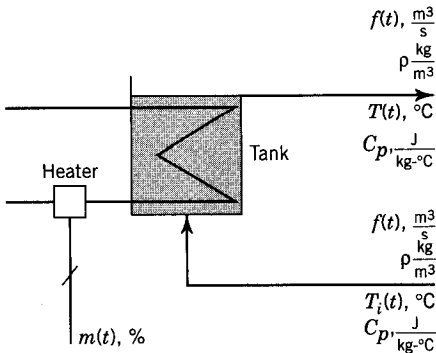
Figure P3-7 Sketch for Problem 3-15.
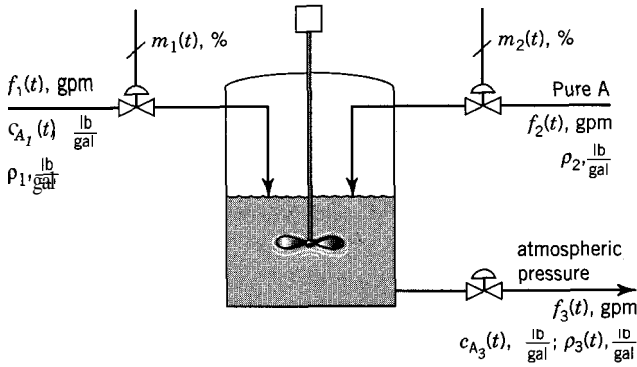
Figure **P3-8** Sketch for Problem 3-16.

mixed in the tank, and that the heat capacity and density of the fluid are constant. Develop the mathematical model that describes how the inlet temperature, $T_i(t)$; the process flow, $f(t)$; and the signal, m(t), affect the outlet temperature T(t). Then determine the transfer functions, and draw the block diagram for this process.

3-16. Consider the mixing process shown in Fig. **P3-8**. The purpose of this process is to blend a stream, weak in component A, with another stream, pure A. The density of stream 1, ρ_1, can be considered constant because the amount of A in this stream is small. The density of the outlet stream is, of course, a function of the concentration and is given by

$$\rho_3(t) = a_3 + b_3 c_{A_3}(t)$$

The flow through valve 1 is given by

$$f_1(t) = C_{v_1} v \rho_1(t) \sqrt{\frac{\Delta P_1}{G_1}}$$

The flow through valve 2 is given by

$$f_2(t) = C_{v_2} v \rho_2(t) \sqrt{\frac{\Delta P_2}{G_2}}$$

Finally, the flow through valve 3 is given by

$$f_3(t) = C_{v_3} \sqrt{\frac{\Delta P_3(t)}{G_3(t)}}$$

The relationship between the valve position and the signal is given by

$$v \rho_1(t) = a_1 + b_1 [m_1(t) - d_1]$$

and

$$v\rho_2(t) = a_2 + b_2[m_2(t) - d_2]$$

where

a,, b_1, d,, a_2, b_2, d_2, a_3, b_3 = known constants

C_{v_1}, C_{v_2}, C_{v_3} = valve coefficients of valve 1, 2, and 3, respectively, $m^3/(s\text{-}psi^{1/2})$

$v\rho_1(t)$, $v\rho_2(t)$ = valve position of valves 1 and 2, respectively, a dimensionless fraction

AP,, ΔP_2 = pressure drop across valves 1 and 2, respectively (constants), psi

$\Delta P_3(t)$ = pressure drop across valve 3, psi

G_1, G_2 = specific gravity of streams 1 and 2, respectively (constants), dimensionless

$G_3(t)$ = specific gravity of stream 3, dimensionless

Develop the mathematical model that describes how the forcing functions m,(t), $m_2(t)$, and $c_{A_1}(t)$ affect $h(t)$ and $c_{A_3}(t)$; determine the transfer functions; and draw the block diagram. Be sure to show the units of all the gains and time constants.

3-17. Consider the tank shown in Fig. P3-9. A 10% (± 0.2%) by weight NaOH solution is being used for a caustic washing process. In order to smooth variations in flow rate and concentration, a 8000-gal tank is being used as surge tank. The steady-state conditions are as follows:

$$\bar{V} = 4000 \; gal \qquad \bar{f}_i = \bar{f}_o = 2500 \; gph \qquad \bar{c}_i = \bar{c}_0 = 10 \; wt\%$$

The tank contents are well mixed, and the density of all streams is 8.8 lbm/gal.
(a) An alarm will sound when the outlet concentration drops to 9.8 wt% (or rises to 10.2 wt%). Assume that the flows are constant.
  (i) Obtain the transfer function relating the outlet concentration to the inlet concentration. Obtain the numerical values of all gains and time constants.

$f_i(t)$, gph
$c_i(t)$, wt%

f_o, gph
$c_o(t)$, wt%
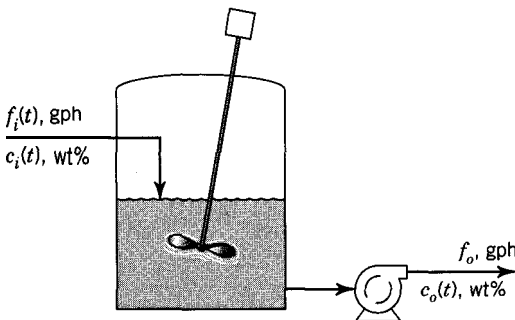
Figure P3-9 Sketch for Problem 3-17.

      (ii) Because of an upset, the inlet concentration, $c_i(t)$, drops to 8% NaOH instantaneously. Determine how long it will take before the alarm sounds.

(b) Consider now that the inlet flow, $f_i(t)$, can vary, whereas the outlet flow is maintained constant at 2500 gph. Therefore, the volume in the tank can also vary.

      (iii) Develop the differential equation that relates the volume in the tank to the flows in and out.

      (iv) Develop the differential equation that relates the outlet concentration of NaOH to the inlet flow and inlet concentration.

      (v) Obtain the transfer function relating the volume in the tank to the inlet flow.

      (vi) Obtain the transfer function relating the outlet concentration to the inlet flow and the inlet concentration. Obtain the numerical values of all gains and time constants.

      (vii) Suppose now that the inlet flow to the tank drops to 1000 gph. Determine how long it takes to empty the tank.

3-18. The blending tank shown in Fig. P3-10 may be assumed to be perfectly mixed. The input variables are the solute concentrations and flows of the inlet streams, $c_1(t)$, $c_2(t)$ [kg/m^3], $f_1(t)$, and $f_2(t)$ [m^3/min]. The volume of liquid in the tank, V [m^3], can be assumed constant, and variation of stream densities with composition may be neglected.

(a) Obtain the transfer functions for the outlet composition C(s), kg/m^3, and outlet flow F(s), m^3/min, to the four input variables, and write the expressions for the time constant and gains of the blender in terms of the parameters of the system.

(b) Draw the block diagram for the blender, showing all transfer functions.

(c) Calculate the numerical values of the time constants and gains for a blender that is initially mixing a stream containing 80 kg/m^3 of solute with a second stream containing 30 kg/m^3 of the solute to produce 4.0 m^3/min of a solution containing 50 kg/m^3 of the solute. The volume of the blender is 40 m^3.

3-19. Draw the block diagram representing the following transfer functions. In each

$f_1(t)$
$c_1(t)$

V

$f(t)$
$c(t)$

$f_2(t)$
$c_2(t)$
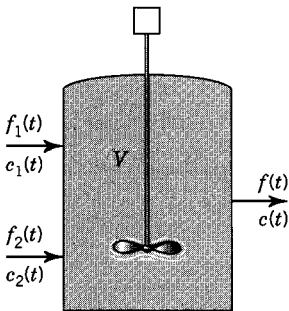
Figure **P3-10** Sketch for Problem 3-18.
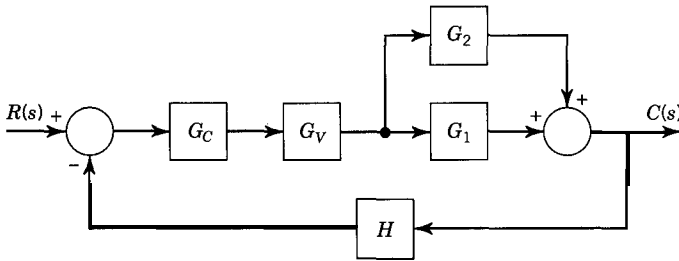
Figure P3-11 Sketch for Problem 3-20.

case, do not do any algebraic manipulations to simplify the transfer functions, but use the rules of block diagram algebra to simplify the diagram if possible.

(a) $Y(s) = \dfrac{K_1}{\tau_1 s + 1} X(s) + \dfrac{K_2}{\tau_2 s + 1} X(s)$

(b) $Y(s) = \dfrac{1}{\tau s + 1} [K_1 F_1(s) - K_2 F_2(s)]$

(c) $Y_1(s) = G_1(s)X(s) + G_3(s)Y_2(s)$
$Y_2(s) = G_2(s)Y_1(s)$

3-20. Determine the transfer function $C(s)/R(s)$ for the system shown in Fig. P3-11.

3-21. Determine the transfer function $C(s)/L(s)$ for the system shown in Fig. P3-12.

3-22. Determine the transfer function $C(s)/R(s)$ for the system shown in Fig. P3-13.

3-23. Obtain the response of a process described by a first-order-plus-dead-time transfer function to the forcing function shown in Fig. P3-14.

3-24. Assume that the following equation describes a certain process

$$\frac{Y(s)}{x(s)} = \frac{3e^{-0.5s}}{5s + 0.2}$$

(a) Obtain the steady-state gain, time constant, and dead time of this process.

(b) The initial condition of the variable y is $y(0) = 2$. For a forcing function as shown in Fig. P3-15, what is the final value of y(t)?

3-25. Obtain the response of a process described by a first-order transfer function to an impulse forcing function.
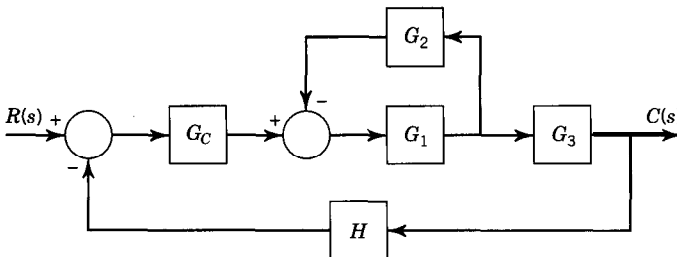
Figure P3-12 Sketch for Problem 3-21.
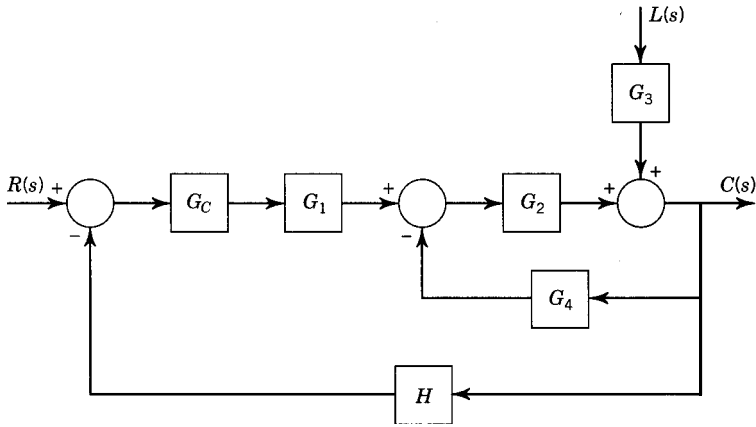
Figure P3-13 Sketch for Problem 3-22.

3-26. A gas detector is used to determine the concentration of flammable gas in a gas stream. Normally the gas concentration is 1% by volume, well below the alarm limit of 4% and the lower flammability limit of 5%. If the gas concentration is above the lower flammability limit, it is flammable. A particular gas detector demonstrates first-order behavior with a time constant of 5 s. At a particular time, the gas stream is flowing at 1 m^3/s through a duct with a cross sectional area of 1 m^2. If the gas concentration suddenly increases from 1% to 7% by volume, how many cubic meters of flammable gas pass the sensor before the alarm is sounded? Is it possible for a plug of flammable gas to pass the detector without the alarm ever being sounded? (Copyright 1992 by the American Institute of Chemical Engineers; reproduced by permission of Center for Chemical Process Safety of AIChE.)

3-27. Consider the chemical reactor shown in Fig. P3-16. In this reactor, an endothermic reaction of the type A + 2B → C takes place. The rate of appearance of A is given by

$$r_A(t) = k_0 e^{-E/RT(t)} c_A(t) c_B(t)$$
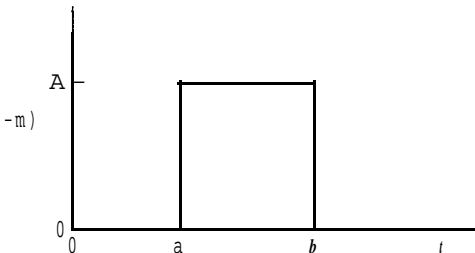
Figure P3-14 Sketch for Problem 3-23.
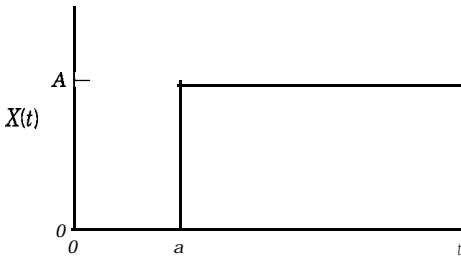
Figure P3-15 Sketch for Problem 3-24.

where

$r_A(t)$ = rate of appearance of A, kmoles of A/(m³-s)
k_0 = frequency factor (constant), m³/(kmole-s)
E = energy of activation (constant), cal/gmole
R = gas law constant, 1.987 cal/(gmole-K)
$T(t)$ = temperature in reactor, K
$c_A(t)$ = concentration of A in reactor, kmoles/m³
$c_B(t)$ = concentration of B in reactor, kmoles/m³
AH, = beat of reaction, J/kmole

The heat input to the reactor is related to the signal to the heater by the expression

$$q(t) = rm_1(t)$$

where

$q(t)$ = heat input to reactor, J/s
r = constant
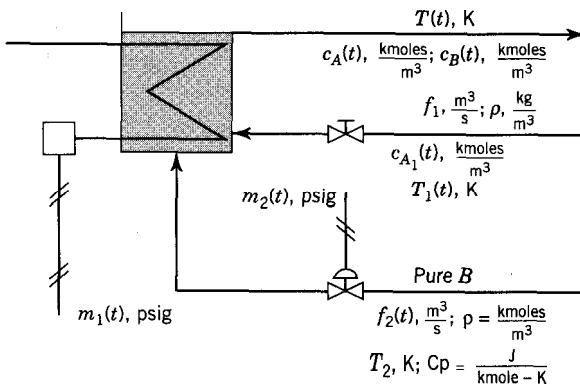
Figure P3-16 Sketch for Problem 3-27.

The flow of pure B through the valve is given by

$$f_2(t) = C_{v_2} vp_2(t) \sqrt{\frac{\Delta P_2}{G_2}}$$

where

C_{v_2} = valve coefficient (constant), $\text{m}^3/(\text{s-psi}^{1/2})$
ΔP_2 = pressure drop across valve (constant), psi
G_2 = specific gravity of B (constant), dimensionless
$vp_2(t)$ = valve position, a fraction

You may assume that the reactor is well insulated and that the physical properties of the reactants and products are similar. The flow rate f_1 can be assumed to be constant. The valve position $vp_2(t)$ is linearly related to the signal $m_2(t)$. Develop the mathematical model that describes the interactions among the input variables m,(t), m,(t), and $c_{Ai}(t)$ and the outlet temperature T(t); determine the transfer functions; and draw the block diagram. Show the units of all gains and time constants.

Chapter 4

Higher-Order Dynamic Systems

The previous chapter investigated the steady-state and dynamic response of simple processes that were all described by first-order ordinary differential equations. The objective of this chapter is to investigate the steady-state and dynamic characteristics of processes described by higher-order ordinary differential equations. Thus the processes shown in this chapter are more intricate; however, they are also more representative of those found in industry.

It is important to remember why we are going through this modeling and analysis procedure. Don't get lost in the mathematics; that's not the reason. Before a control system is designed and implemented, it is imperative to understand the characteristics and behavior of processes. Instilling this understanding is the objective of both chapters. Mathematical methods enable us to quantify the process characteristics.

4-1   NONINTERACTING   SYSTEMS

Higher-order processes and systems are classified as either noninteracting or interacting. This section presents two examples of noninteracting systems, and Section 4-2 presents three examples of interacting systems. These terms are explained in the respective sections.

4-1.1   Noninteracting   Level   Process

Consider the set of tanks shown in Fig. 4-1.1. In this process all the tanks are open to the atmosphere, and the temperature is constant. The openings of the valves remain constant, and the flow of liquid through the valves is given by

$$f(t) = C_v \sqrt{\frac{\Delta P(t)}{G_f}}$$
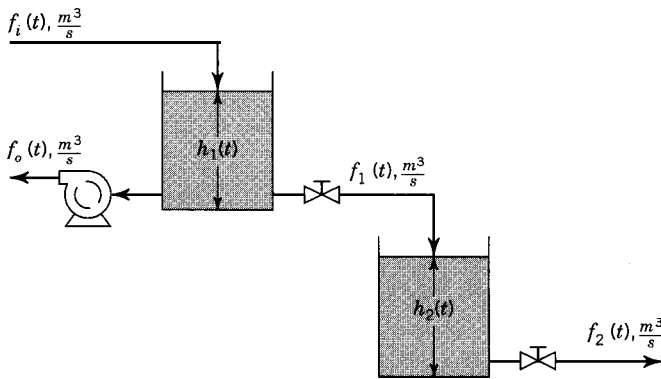
Figure 4-1.1 Tanks in series-noninteracting system.

where

$$f(t) = \text{flow through valve, } m^3/s$$
$$C_v = \text{valve coefficient, } m^3/s\text{-Pa}^{1/2}$$
$$\Delta P(t) = \text{pressure drop across valve, Pa}$$
$$G_f = \text{specific gravity of liquid, dimensionless}$$

Because the tanks are open to the atmosphere and the valves discharge to atmospheric pressure, the pressure drop across each valve is given by

$$\Delta P(t) = P,(t) - P_d = P_a + \rho g h(t) - P_a = \rho g h(t)$$

where

$$P,(t) = \text{upstream pressure from valve, Pa}$$
$$P_d = \text{downstream pressure from valve, Pa}$$
$$P_a = \text{atmospheric pressure, Pa}$$
$$\rho = \text{density of liquid, } kg/m^3$$
$$g = \text{acceleration due to gravity, 9.8 } m/s^2$$
$$h(t) = \text{liquid level in tank, m}$$

Thus the valve equation for this process becomes

$$f(t) = C_v \sqrt{\frac{\Delta P(t)}{G_f}} = C_v \sqrt{\frac{\rho g h(t)}{G_f}} = C_v' \sqrt{h(t)}$$

It is desired to know how the level in the second tank, h,(t), is affected by the inlet flow into the first tank, $f_i(t)$, and by the pump flow, $f_o(t)$. The objective is to develop the mathematical model, determine the transfer functions relating $h_2(t)$ to $f_i(t)$ and $f_o(t)$, and draw the block diagram.

Writing an unsteady-state mass balance around the first tank gives

$$\begin{array}{ccc} \text{Rate of mass into} & \text{Rate of mass out of} & = & \text{Rate of accumulation of} \\ \text{the tank} & \text{the tank} & & \text{mass in tank} \end{array}$$

or, in equation form,

$$\rho f_i(t) - \rho f_1(t) - \rho f_o(t) = \frac{dm_1(t)}{dt}$$

where $m_1(t)$ = mass of liquid accumulated in the first tank, kg. This mass is given by

$$m_1(t) = \rho A_1 h_1(t)$$

where

   $A,$ = cross-sectional area of first tank, uniform throughout, m^2
   $h,(t)$ = liquid level in first tank, m

Then substituting the expression for m,(t) into the mass balance yields

$$\rho f_i(t) - \rho f_1(t) - \rho f_o(t) = \rho A_1 \frac{dh_1(t)}{dt} \tag{4-1.1}$$

1 eq., 2 unk. $[f_1(t), h_1(t)]$

As in the previous chapter, we do not consider the input variables, $f_i(t)$ and $f_o(t)$, un-knowns; it is up to us to specify how they will change.
   The valve expression provides another equation:

$$f_1(t) = C'_{v_1} \sqrt{h_1(t)} \tag{4-1.2}$$

2 eqs., 2 unk.

   Equations 4-1.1 and 4-1.2 describe the first tank. We now proceed to the second tank. An unsteady-state mass balance around the second tank gives

$$\rho f_1(t) - \rho f_2(t) = \rho A_2 \frac{dh_2(t)}{dt} \tag{4-1.3}$$

3 eq., 4 unk. $[f_2(t), h_2(t)]$

Again, the valve expression provides another equation:

$$f_2(t) = C'_{v_2} \sqrt{h_2(t)} \tag{'4-1.4}$$

4 eqs., 4 unk.

   The set of Eqs. 4-1.1 through 4-1.4 describes the process; this set is the mathematical model of the process.

We now proceed to obtain the transfer functions. Because Eqs. 4-1.2 and 4-1.4 are nonlinear, they must first be linearized. This yields

$$f_1(t) \approx \bar{f}_1 + C_1[h_1(t) - \bar{h}_1]$$
(4-1.5)

and

$$f_2(t) \approx \bar{f}_2 + C_2[h_2(t) - \bar{h}_2]$$
(4-1.6)

where

$$C_1 = \frac{\partial f_1(t)}{\partial h_1(t)}\bigg|_{ss} = \frac{1}{2}C'_{v_1}(\bar{h}_1)^{-1/2}, \frac{m^3/s}{m}$$

$$C_2 = \frac{\partial f_2(t)}{\partial h_2(t)}\bigg|_{ss} = \frac{1}{2}C'_{v_2}(\bar{h}_2)^{-1/2}, \frac{m^3/s}{m}$$

Equations 4-1.1, 4-1.3, 4-1.5, and 4-1.6 provide a set of linear equations that describes the process around the linearization values $\bar{h}_1$ and $\bar{h}_2$. Substituting Eq. 4-1.5 into Eq. 4-1.1, substituting Eq. 4-1.6 into Eq. 4-1.3, writing the steady-state mass balances, defining the deviation variables, and rearranging yield

$$\tau_1 \frac{dH_1(t)}{dt} + H_1(t) = K_1 F_i(t) - K_1 F_o(t)$$
(4-1.7)

and

$$\tau_2 \frac{dH_2(t)}{dt} + H_2(t) = K_2 H_1(t)$$
(4-1.8)

where

$$H_1(t) = h_1(t) - \bar{h}_1 \qquad H_2(t) = h_2(t) - \bar{h}_2$$
$$F_i(t) = f_i(t) - \bar{f}_i \qquad F_o(t) = f_o(t) - \bar{f}_o$$
$$F_1(t) = f_1(t) - \bar{f}_1$$

and

$$\tau_1 = \frac{A_1}{C_1}, \text{ seconds} \qquad \tau_2 = \frac{A_2}{C_2}, \text{ seconds}$$

$$K_1 = \frac{1}{C_1}, \frac{m}{m^3/s} \qquad K_2 = \frac{C_1}{C_2}, \text{ dimensionless}$$

Equation 4-1.7 relates the level in the first tank to the inlet and pump flows. Equation 4-1.8 relates the level in the second tank to the level in the first tank.

Taking the Laplace transform of Eqs. 4-1.7 and 4-1.8 and rearranging, we get

$$H_1(s) = \frac{K_1}{\tau_1 s + 1} F_i(s) - \frac{K_1}{\tau_1 s + 1} F_o(s)$$
(4-1.9)

$$H_2(s) = \frac{K_2}{\tau_2 s + 1} H_1(s) \tag{4-1.10}$$

To determine the desired transfer functions, we substitute Eq. 4-1.9 into Eq. 4-1.10, which yields

$$H_2(s) = \frac{K_1 K_2}{(\tau_1 s + 1)(\tau_2 s + 1)} [F_i(s) - F_o(s)] \tag{4-1.11}$$

from which the individual desired transfer functions can be obtained:

$$\frac{H_2(s)}{F_i(s)} = \frac{K_1 K_2}{(\tau_1 s + 1)(\tau_2 s + 1)} \tag{4-1.12}$$

and

$$\frac{H_2(s)}{F_o(s)} = \frac{-K_1 K_2}{(\tau_1 s + 1)(\tau_2 s + 1)} \tag{4-1.13}$$

When the denominator of these two transfer functions is expanded into a polynomial form, the power on the s operator is two. Thus these transfer functions are called *second-order transfer functions* or *second-order lags.* Their development shows that they are "formed" by two first-order transfer functions, or differential equations, in series.

The block diagram for this system can be represented in different forms, as shown in Fig. 4-1.2. The block diagram of Fig. **4-1.2a** is developed by "chaining" Eqs. 4-1.9 and 4-1.10. The diagram shows that the inlet and pump flows initially affect the level in the first tank. A change in this level then affects the level in the second tank. Figure **4-1.2b** shows a more compact diagram. Even though the block diagram of Fig. **4-1.2a**

(a)

(b)
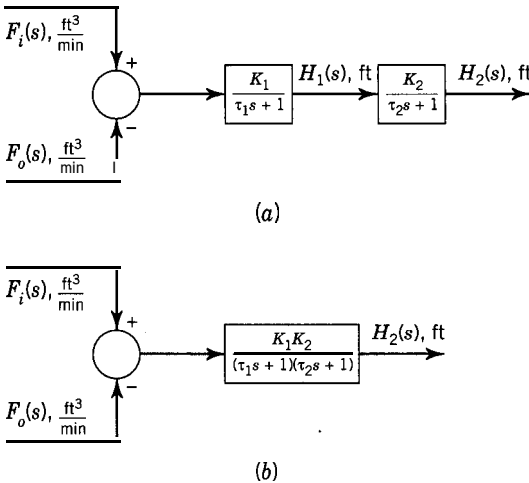
Figure 4-1.2 Block diagrams for two noninteracting tanks.
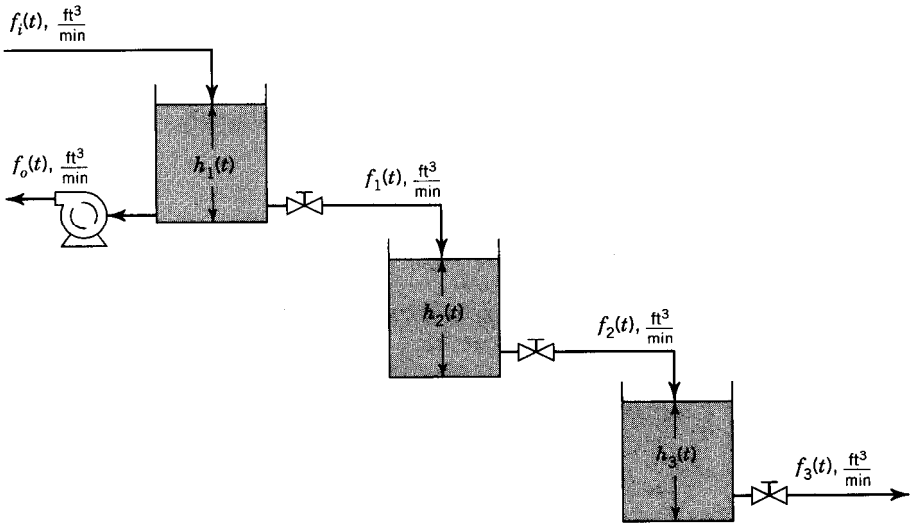
Figure 4-1.3 Tanks in series-noninteracting system.

provides a better description of the physics involved (how things really happen), both diagrams are used without any preference.

Now let us extend the process shown in Fig. 4-1.1 by one more tank, as shown in Fig. 4-1.3. For this new process, the objective is to develop the mathematical model, determine the transfer functions relating the level in the third tank to the inlet flow and to the pump flow, and draw the block diagram.

Because the first two tanks have already been modeled (Eqs. 4-1.1 through 4-1.4), the third tank is now modeled. Writing an unsteady-state mass balance around the third tank results in

$$\rho f_2(t) - \rho f_3(t) = \rho A_3 \frac{dh_3(t)}{dt} \tag{4-1.14}$$

5 eq., 6 unk. $[f_3(t), h_3(t)]$

The valve expression provides the next required equation:

$$f_3(t) = C'_{v_3} \sqrt{h_3(t)} \tag{4-1.15}$$

6 eq., 6 unk.

The new process, Fig. 4-1.3, is now modeled by Eqs 4-1.1 through 4-1.4, 4-1.14, and 4-1.15.

Proceeding as before, we get from Eq. 4-1.14 and the linearized form of Eq. 4-1.15 the equation

$$\tau_3 \frac{dH_3(t)}{dt} + H_3(t) = K_3 H_2(t) \tag{4-1.16}$$

where

$$H_3(t) = h_3(t) - \bar{h}_3$$

$$C_3 = -\frac{\partial f_3(t)}{\partial h_3(t)}\bigg|_{ss} = \frac{1}{2} C'_{v_3}(\bar{h}_3)^{-1/2}, \frac{m^3/s}{m}$$

$$\tau_3 = \frac{A_3}{C_3}, \text{ seconds}$$

$$K_3 = \frac{C_2}{C_3}, \text{ dimensionless}$$

Taking the Laplace transform of Eq. 4-1.16 and rearranging, we obtain

$$H_3(s) = \frac{K_3}{\tau_3 s + 1} H_2(s) \tag{4-1.17}$$

Finally, substituting Eq. 4-1.11 into this equation gives

$$H_3(s) = \frac{K_1 K_2 K_3}{(\tau_1 s + 1)(\tau_2 s + 1)(\tau_3 s + 1)} [F_i(s) - F_o(s)] \tag{4-1.18}$$

from which the following transfer functions are determined:

$$\frac{H_3(s)}{F_i(s)} = \frac{K_1 K_2 K_3}{(\tau_1 s + 1)(\tau_2 s + 1)(\tau_3 s + 1)} \tag{4-1.19}$$

and

$$\frac{H_3(s)}{F_o(s)} = \frac{-K_1 K_2 K_3}{(\tau_1 s + 1)(\tau_2 s + 1)(\tau_3 s + 1)} \tag{4-1.20}$$

When the denominator of these two transfer functions is expanded into a polynomial form, the power on the s operator is three. Thus they are referred to as **third-order transfer functions** or **third order-lags.** Figure 4-1.4 shows a block diagram for this process.

The processes shown in Figs. 4-1.1 and 4-1.3 are referred to as **noninteracting systems** because there is no full interaction between the variables. That is, the level in the first tank affects the level in the second tank, but the level in the second tank does not in
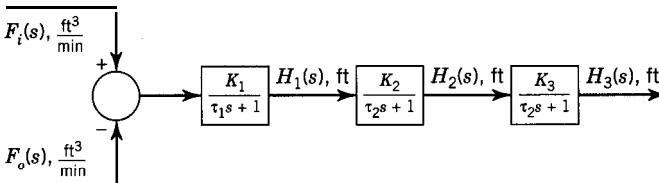
Figure 4-1.4 Block diagram for three noninteracting tanks.

turn affect the level in the first tank. The level in the second tank does not "feed back" to affect the level in the first tank. The cause-and-effect relationship is a one-way path. The same is true for the levels in the second and third tanks.

It is important to remember what we said about transfer functions in Chapters 2 and 3. Transfer functions completely describe the characteristics of linear processes and those around the linearization values for nonlinear processes. Equation 4- 1.20, for example, shows that if the pump flow increases by 10 m^3/s, then the level in the third tank will change by $- 10K_1K_2K_3$ m; that is, it will decrease by $10K_1K_2K_3$ m. The dynamics of the change will depend on τ_1, τ_2, and τ_3. These dynamics are discussed in detail in Section 4-3. Transfer functions quantify the process characteristics, or behavior.

Note that the transfer functions presented in this section were obtained by multiplying first-order transfer functions in series. For example,

$$\frac{H_2(s)}{F_i(s)} = \frac{H_1(s)}{F_i(s)} \cdot \frac{H_2(s)}{H_1(s)}$$

and

$$\frac{H_3(s)}{F_o(s)} = \frac{H_1(s)}{F_o(s)} \cdot \frac{H_2(s)}{H_1(s)} \cdot \frac{H_3(s)}{H_2(s)}$$

In general, this is the case for noninteracting systems only. It can be generalized by writing

$$G(s) = \prod_{i=1}^{n} G_i(s) \qquad \qquad \textbf{(4-1.21)}$$

where

      n = number of noninteracting systems in series
   $G(s)$ = transfer function relating the output from the last system, the
           nth system, to the input to the first system
   $G_i(s)$ = the individual transfer function of each system

Remember, Eq. 4-1.21 is true only for noninteracting systems.

4-1.2   Thermal Tanks in Series

Consider the set of tanks shown in Fig. 4-1.5. The first tank provides some mixing and residence time to stream A. Tank 2 provides mixing of streams A and B. Let us assume that the volumetric flows of these streams, f_A and f_B, are constant; that the density and heat capacity, ρ and C_p, of the streams are equal to each other and are constants; that because the fluids are liquid, $C_p = C_v$; that the two tanks are next to each other; and, finally, that the heat losses to the surroundings and the paddle work are negligible. The volumes of tanks 1 and 2 are V_1 and V_2, respectively.

It is desired to know how the outlet temperature from the second tank, $T_4(t)$, is affected by the inlet temperature of stream A, $T_1(t)$ and by that of stream B, $T_3(t)$. For
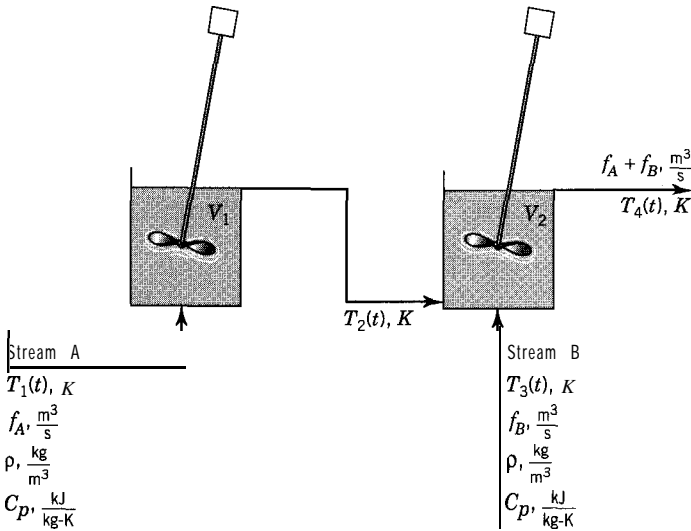
Figure 4-1.5 Thermal tanks in series-noninteracting system.

this process let us develop the mathematical model, determine the transfer functions that relate $T_4(t)$ to T,(t) and $T_3(t)$, and draw the block diagram.

The flows and densities are constants, so the mass accumulated in each tank is also constant. Therefore, a total mass balance around both tanks indicates that the total flow out of the second tank is equal to the sum of the individual inlet streams, or $f_A + f_B$.

We start by writing an unsteady-state energy balance on the contents of the first tank:

$$f_A \rho h_1(t) - f_A \rho h_2(t) = V_1 \rho \frac{du_2(t)}{dt}$$

where

h(t) = specific enthalpy, kJ/kg
$u(t)$ = specific internal energy, kJ/kg

Or, in terms of temperature, using as reference state for h(t) and u(t) the components in the liquid phase at 0 K,

$$f_A \rho C_p T_1(t) - f_A \rho C_p T_2(t) = V_1 \rho C_p \frac{dT_2(t)}{dt} \tag{4-1.22}$$

1 eq., 1 unk. $[T,(t)]$

Another unsteady-state energy balance on the contents of the second tank yields

$$f_A \rho C_p T_2(t) + f_B \rho C_p T_3(t) - (f_A + f_B) \rho C_p T_4(t) = V_2 \rho C_p \frac{dT_4(t)}{dt} \tag{4-1.23}$$

2 eq., 2 unk. $[T_4(t)]$

Equations 4-1.22 and 4-1.23 are the mathematical model that relates the output variable, $T_4(t)$, to the inputs of interest, $T_1(t)$ and $T_3(t)$.

To develop the transfer functions and block diagrams, we first realize that this model is a set of linear equations and that, accordingly, there is no need for linearization. Thus we proceed by writing the steady-state energy balances, defining deviation variables, taking Laplace tranforms, and rearranging to yield from Eq. 4-1.22

$$\Gamma_2(s) = \frac{1}{\tau_1 s + 1} \Gamma_1(s) \tag{4-1.24}$$

and from Eq. 4-1.23

$$\Gamma_4(s) = \frac{K_1}{\tau_2 s + 1} \Gamma_2(s) + \frac{K_2}{\tau_2 s + 1} \Gamma_3(s) \tag{4-1.25}$$

where

$$K_1 = \frac{f_A}{f_A + f_B}, \quad \text{dimensionless} \qquad K_2 = \frac{f_B}{f_A + f_B}, \quad \text{dimensionless}$$

$$\tau_1 = \frac{V_1}{f_A} \text{ seconds} \qquad \tau_2 = \frac{V_2}{f_A + f_B}, \text{ seconds}$$

Substituting Eq. 4-1.24 into Eq. 4-1.25 yields

$$\Gamma_4(s) = \frac{K_1}{(\tau_1 s + 1)(\tau_2 s + 1)} \Gamma_1(s) + \frac{K_2}{\tau_2 s + 1} \Gamma_3(s) \tag{4-1.26}$$

from which the two required transfer functions are obtained:

$$\frac{\Gamma_4(s)}{\Gamma_1(s)} = \frac{K_1}{(\tau_1 s + 1)(\tau_2 s + 1)} \tag{4-1.27}$$

and

$$\frac{\Gamma_4(s)}{\Gamma_3(s)} = \frac{K_2}{\tau_2 s + 1} \tag{4-1.28}$$

Equation 4-1.27 is the transfer function relating the outlet temperature to the inlet temperature of stream A; it is a second-order transfer function. Equation 4-1.28 is the transfer function relating the outlet temperature to the inlet temperature of stream B; it is a first-order transfer function. The block diagram is shown in Fig. 4-1.6.

On the basis of what we learned in Chapters 2 and 3, we know that the dynamic response of $\Gamma_4(t)$ to changes in $r_1(t)$, expressed by Eq. 4-1.27, is different from the response of $\Gamma_4(t)$ to changes in $r_1(t)$, expressed by Eq. 4-1.28. The reader must also try to understand this difference by "looking at" the physical system. If $\Gamma_1(t)$ changes, it
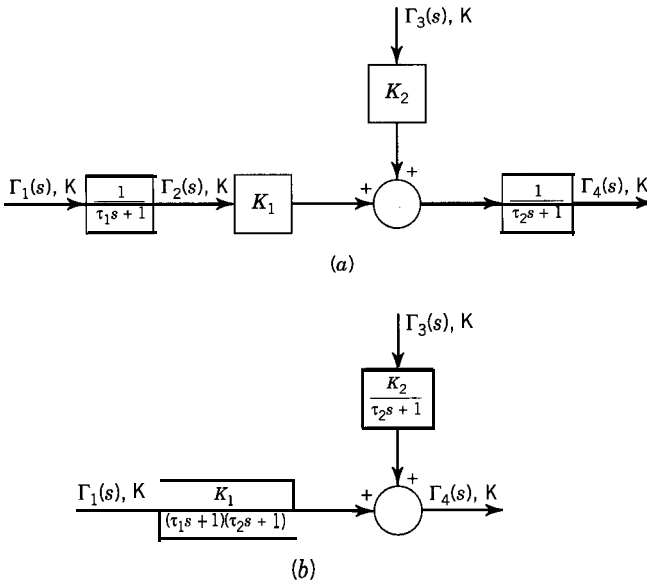
Figure 4-1.6 Block diagrams for thermal tanks in series.

must affect $\Gamma_2(t)$ before $\Gamma_4(t)$ starts to feel the effect; this is shown in Fig. 4-1.6*a*. However, if $\Gamma_3(t)$ changes, it will start affecting r,(t) right away. r,(t) will respond more slowly to changes in r,(t) than to changes in r,(t). How much slower the response is given by the time constant τ_1, which, as shown by its definition, is related to the tank volume and the flow.

We have simplified this process by the assumptions made. However, you may want to think how the development-and indeed the final form-of the transfer functions would be affected by removing some of the assumptions. For example, what if a long pipe exists between the tanks? And what if we allow variations in the flows of streams A and B?

4-2 INTERACTING SYSTEMS

Interacting systems are more frequently encountered in industry than noninteracting systems; this section presents three examples. The differences in dynamic response between the noninteracting and interacting systems are also presented.

4-2.1 Interacting Level Process

Let us rearrange the tanks of Fig. 4-1.1 to result in the new process shown in Fig. 4-2.1. In this case the pressure drop, $\Delta P(t)$, across the valve between the two tanks is given by

$$\Delta P(t) = P_u(t) - P_d(t) = [P_a + \rho g h_1(t)] - [P_a + \rho g h_2(t)] = \rho g[h_1(t) - h_2(t)]$$
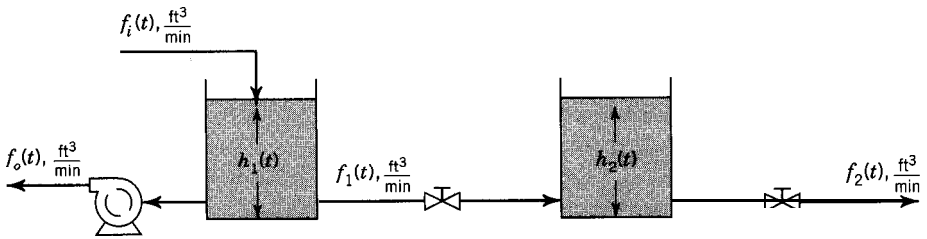
Figure 4-2.1 Tanks in series-interacting system.

Substituting this pressure drop expression into the valve equation results in

$$f_1(t) = C_{v_1} \sqrt{\frac{\Delta P(t)}{G_f}} = C_{v_1} \sqrt{\frac{\rho g[h_1(t) - h_2(t)]}{G_f}}$$

$$= C'_{v_1} \sqrt{h_1(t) - h_2(t)}$$

This new process is referred to as an *interacting system.* The valve equation shows that the flow between the two tanks depends on the levels in *both* tanks, each affecting the other. That is, the level in the first tank affects the level in the second tank, and at the same time, the level in the second tank affects that in the first tank. Each element of the process affects the other. The cause-and-effect relationship is a two-way path.

We are still interested in determining how the level in the second tank is affected by the flow into the first tank and by the pump flow. Let us develop the mathematical model, determine the transfer functions, and draw the block diagram for this new process.

We start by writing an unsteady-state mass balance around the first tank; this is given by Eq. (4-1.1).

$$\rho f_i(t) - \rho f_1(t) - \rho f_o(t) = \rho A_1 \frac{dh_1(t)}{dt} \qquad (4\text{-}1.1)$$

1 eq., 2 unk. $[f_1(t), h_1(t)]$

The valve equation provides the next equation:

$$f_1(t) = C'_{v_1} \sqrt{h_1(t) - h_2(t)} \qquad (4\text{-}2.1)$$

2 eq., 3 unk. $[h_2(t)]$

Another independent equation is still needed. An unsteady-state mass balance around the second tank, Eq. 4-1.3, provides the needed equation:

$$\rho f_1(t) - \rho f_2(t) = \rho A_2 \frac{dh_2(t)}{dt} \qquad (4\text{-}1.3)$$

3 eq., 4 unk. $[f_2(t)]$

The expression for the flow $f_2(t)$ is given by Eq. 4-1.4.

$$f_2(t) = C'_{v_2}\sqrt{h_2(t)} \qquad \text{(4-1.4)}$$

4 eq., 4 unk.

Equations 4-1.1, 4-2.1, 4-1.3, and 4- 1.4 constitute the mathematical model.

We continue with the usual procedure to obtain the transfer functions and block diagram. Because Eq. (4-2.1) is nonlinear, it is linearized as

$$f_1(t) \approx \bar{f}_1 + C_4[h_1(t) - \bar{h}_1] - C_4[h_2(t) - \bar{h}_2] \qquad \text{(4-2.2)}$$

where

$$C_4 = \left.\frac{\partial f_1(t)}{\partial h_1(t)}\right|_{ss} = -\left.\frac{\partial f_1(t)}{\partial h_2(t)}\right|_{ss} = \frac{1}{2}C'_{v_1}(\bar{h}_1 - \bar{h}_2)^{-1/2}, \frac{m^3/s}{m}$$

Eq. 4-1.4 is linearized as given by Eq. 4-1.6.

$$f_2(t) \approx \bar{f}_2 + C_2[h_2(t) - \bar{h}_2] \qquad \text{(4-1.6)}$$

Equations 4-1.1, 4-1.3, 4-2.2, and 4-1.6 provide the set of linear equations that describe the process around the linearization values $\bar{h}_1$ and $\bar{h}_2$.

Substituting Eq. 4-2.2 into Eq. 4-1.1, writing the steady-state mass balance around the first tank, defining deviation variables, taking **Laplace** transforms, and rearranging yield

$$H_1(s) = \frac{K_4}{\tau_4 s + 1}[F_i(s) - F_o(s)] + \frac{1}{\tau_4 s + 1}H_2(s) \qquad \text{(4-2.3)}$$

where

$$K_4 = \frac{1}{C_4}, \frac{m}{m^3/s} \qquad \tau_4 = \frac{A_1}{C_4}, \text{ seconds}$$

Following the same procedure for the second tank gives

$$H_2(s) = \frac{K_5}{\tau_5 s + 1}H_1(s) \qquad \text{(4-2.4)}$$

where

$$K_5 = \frac{C_4}{C_4 + C_2}, \text{ dimensionless} \qquad \tau_5 = \frac{A_2}{C_4 + C_2}, \text{seconds}$$

Finally, substituting Eq. 4-2.3 into Eq. 4-2.4, we get

$$H_2(s) = \frac{K_4 K_5}{(\tau_4 s + 1)(\tau_5 s + 1)}[F_i(s) - F_o(s)] + \frac{K_5}{(\tau_4 s + 1)(\tau_5 s + 1)}H_2(s)$$

and rearranging yields

$$H_2(s) = \frac{\dfrac{K_4 K_5}{1 - K_5}}{\dfrac{(\tau_4 \tau_5)}{(1 - K_5)^2} s^2 + \dfrac{(\tau_4 + \tau_5)}{(1 - K_5)} s + 1} [F_i(s) - F_o(s)]$$

(4-2.5)

from which the desired transfer functions are obtained:

$$\frac{H_2(s)}{F_i(s)} = \frac{\dfrac{K_4 K_5}{1 - K_5}}{\dfrac{(\tau_4 \tau_5)}{(1 - K_5)} s^2 + \dfrac{(\tau_4 \tau_5)}{(1 - K_5)} s + 1}$$

(4-2.6)

and

$$\frac{H_2(s)}{F_o(s)} = \frac{\dfrac{- K_4 K_5}{1 - K_5}}{\dfrac{(\tau_4 \tau_5)}{(1 - K_5)} s^2 + \dfrac{(\tau_4 + \tau_5)}{(1 - K_5)} s + 1}$$

(4-2.7)

These transfer functions are of second order. Block diagrams depicting this interacting
process are shown in Fig. 4-2.2. Fig. 4-2.2a is developed directly from Eq. 4-2.5. Fig.
4-2.2b is developed by "chaining" Eqs. 4-2.3 and 4-2.4. Note also that Fig. 4-2.2a can
be obtained directly from Fig. 4-2.26 by using the "positive feedback" rule of block
diagrams presented in Chapter 3. The interacting nature of this process is clearly shown
in Fig. 4-2.2b. The figure shows that $H_i(s)$ is the input to obtain $H_i(s)$ but also that

(a)

(b)
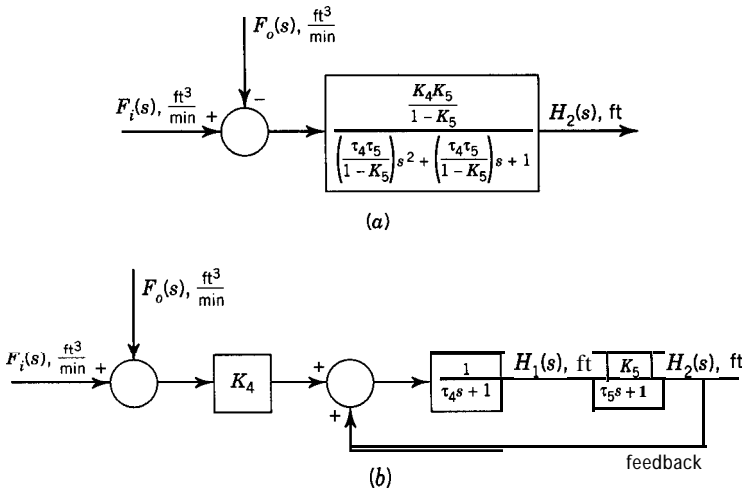
Figure 4-2.2 Block diagrams for an interacting two-tanks system.

(a)

(b)
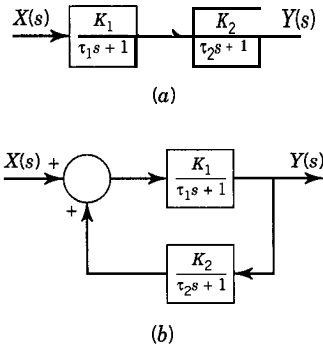
Figure 4-2.3 (a) Block diagram of a noninteracting system. *(b)* Block diagram of an interacting system.

$H_2(s)$ is another input to obtain H,(s), as indicated by the "feedback path." Oftentimes, we referred to this type of system as "interacting lags."

At this time, there are several things we can learn by comparing the transfer functions of the interacting and noninteracting systems. Consider Fig. 4-2.3, which shows a block diagram of a noninteracting system and one of an interacting system. For the noninteracting system, the transfer function is

$$\frac{Y(s)}{X(s)} = \frac{K_1 K_2}{(\tau_1 s + 1)(\tau_2 s + 1)} \qquad (4\text{-}2.8)$$

As presented in Section 2-5, the "effective" time constants are the negatives of the reciprocals of the roots of the denominator of the transfer function. For the foregoing transfer function, the effective time constants are equal to the individual τ values; that is, $\tau_{1\text{eff}} = \tau_1$ and $\tau_{2\text{eff}} = \tau_2$.

For the interacting system, the transfer function is

$$\frac{Y(s)}{X(s)} = \frac{\dfrac{K_1}{(\tau_1 s + 1)}}{1 - \dfrac{K_1 K_2}{(\tau_1 s + 1)(\tau_2 s + 1)}} \qquad (4\text{-}2.9)$$

or

$$\frac{Y(s)}{X(s)} = \frac{K_1(\tau_2 s + 1)}{(\tau_1 s + 1)(\tau_2 s + 1) - K_1 K_2} = \frac{K_1(\tau_2 s + 1)}{\tau_1 \tau_2 s + (\tau_1 + \tau_2)s + (1 - K_1 K_2)} \qquad (4\text{-}2.10)$$

The roots of the denominator are

$$\text{Roots} = \frac{-(\tau_1 + \tau_2) \pm \sqrt{(\tau_1 + \tau_2)^2 - 4\tau_1 \tau_2(1 - K_1 K_2)}}{2\tau_1 \tau_2} \qquad (4\text{-}2.11)$$

or, making use of the assumption $\tau_1 = \tau_2 = \tau$,

$$\text{Roots} = \frac{-(1 + \sqrt{K_1 K_2})}{\tau} \frac{-(1 - \sqrt{K_1 K_2})}{\tau}$$

from which the "effective" time constants for the interacting system can be obtained as

$$\tau_{1_{\text{eff}}} = \frac{\tau}{1 + \sqrt{K_1 K_2}} \qquad \text{and} \qquad \tau_{2_{\text{eff}}} = \frac{\tau}{1 - \sqrt{K_1 K_2}}$$

The ratio of these two terms is

$$\frac{\tau_{2_{\text{eff}}}}{\tau_{1_{\text{eff}}}} = \frac{1 + \sqrt{K_1 K_2}}{1 - \sqrt{K_1 K_2}}$$

which is a number greater than 1 even though $\tau_1 = \tau_2$! This result clearly shows that the larger τ the interacting system "experiences," $\tau_{2_{\text{eff}}}$, is larger than any individual τ.

The following observations, conclusions, and comments are related to this analysis and to the general subject of higher-order systems.

1. Most times the "effective" time constants are real, yielding a non-oscillatory response to step change in input. The roots are real if in, Eq. 4-2.11,

$$(\tau_1 + \tau_2)^2 \quad 4\tau_1 \tau_2 (1 \quad K_1 K_2) > 0$$

or

$$\tau_1^2 + 2\tau_1 \tau_2 + \tau_2^2 - 4\tau_1 \tau_2 + 4\tau_1 \tau_2 K_1 K_2 > 0$$

or

$$\tau_1^2 \quad 2\tau_1 \tau_2 + \tau_2^2 + 4\tau_1 \tau_2 K_1 K_2 > 0$$

or

$$(\tau_1 \quad \tau_2)^2 + 4\tau_1 \tau_2 K_1 K_2 > 0$$

and this is true if $\tau_1 \tau_2 K_1 K_2 > 0$. Because for most cases $\tau_1 > 0$, $\tau_2 > 0$, and $K_1 K_2 > 0$, the roots are real.

The exception to the above statement is the exothermic continuous stirred tank reactor where sometimes one of the τ values is negative. Refer to Section 4-2.3 where a reactor is presented and a τ is negative. Section 4-4.2 also presents another reactor and shows the oscillatory response.

Shinskey (1988) points out that for interacting systems, the higher the interaction, the more different the two effective time constants are, and therefore the more controllable the process.

2. In Chapter 2 and in the present chapter, we have defined and used several times the term *effective time constant.* Let's discuss this term a bit further.

Chapters 2 and 3 showed that when the input to a first-order system changes in a step fashion, the time constant (τ) is the time required for the system to reach 63.2% of its total change. This definition applies to first-order systems. In higher-order systems, there is no one time constant. That is, we cannot say that any one of the τ values in a higher-order system represents the time to reach 63.2% of the total change. However, the values of τ in the transfer functions of these systems are still an indication of the dynamics of the system. The slower the system, the larger these values, and the faster the system, the smaller they are. This is why *we* use the term *effective time constant* instead of just *time constant;* we still use the same representation. Often, in everyday conversation, we drop the word *effective.* What is important to remember is that τ is a parameter of the system related to its dynamics; that for first-order systems it has a definite definition, and that for higher-order systems it is only an indication.

4-2.2 Thermal Tanks with Recycle

Consider the process shown in Fig. 4-2.4. This process is essentially the same one described in Section 4-1.2 except that a recycle stream to the first tank has been added. Let us suppose that this recycle stream is a constant 20% of the total flow out from the process. In addition, let us accept the same assumptions as in Section 4-1.2.
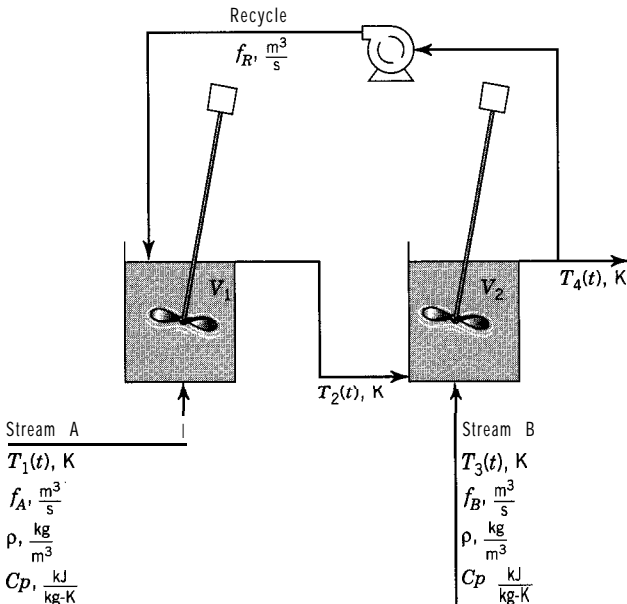
Figure 4-2.4 Thermal tanks with recycle.

It is required to know how the outlet temperature from the second tank, $T_4(t)$, responds to changes in the inlet temperatures of streams A and B. Develop the mathematical model, determine the transfer functions that relate $T_4(t)$ to T,(t) and $T_3(t)$, and draw the block diagram for this process.

As in Section 4-1.2, we start by writing an unsteady-state energy balance on the contents of the first tank.

$$f_A \rho C_P T_1(t) + 0.2(f_A + f_B)\rho C_P T_4(t) - [f_A + 0.2(f_A + f_B)]\rho C_P T_2(t) = V_1 \rho C_P \frac{dT_2(t)}{dt}$$

$$(4\text{-}2.12)$$

1 eq., 2 unk. $[T_2(t), T_4(t)]$

Then we write an unsteady-state energy balance on the contents of the second tank.

$$[f_A + 0.2(f_A + f_B)]\rho C_P T_2(t) + f_B \rho C_P T_3(t) - 1.2(f_A + f_B)\rho C_P T_4(t) = V_2 \rho C_P \frac{dT_4(t)}{dt}$$

$$(4\text{-}2.13)$$

2 eq., 2 unk

The mathematical model for this process is given by Eqs. 4-2.12 and 4-2.13.

To obtain the required transfer functions and block diagram, we proceed in the usual way and obtain from Eq. 4-2.12

$$\Gamma_2(s) = \frac{K_1}{\tau_1 s + 1} \Gamma_1(s) + \frac{K_2}{\tau_1 s + 1} \Gamma_4(s) \tag{4-2.14}$$

and from Eq. 4-2.13

$$\Gamma_4(s) = \frac{K_3}{\tau_2 s + 1} \Gamma_2(s) + \frac{K_4}{\tau_2 s + 1} \Gamma_3(s) \tag{4-2.15}$$

where

$$K, = \frac{f_A}{f_A + 0.2(f_A + f_B)}, \text{ dimensionless} \qquad K_2 = \frac{0.2(f_A + f_B)}{f_A + 0.2(f_A + f_B)}, \text{ dimensionless}$$

$$K_3 = \frac{f_A + 0.2(f_A + f_B)}{1.2(f_A + f_B)}, \text{ dimensionless} \qquad K_4 = \frac{f_B}{1.2(f_A + f_B)}, \text{ dimensionless}$$

$$\tau_1 = \frac{V,}{f_A + 0.2(f_A + f_B)}, \text{ seconds} \qquad \tau_2 = \frac{V_2}{1.2(f_A + f_B)}, \text{ seconds}$$

Substituting Eq. 4-2.14 into Eq. 4-2.15 and solving for $T_4(s)$ give

$$\Gamma_4(s) = \frac{K_3 K_1}{(\tau_1 s + 1)(\tau_2 s + 1) - K_2 K_3} \Gamma_1(s) + \frac{K_4(\tau_1 s + 1)}{(\tau_1 s + 1)(\tau_2 s + 1) - K_2 K_3} \Gamma_3(s)$$

$$(4\text{-}2.16)$$

from which the two required transfer functions can be obtained:

$$\frac{\Gamma_4(s)}{\Gamma_1(s)} = \frac{K_3 K_1}{(\tau_1 s + 1)(\tau_2 s + 1) - K_2 K_3} \tag{4-2.17}$$

and

$$\frac{\Gamma_4(s)}{\Gamma_3(s)} = \frac{K_4(\tau_1 s + 1)}{(\tau_1 s + 1)(\tau_2 s + 1) - K_2 K_3} \tag{4-2.18}$$

Figure 4-2.5 shows two different ways to draw the block diagram. Figure 4-2.5a is developed by chaining Eqs. 4-2.14 and 4-2.15. Figure 4-2.5b is the graphical representation of Eq. 4-2.16. The feedback path in Fig. 4-2.5a shows graphically the interactive nature of this process.

The transfer functions given by Eqs. 4-2.17 and 4-2.18 are of second order, as expressed by the denominator terms. Even though both denominators are the same, the dynamic response of r,(t) to a change in r,(t) is different from the response to a change in $\Gamma_3(t)$. The fact that the term $(\tau_1 s + 1)$ appears in the numerator of Eq. 4-2.18, and not in that of Eq. 4-2.17, shows this difference. The presence of this term, as will be shown in Section 4-4.1, results in a faster dynamic response. Thus, Eqs. 4-2.17 and 4-2.18 tell us that $\Gamma_4(t)$ responds faster to a change in $\Gamma_3(t)$ than to a change in r,(t). From a physical point of view, this makes sense. Looking at Fig. 4-2.4, we notice that a change in T,(t) affects first the temperature in the first tank, $T_2(t)$, and then the temperature in the second tank, $T_4(t)$. A change in $T_3(t)$ however, affects $T_4(t)$ directly.
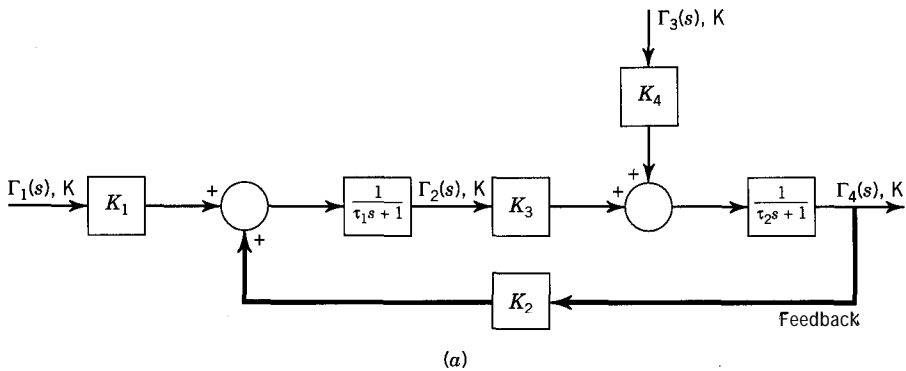
(a)
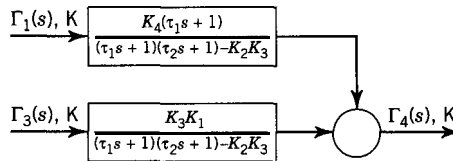
(b)

Figure 4-2.5 Block diagrams for thermal tanks with recycle.

The previous paragraph brings up a most important consideration. We must never, during any mathematical analysis, forget the physics of the process. If our analysis does not describe what happens in reality, then it is no good. Mathematics is a "tool" to describe   nature.

4-2.3 Nonisothermal Chemical Reactor

Consider the reactor shown in Fig. 4-2.6. The reactor is a continuous stirred tank (CSTR) where the exothermic reaction A $\rightarrow$ B occurs. To remove the heat of reaction, the reactor is surrounded by a jacket through which a cooling liquid flows. Let us assume that the heat losses to the surroundings are negligible and that the thermodynamic properties, densities, and heat capacities of the reactants and products are both equal and constant. The heat of reaction is constant and is given by ΔH_r in BTU/lbmole of A reacted. Let us also assume that the level of liquid in the reactor tank is constant; that is, the rate of mass into the tank is equal to the rate of mass out of the tank. Finally, the rate of reaction is given by

$$r_A(t) = k_o e^{-E/RT(t)} c_A^2(t), \quad \frac{\text{lbmoles of A produced}}{ft^3\text{-min}}$$

where the frequency factor, k_o, and the energy of activation, E, are constants. Table 4-2.1 gives the steady-state values of the variables and other process specifications.

It is desired to find out how the outlet concentration of A, $c_A(t)$, and the outlet temperature, T(t), respond to changes in the inlet concentration of A, $c_{Ai}(t)$; the inlet temperature of the reactant, $T_i(t)$; the inlet temperature of the cooling liquid, $T_{ci}(t)$; and the flows $f(t)$ and $f_c(t)$. The objective, therefore, is to develop the mathematical model, determine the transfer functions relating $c_A(t)$ and T(t) to $c_{Ai}(t)$, $T_i(t)$, $T_{ci}(t)$, $f(t)$, and $f_c(t)$, and draw the block diagram for this process.

Before we accomplish the objectives, it might be wise to discuss why we are interested in learning how the outlet temperature responds to the different inputs. This
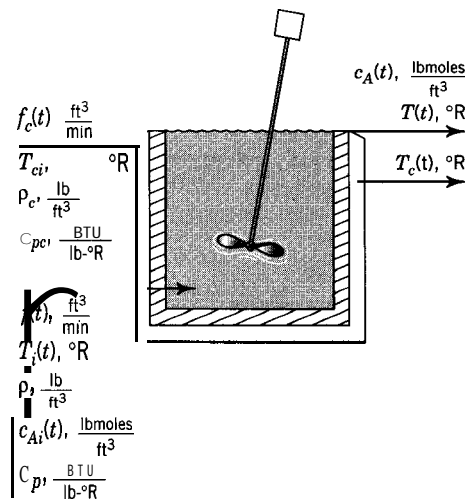
Figure 4-2.6 Nonisothermal chemical reactor.

Table 4-2.1 Process Information and Steady-State Values

Process Information

$V = 13.26 \text{ ft}^3$ $k_o = 8.33 \times 10^8 \text{ ft}^3/\text{(lbmole-min)}$

$E = 27,820 \text{ Btu/lbmole}$ $R = 1.987 \text{ Btu/(lbmole-°R)}$

$\rho = 55 \text{ lbm/ft}^3$ $C_p = 0.88 \text{ Btu/(lbm-°F)}$

$AH, . = -12,000 \text{ Btu/lbmole}$ $U = 75 \text{ Btu/(h-ft}^2\text{-°F)}$

$A = 36 \text{ ft}^2; \rho_c = 62.4 \text{ lbm/ft}^3$ $C_{pc} = 1.0 \text{ Btu/(lbm-°F)}$ $V_c = 1.56 \text{ ft}^3$

Steady-State Values

$c_{Ai}(t) = 0.5975 \text{ lbmole/ft}^3$ $T_i(t) = 635 \text{ °R}$

$T_c = 602.7 \text{ °R}$ $f = 1.3364 \text{ ft}^3/\text{min}$

$c_A(t) = 0.2068 \text{ lbmole/ft}^3$ $T(t) = 678.9 \text{ °R}$

$T_{ci}(t) = 540 \text{ °R}$ $f,(t) = 0.8771 \text{ ft}^3/\text{min}$

---

temperature is most often economically unimportant; however, it is related to safety, production rate, yield, and other operational objectives. Because temperature is easy to measure, it is usually controlled as a way to control the reactor performance.

Returning to our objectives, we start by writing an unsteady-state mole balance on component A as discussed in Section 3-6.1.

$$f(t)c_{Ai}(t) - Vr_A(t) - f(t)c_A(t) = V \frac{dc_A(t)}{dt} \tag{4-2.19}$$

1 eq., 2 unk. $[r_A(t), c_A(t)]$

where $V =$ volume of liquid in reactor, ft^3. The rate expression provides another equation:

$$r_A(t) = k_o e^{-E/RT(t)} c_A^2(t) \tag{4-2.20}$$

2 eq., 3 unk. $[T(t)]$

We still need another equation, specifically, an equation to obtain temperature. Usually, an energy balance provides this necessary equation. Thus, writing an unsteady-state energy balance on the contents of the reactor, as also presented in Section 3-6.1, gives

$$f(t)\rho C_p T_i(t) - Vr_A(t)(\Delta H_r) - UA[T(t) \quad T_c(t)] \quad f(t)\rho C_p T(t) = V\rho C_v \frac{dT(t)}{dt} \tag{4-2.21}$$

3 eq., 4 unk. $[T_c(t)]$

where

$U =$ overall heat transfer coefficient, assumed constant, $\text{Btu/ft}^2\text{-°R-min}$

$A =$ heat transfer area, ft^2

$AH, =$ heat of reaction, Btu/lbmole of A reacted

$C_v =$ heat capacity at constant volume, Btu/lb-°R

Writing an unsteady-state energy balance on the contents of the cooling jacket provides another equation:

$$f_c(t)\rho_c C_{pc} T_{ci}(t) + UA[T(t) - T_c(t)] - f_c(t)\rho_c C_{pc} T_c(t) = V_c \rho_c C_{vc} \frac{dT_c(t)}{dt} \qquad (4\text{-}2.22)$$

4 eq., 4 unk.

where

V_c = volume of cooling jacket, m^3

C_{vc} = heat capacity at constant volume of cooling liquid, assumed constant, Btu/lb-°R

Equations 4-2.19 through 4-2.22 constitute the model of the process.

To obtain the transfer functions and the block diagram, we realize that this set of equations is nonlinear, so the nonlinear terms must first be linearized. Doing so, and defining the following deviation variables

$$C_{A_i}(t) = c_{Ai}(t) - \bar{c}_{Ai} \quad C_A(t) = c_A(t) \quad \bar{c}_A \qquad \Gamma(t) = T(t) \quad \bar{T} \qquad F(t) = f(t) - \bar{f}$$

$$\Gamma_i(t) = T_i(t) \quad \bar{T}_i \qquad F_c(t) = f_c(t) - \bar{f}_c \qquad \Gamma_c(t) = T_c(t) \quad \bar{T}_c \qquad \Gamma_{ci}(t) = T_{ci}(t) - \bar{T}_{ci}$$

we get, from Eq. (4-2.19),

$$C_A(s) = \frac{K_1}{\tau_i s + 1} C_{Ai}(s) + \frac{K_2}{\tau_1 s + 1} F(s) - \frac{K_3}{\tau_1 s + 1} \Gamma(s) \qquad (4\text{-}2.23)$$

where

$$\tau_1 = \frac{V}{\bar{f} + 2Vk_o e^{-E/R\bar{T}} \bar{c}_A} = 2.07 \text{ minutes}$$

$$K, = \frac{\bar{f}}{\bar{f} + 2Vk_o e^{-E/R\bar{T}} \bar{c}_A} = 0.209$$

$$K_2 = \frac{\bar{c}_{Ai} \quad \bar{c}_A}{\bar{f} + 2Vk_o e^{-E/R\bar{T}} \bar{c}_A} = 0.0612 \frac{\text{lbmole/ft}^3}{\text{ft}^3/\text{min}}$$

$$K_3 = \frac{V \dfrac{E}{R\bar{T}^2} \bar{r}_A}{\bar{f} + 2Vk_o e^{-E/R\bar{T}} \bar{c}_A} = 0.00248 \frac{\text{lbmole/ft}^3}{\text{°R}}$$

From Eq. 4-2.21,

$$\Gamma(s) = \frac{K_4}{\tau_2 s + 1} F(s) + \frac{K_5}{\tau_2 s + 1} \Gamma_i(s) \quad \frac{K_6}{\tau_2 s + 1} C_A(s) + \frac{K_7}{\tau_2 s + 1} \, r \,, \, (s) \qquad (4\text{-}2.24)$$

where

$$\tau_2 = \frac{V\rho C_v}{V(\Delta H_r)\bar{r}_A \dfrac{E}{R\bar{T}^2} + UA + \bar{f}\rho C_p} = 7.96 \text{ min}$$

$$K_4 = \frac{\rho C_p(\bar{T}_i - \bar{T})}{V(\Delta H_r)\bar{r}_A \dfrac{E}{R\bar{T}^2} + UA + \bar{f}\rho C_p} = 26.35 \ \frac{°R}{ft^3/min}$$

$$K_5 = \frac{\bar{f}\rho C_p}{V(\Delta H_r)\bar{r}_A \dfrac{E}{R\bar{T}^2} + UA + \bar{f}\rho C_p} = -0.802$$

$$K_6 = \frac{2V(\Delta H_r)k_o e^{-E/R\bar{T}}C_A}{V(\Delta H_r)\bar{r}_A \dfrac{E}{R\bar{T}^2} + UA + \bar{f}\rho C_p} = 751.48 \ \frac{°R}{lbmoles/ft^3}$$

$$K_7 = \frac{UA}{V(\Delta H_r)\bar{r}_A \dfrac{E}{R\bar{T}^2} + UA + \bar{f}\rho C_p} = -0.558$$

Finally, Eq. 4-2.22 yields

$$\Gamma_c(s) = \frac{K_8}{\tau_3 s + 1} F_c(s) + \frac{K_9}{\tau_3 s + 1} \Gamma_{ci}(s) + \frac{K_{10}}{\tau_3 s + 1} \Gamma(s) \qquad \textbf{(4-2.25)}$$

where

$$\tau_3 = \frac{V_c \rho_c C_{pc}}{UA + \bar{f}_c \rho_c C_{pc}} = 0.976 \text{ min}$$

$$K_8 = \frac{\rho_c C_{pc}(\bar{T}_{ci} - \bar{T}_c)}{UA + \bar{f}_c \rho_c C_{pc}} = -39.23 \ \frac{°R}{ft^3/min}$$

$$K_9 = \frac{\bar{f}_c \rho_c C_{pc}}{UA + \bar{f}_c \rho_c C_{pc}} = 0.5488$$

$$K_{10} = \frac{UA}{UA + \bar{f}_c \rho_c C_{pc}} = 0.4512$$

Substituting Eq. 4-2.25 into Eq. 4-2.24 gives

$$\Gamma(s) = \frac{(\tau_3 s + 1)}{(\tau_2 s + 1)(\tau_3 s + 1) - K_7 K_{10}} [K_4 F(s) + K_5 \Gamma_i(s) - K_6 C_A(s)]$$

$$+ \frac{K_7}{(\tau_2 s + 1)(\tau_3 s + 1) - K_7 K_{10}} [K_8 F_c(s) + K_9 \Gamma_{ci}(s)] \qquad \textbf{(4-2.26)}$$

Substituting   Eq.   4-2.26   into   Eq.   4-2.23   yields

$$C_A(s) = \frac{K_1[(\tau_2 s + 1)(\tau_3 s + 1) - K_7 K_{10}]}{(\tau_1 s + 1)(\tau_2 s + 1)(\tau_3 s + 1) - K_7 K_{10}(\tau_1 s + 1) - K_3 K_6(\tau_3 s + 1)} C_{Ai}(s)$$

$$+ \frac{K_2[(\tau_2 s + 1)(\tau_3 s + 1) - K_7 K_{10}] - K_3 K_4(\tau_3 s + 1)}{(\tau_1 s + 1)(\tau_2 s + 1)(\tau_3 s + 1) - K_7 K_{10}(\tau_1 s + 1) - K_3 K_6(\tau_3 s + 1)} F(s)$$

$$- \frac{K_3 K_5(\tau_3 s + 1)}{(\tau_1 s + 1)(\tau_2 s + 1)(\tau_3 s + 1) - K_7 K_{10}(\tau_1 s + 1) - K_3 K_6(\tau_3 s + 1)} \Gamma_i(s)$$

$$- \frac{K_3 K_7}{(\tau_1 s + 1)(\tau_2 s + 1)(\tau_3 s + 1) - K_7 K_{10}(\tau_1 s + 1) - K_3 K_6(\tau_3 s + 1)}$$

$$[K_8 F_c(s) + K_9 \Gamma_{ci}(s)] \quad \textbf{(4-2.27)}$$

and   from   Eq.   4-2.27,   the   following   required   transfer   functions   can   be   obtained.

$$\frac{C_A(s)}{C_{Ai}(s)} = \frac{K_1[(\tau_2 s + 1)(\tau_3 s + 1) - K_7 K_{10}]}{(\tau_1 s + 1)(\tau_2 s + 1)(\tau_3 s + 1) - K_7 K_{10}(\tau_1 s + 1) - K_3 K_6(\tau_3 s + 1)} \tag{4-2.28}$$

$$= \frac{0.427(0.95s + 1)(6.54s - 1)}{26.27s^3 + 36.31s^2 + 10.14s + 1} \tag{4-2.29}$$

$$\frac{C_A(s)}{F(s)} = \frac{K_2[(\tau_2 s + 1)(\tau_3 s + 1) - K_7 K_{10}] - K_3 K_4(\tau_3 s + 1)}{(\tau_1 s + 1)(\tau_2 s + 1)(\tau_3 s + 1) - K_7 K_{10}(\tau_1 s + 1) - K_3 K_6(\tau_3 s + 1)} \tag{4-2.30}$$

$$= \frac{0.0182(0.95s + 1)(44.75s - 1)}{26.27s^3 + 36.31s^2 + 10.14s + 1} \tag{4-2.31}$$

$$\frac{C_A(s)}{\Gamma_i(s)} = \frac{- K_3 K_5(\tau_3 s + 1)}{(\tau_1 s + 1)(\tau_2 s + 1)(\tau_3 s + 1) - K_7 K_{10}(\tau_1 s + 1) - K_3 K_6(\tau_3 s + 1)} \tag{4-2.32}$$

$$= \frac{- 0.0032(0.976s + 1)}{26.27s^3 + 36.31s^2 + 10.14s + 1} \tag{4-2.33}$$

$$\frac{C_A(s)}{F_c(s)} = \frac{- K_3 K_7 K_8}{(\tau_1 s + 1)(\tau_2 s + 1)(\tau_3 s + 1) - K_7 K_{10}(\tau_1 s + 1) - K_3 K_6(\tau_3 s + 1)} \tag{4-2.34}$$

$$= \frac{0.0887}{26.27s^3 + 36.31s^2 + 10.14s + 1} \tag{4-2.35}$$

$$\frac{C_A(s)}{\Gamma_{ci}(s)} = \frac{- K_3 K_7 K_9}{(\tau_1 s + 1)(\tau_2 s + 1)(\tau_3 s + 1) \quad K_7 K_{10}(\tau_1 s + 1) \quad K_3 K_6(\tau_3 s + 1)} \tag{4-2.36}$$

$$= \frac{- 0.00124}{26.27s^3 + 36.31s^2 + 10.14s + 1} \tag{4-2.37}$$

From Eqs. 4-2.23, 4-2.24, and 4-2.25, we also obtain

$$\frac{\Gamma(s)}{\Gamma_i(s)} = \frac{1.31(2.07s + 1)(0.976s + 1)}{26.27s^3 + 36.31s^2 + 10.14s + 1} \tag{4-2.38}$$

$$\frac{\Gamma(s)}{F(s)} = \frac{-31.79(0.976s + 1)(2.77s - 1)}{26.27s^3 + 36.31s^2 + 10.14s + 1} \tag{4-2.39}$$

$$\frac{\Gamma(s)}{C_{Ai}(s)} = \frac{256(0.976s + 1)}{26.27s^3 + 36.31s^2 + 10.14s + 1} \tag{4-2.40}$$

$$\frac{\Gamma(s)}{F_c(s)} = \frac{-35.77(2.07s + 1)}{26.27s^3 + 36.31s^2 + 10.14s + 1} \tag{4-2.41}$$

$$\frac{\Gamma(s)}{\Gamma_{ci}(s)} = \frac{0.5(2.07s + 1)}{26.27s^3 + 36.31s^2 + 10.14s + 1} \tag{4-2.42}$$

All of the transfer functions developed are of third order. However, the dynamic behavior of the responding variables varies significantly depending on the forcing function. The differences are due to the terms in the numerator. Equations 4-2.35 and 4-2.37 show that the dynamic behavior of $C_A(t)$ in response to changes in $F_c(t)$ is the same as that in response to r,,(t) but different from the behavior due to changes in $C_{Ai}(t)$, $F(t)$, or $\Gamma_i(t)$. Furthermore, Eqs. 4-2.29, 4-2.31, and 4-2.33 also indicate different dynamic behavior. Similarly, Eqs. 4-2.41 and 4-2.42 indicate the same dynamic behavior of $\Gamma(t)$ in response to changes in $F_c(t)$ and $\Gamma_{ci}(t)$. Note that the dynamic response of $C_A(t)$ to a change in $\Gamma_i(t)$ is the same as the dynamic response of r(t) to a change in $C_{Ai}(t)$, as indicated by Eqs. 4-2.33 and 4-2.40. Section 4-3 explains in detail the effect of the term $(\tau s + 1)$ in the numerator of the transfer function, and Section 4-4.2 explains the significance of the similar, but distinctly different, term $(\tau s - 1)$.

Figure 4-2.7 shows different ways to draw the block diagram for this reactor. Although Fig. 4-2.7b seems a bit less complex, Fig. 4-2.7a clearly shows the feedback paths indicating the interactions.

In Chapter 3 the nonlinear characteristics of processes were presented and discussed. Chemical reactors are nonlinear in their behavior, so it is appropriate to use this reactor to demonstrate once more the nonlinear characteristics of processes. Figure 4-2.8 shows how four of the terms that describe the process vary as the concentration in the reactor, $c_A(t)$, is operated at different conditions. To obtain these different conditions, the coolant flow, $f_c(t)$, was varied, which also resulted in a variation of the temperatures, T(t) and T,(t), in the reactor. Figure 4-2.8a shows how the gain in Eq. 4-2.29 varies. This gain is calculated from Eq. 4-2.28 as $\dfrac{K_1}{1 - K_7 K_{10} - K_3 K_6}$ $K_1 K_7 K_{10}$. At the original steady-state condition, the value of this gain is -0.427. The figure shows that sometimes the variation is as much as a factor of 2. The figure also shows that after some minimum, the gain starts to increase again. Figure 4-2.8b shows the gain of Eq. 4-2.35, which is calculated from Eq. 4-2.34 as $\dfrac{-K_3 K_7 K_8}{1 - K_7 K_{10} - K_3 K_6}$. The figure shows a change of greater than a
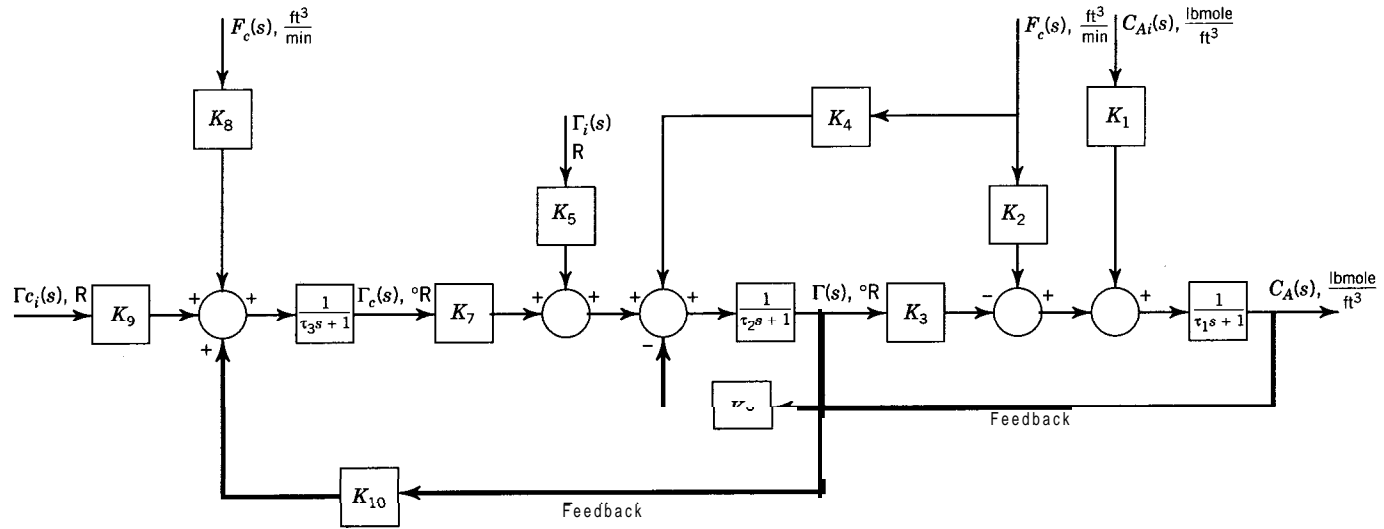
$$(a)$$

Figure **4-2.7a** Block diagram for nonisothermal chemical reactor.

$$\frac{C_{Ai}(s)}{\frac{\text{lbmole}}{\text{ft}^3}} \qquad \frac{K_1\left[(\tau_1 s + 1)(\tau_3 s + 1) - K_7 K_{10}\right]}{(\tau_1 s + 1)(\tau_2 s + 1)(\tau_3 s + 1) - K_7 K_{10}(\tau_1 s + 1) - K_3 K_6(\tau_3 s + 1)}$$

$$\frac{F(s)}{\frac{\text{ft}^3}{\text{min}}} \qquad \frac{K_2\left[(\tau_1 s + 1)(\tau_3 s + 1) - K_7 K_{10}\right] - K_3 K_4(\tau_3 s + 1)}{(\tau_1 s + 1)(\tau_2 s + 1)(\tau_3 s + 1) - K_7 K_{10}(\tau_1 s + 1) - K_3 K_6(\tau_3 s + 1)}$$

$$\frac{\Gamma_i(s)}{°R} \qquad \frac{K_3 K_5(\tau_3 s + 1)}{(\tau_1 s + 1)(\tau_2 s + 1)(\tau_3 s + 1) - K_7 K_{10}(\tau_1 s + 1) - K_3 K_6(\tau_3 s + 1)}$$

$$\frac{F_c(s)}{\frac{\text{ft}^3}{\text{min}}} \qquad \frac{K_3 K_7 K_8}{(\tau_1 s + 1)(\tau_2 s + 1)(\tau_3 s + 1) - K_7 K_{10}(\tau_1 s + 1) - K_3 K_6(\tau_3 s + 1)}$$

$$\frac{\Gamma_{ci}(s)}{°R} \qquad \frac{K_3 K_7 K_9}{(\tau_1 s + 1)(\tau_2 s + 1)(\tau_3 s + 1) - K_7 K_{10}(\tau_1 s + 1) - K_3 K_6(\tau_3 s + 1)}$$

$$\frac{C_A(s)}{\frac{\text{lbmole}}{\text{ft}^3}}$$
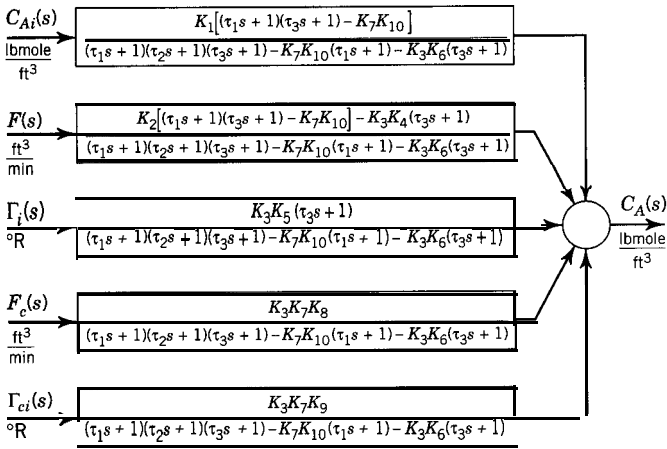
Figure 4-2.76 Block diagram for nonisothermal chemical reactor.
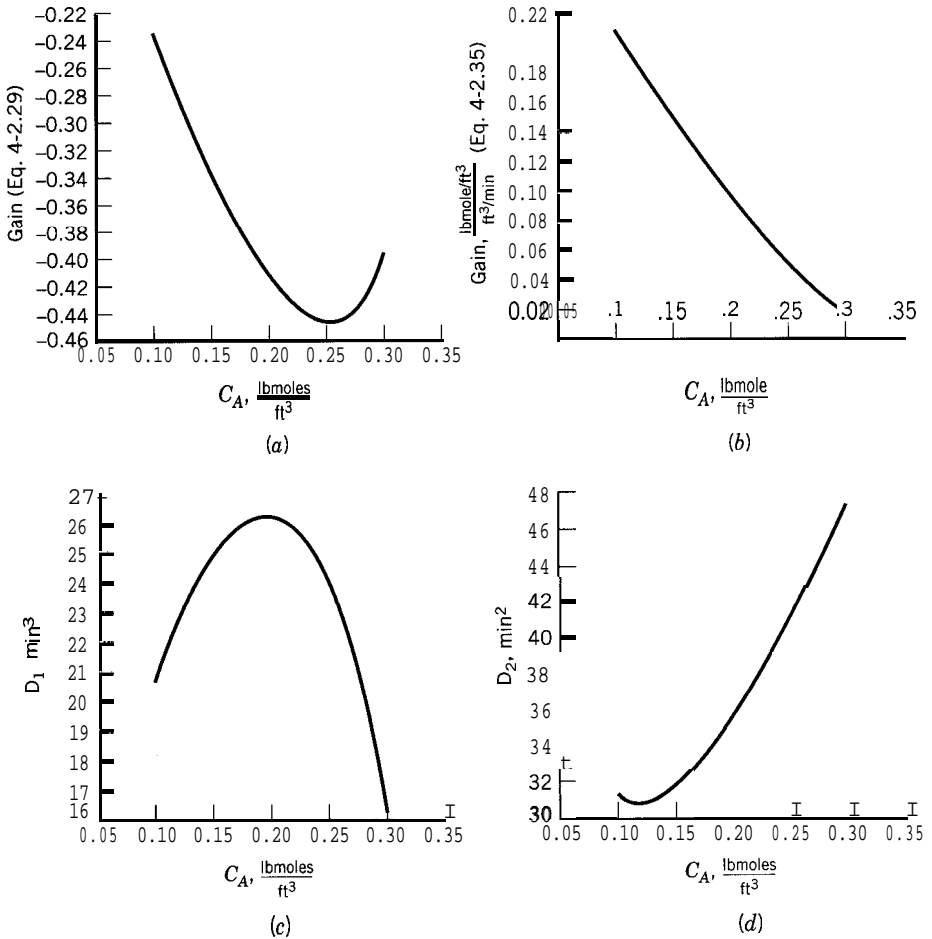
Figure 4-2.8 Variations of process parameters with operating conditions.

factor of 10. Figure 4-2.8c shows how the first term of the denominator, referred to as D,, of all the transfer functions developed varies. This term is calculated as $\dfrac{\tau_1 \tau_2 \tau_3}{1 - K_7 K_{10} \quad K_3 K_6}$, and has a value of 26.27 min^3 at the original steady state. Figure 4-2.8d shows the second term of the denominator, referred to as D,, of the transfer functions. This term is calculated as $\dfrac{\tau_1 \tau_2 + \tau_1 \tau_3 + \tau_2 \tau_3}{- K_7 K_{10} - K_3 K_6}$ and has a value of 36.31 min^2 at the original steady state. As we discussed in Chapter 3, the variations in process behavior, the nonlinearities, as shown in Fig. 4-2.8, have a significant effect on the control of the process.

In this reactor example, the heat transfer rate expression $UA[T(t) - T,(t)]$ has been used. This expression says that once the cooling water temperature, $T_c(t)$, changes, the contents of the reactor immediately undergo a change in heat transfer. Thus the dynamics of the wall have been neglected. In reality, however, these dynamics may be significant. When the cooling water temperature changes, the heat transfer to the wall changes. As the wall temperature changes, the heat transfer from the wall to the reactants then changes. Thus it is only after the wall temperature exhibits the change that the heat transfer to the reacting mass starts to change. Therefore, the wall represents another capacitance in the system, the magnitude of which depends on thickness, density, heat capacity, and other physical properties of the material of the construction of the wall.

Taking the wall into consideration gives a better understanding of the capacitance. We will assume that the wall is at a uniform temperature, $T_m(t)$, because the heat transfer resistance of the wall is small compared to the resistances of the films on each side. Sometimes, one of the two resistances is much larger than the other. In this case, the capacitance of the wall can be lumped with the capacitance of the side of the smaller resistance, and they are assumed to be at the same temperature.

When we consider the reactor wall, the unsteady-state mole balance on component A and the rate of reaction remain the same, thus providing two equations (Eqs. 4-2.19 and 4-2.20), with three unknowns: $r_A(t)$, c,(t), and T(t). The unsteady-state energy balance on the contents of the reactor is changed to

$$f(t)\rho C_p T_i(t) - V r_A(t)(\Delta H_r) - h_i A_i[T(t) - T_m(t)] - f(t)\rho C_p T(t) = V \rho C_v \frac{dT(t)}{dt} \quad (4\text{-}2.43)$$

3 eq., 4 unk. [T,(t)]

where

h_i = inside film heat transfer coefficient, assumed constant, $Btu/ft^2\text{-}min\text{-}°R$
$A,$ = inside heat transfer area, ft^2
T,(t) = temperature of metal wall, $°R$

Proceeding with an unsteady-state energy balance on the wall, we can write

$$h_i A_i[T(t) - T_m(t)] - h_o A_o[T_m(t) - T_c(t)] = V_m \rho_m C_{vm} \frac{dT_m(t)}{dt} \quad (4\text{-}2.44)$$

4 eq., 5 unk. [T,(t)]

where

h_o = outside film heat transfer coefficient, assumed constant, $\text{Btu/ft}^2\text{-min-}°R$
$A,$ = outside heat transfer area, ft^2
V_m = volume of the metal wall, ft^3
ρ_m = density of the metal wall, lb/ft^3
C_{vm} = heat capacity at constant volume of the metal wall, $\text{Btu/lb-}°R$

Finally, an unsteady-state energy balance on the cooling water gives the other required equation:

$$f_c(t)\rho_c C_{pc} T_{ci}(t) + h_o A_o[T_m(t) - T_c(t)] - f_c(t)\rho_c C_{pc} T_c(t) = V_c \rho_c C_{vc} \frac{dT_c(t)}{dt} \quad \text{(4-2.45)}$$

5 eq., 5 unk.

Five equations are now required to describe the reactor. Eq. 4-2.44 is the new equation describing the dynamics of the wall.

From Eqs. 4-2.19 and 4-2.20, Eq. 4-2.23 is obtained as previously shown. We write this equation again for convenience.

$$C_A(s) = \frac{K_1}{\tau_1 s + 1} C_{Ai}(s) + \frac{K_2}{\tau_1 s + 1} F(s) - \frac{K_3}{\tau_1 s + 1} \Gamma(s) \quad \text{(4-2.23)}$$

From Eq. 4-2.43 and using the procedure previously learned, we obtain

$$\Gamma(s) = \frac{K_{11}}{\tau_4 s + 1} F(s) + \frac{K_{12}}{\tau_4 s + 1} \Gamma_i(s) - \frac{K_{13}}{\tau_4 s + 1} C_A(s) + \frac{K_{14}}{\tau_4 s + 1} \Gamma_m(s) \quad \text{(4-2.46)}$$

where

$$\tau_4 = \frac{V\rho C_v}{V(\Delta H_r)\bar{r}_A \dfrac{E}{R\bar{T}^2} + h_i A_i + \bar{f}\rho C_p}, \quad \text{minutes}$$

$$K_{11} = \frac{\rho C_p(\bar{T}_i - \bar{T})}{V(\Delta H_r)\bar{r}_A \dfrac{E}{R\bar{T}^2} + h_i A_i + \bar{f}\rho C_p}, \quad \frac{°R}{\text{ft}^3/\text{min}}$$

$$K_{12} = \frac{\bar{f}\rho C_p}{V(\Delta H_r)\bar{r}_A \dfrac{E}{R\bar{T}^2} + h_i A_i + \bar{f}\rho C_p}, \quad \text{dimensionless}$$

$$K_{13} = \frac{V(\Delta H_r)k_o e^{-E/R\bar{T}}}{V(\Delta H_r)\bar{r}_A \dfrac{E}{R\bar{T}^2} + h_i A_i + \bar{f}\rho C_p}, \quad \frac{R}{\text{lbmol/ft}^3}$$

$$K_{14} = \frac{h_i A_i}{V(\Delta H_r)\bar{r}_A \dfrac{E}{R\bar{T}^2} + h_i A_i + \bar{f}\rho C_p}, \quad \text{dimensionless}$$

From Eq. 4-2.44 and the usual procedure, we obtain

$$\Gamma_m(s) = \frac{K_{15}}{\tau_5 s + 1} \Gamma(s) + \frac{K_{16}}{\tau_5 s + 1} \Gamma_c(s) \tag{4-2.47}$$

where

$$\tau_5 = \frac{V_m \rho_m C_{vm}}{h_i A_i + h_o A_o}, \text{ minutes}$$

$$K_{15} = \frac{h_i A_i}{h_i A_i + h_o A_o}, \text{ dimensionless}$$

$$K_{16} = \frac{h_o A_o}{h_i A_i + h_o A_o}, \text{ dimensionless}$$

Finally, from Eq. (4-2.45)

$$\Gamma_c(s) = \frac{K_{17}}{\tau_6 s + 1} F_c(s) + \frac{K_{18}}{\tau_6 s + 1} \Gamma_{ci}(s) + \frac{K_{19}}{\tau_6 s + 1} \Gamma_m(s) \tag{4-2.48}$$

where

$$\tau_6 = \frac{V_c \rho_c C_{vc}}{h_o A_o + \bar{f}_c \rho_c C_{pc}}, \text{ minutes}$$

$$K_{17} = \frac{\rho_c C_{pc}(\bar{T}_{ci} - \bar{T}_c)}{h_o A_o + \bar{f}_c \rho_c C_{pc}}, \frac{°R}{ft^3/min}$$

$$K_{18} = \frac{\bar{f}_c \rho_c C_{pc}}{h_o A_o + \bar{f}_c \rho_c C_{pc}}, \text{ dimensionless}$$

$$K_{19} = \frac{h_o A_o}{h_o A_o + \bar{f}_c \rho_c C_{pc}}, \text{ dimensionless}$$

With Eqs. 4-2.23, 4-2.46, 4-2.47 and 4-2.48, the block diagram for this process can be developed. This block diagram, shown in Fig. 4-2.9, shows that now there are three feedback paths, indicating the interactive nature of the process.

Finally, as the reader has undoubtedly noticed, the development of any desired transfer function for this system is more complex (even though only algebraic manipulation is required) than for the previous case. As any good textbook would say, the development of these transfer functions from the foregoing equations is "left to the reader as an exercise."

4-3 RESPONSE OF HIGHER-ORDER SYSTEMS

Several types of higher-order transfer functions were developed in the previous sections. Two of the most common are

$$G(s) = \frac{Y(s)}{X(s)} = \prod_{i=1}^{n} G_i(s) = \frac{K}{\prod_{i=1}^{n} (\tau_i s + 1)} \tag{4-3.1}$$
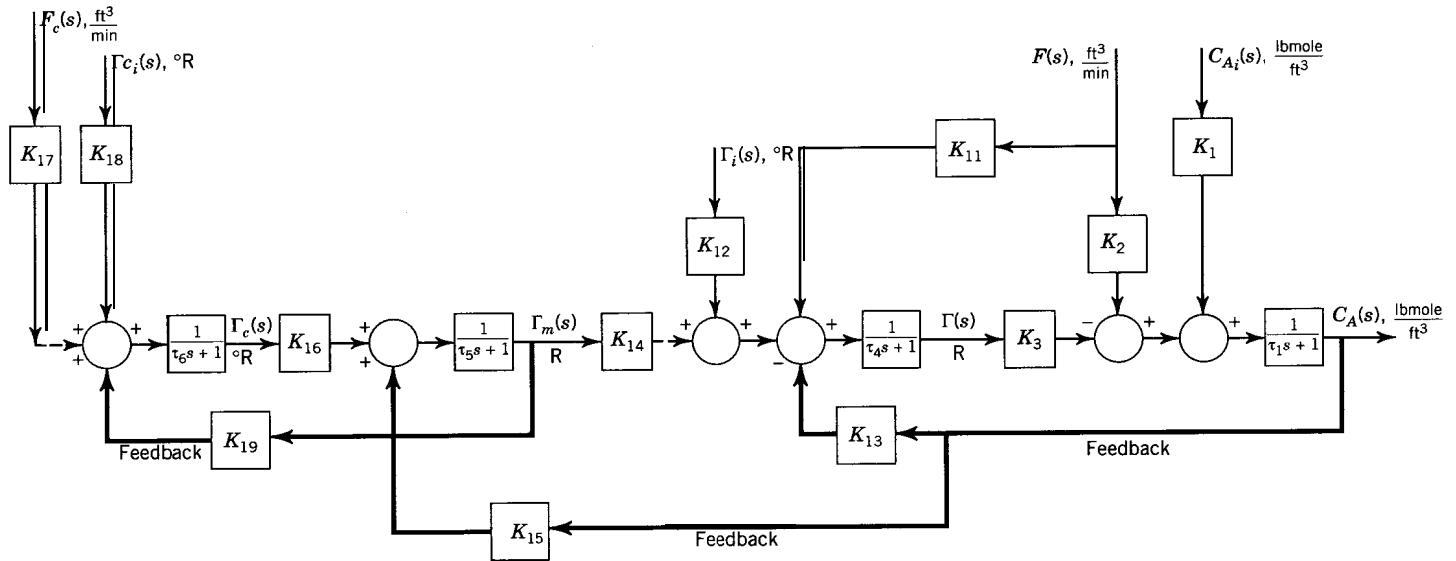
Figure 4-2.9 Block diagram for nonisothermal chemical reactor—wall considered.

and

$$G(s) = \frac{Y(s)}{X(s)} = \frac{K \prod_{j=1}^{m} (\tau_{ldj}s + 1)}{\prod_{i=1}^{n} (\tau_{lgi}s + 1)} \qquad n > m \qquad (4\text{-}3.2)$$

A third type of transfer function developed, the one with the term $(\tau s - 1)$ in the numerator, is discussed in Section 4-4.3

Section 2-5 presented the response of higher-order systems. This section presents a brief review of the response of Eqs. 4-3.1 and 4-3.2 to step change in input. We believe this brief presentation makes it easier to understand the difference in dynamic response between the systems studied in this and the previous chapter.

Consider the transfer function given by Eq. 4-3.1 with real and distinct roots. In the time domain, the response to a step change of unit magnitude is given by Eq. 4-3.3.

$$Y(t) = K \left[1 - \sum_{i=1}^{n} \frac{\tau_i^{n-1} e^{-t/\tau_i}}{\prod_{\substack{j=1 \\ j \neq i}}^{n} (\tau_i - \tau_j)} \right] \qquad (4\text{-}3.3)$$

The general method for solving transfer functions with other types of roots is presented in Chapter 2.

Figure 4-3.1 shows the response of systems of Eq. 4-3.1 with $n = 2$ through $n = 6$ to a step change in forcing function, $X(s) = 1/s$, where all the time constants are equal to 1 min (Eq. 4-3.3 does not apply in this case because the roots are not distinct). From the figure, it is clear that as the order of the system increases, the initial response of the system is slower and slower. That is, there is an "apparent" dead time that also seems
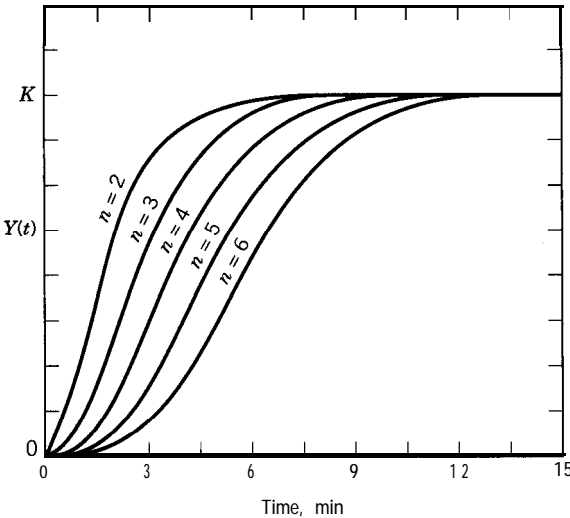
Figure 4-3.1 Response of overdamped higher-order systems to a unit step change in input.

to increase. This is important in the study of automatic process control because most industrial processes are composed of first-order systems in series. Analyzing Fig. 4-3.1 in more detail, we may realize that the response of third-order and higher-order systems looks similar to the response of a second-order overdamped system with some amount of dead time. Because of this similarity, the response of these systems can be approximated by that of a second-order-plus-dead-time (SOPDT). Mathematically, this is shown as follows:

$$\frac{Y(s)}{X(s)} = \frac{K}{\displaystyle\prod_{i=1}^{n} (\tau_i s + 1)} \approx \frac{Ke^{-t_{0}s}}{(\tau_a s + 1)(\tau_b s + 1)} \tag{4-3.4}$$

for $n > 2$.

The response of processes described by Eq. 4-3.2, with real and distinct roots, to a step change of unit magnitude in forcing function is given by Eq. 4-3.5.

$$Y(t) = K\left[1 - \sum_{i=1}^{n} \frac{\displaystyle\prod_{j=1}^{m} (\tau_{lgi} - \tau_{ldj})\tau_{lgi}^{n-m-1}}{\displaystyle\prod_{\substack{j=1 \\ j \neq i}}^{n} (\tau_{lgi} - \tau_{ldj})} e^{-t/\tau_{lgi}} \right] \tag{4-3.5}$$

To obtain a better understanding of the term $(\tau_{ld}s + 1)$, let us compare the responses of the following two processes to a step change of unit magnitude in X(s).

$$\frac{Y_1(s)}{X(s)} = \frac{1}{(s + 1)(2s + 1)(3s + 1)} \tag{4-3.6}$$

$$\frac{Y_2(s)}{X(s)} = \frac{(0.5s + 1)}{(s + 1)(2s + 1)(3s + 1)} \tag{4-3.7}$$

Figure 4-3.2 shows the two responses. The effect of the term $(\tau_{ld}s + 1)$ is to "speed up" the response of the process. This is opposite to the effect of $1/(\tau_{lg}s + 1)$. In Chapter 3, the term $1/(\tau_{lg}s + 1)$ was referred to as a *first-order* lug. Consequently, we refer to the term $(\tau_{ld}s + 1)$ as a *first-order lead.* This is why the notation τ_{lg}, indicating a "lag" time constant, and τ_{ld}, indicating a "lead" time constant, is used. Note that when τ_{ld} becomes equal to τ_{lg}, the transfer function becomes of one order less. Chapter 2 also presented the concepts of lead and lag using the results of a response to a ramp function.

A common characteristic of all the responses presented so far is that they all reach a new steady state, or operating condition. Processes that show this characteristic— that is, those processes that after a step change in input reach a new steady state-are sometimes classified as *self-regulating processes;* most processes are of this type. Section 4-4 presents two examples of *non-self-regulating processes.*

4-4 OTHER TYPES OF PROCESS RESPONSES

This section presents some systems that cannot be classified as any of the types presented so far. The first two systems presented are sometimes classified under the general
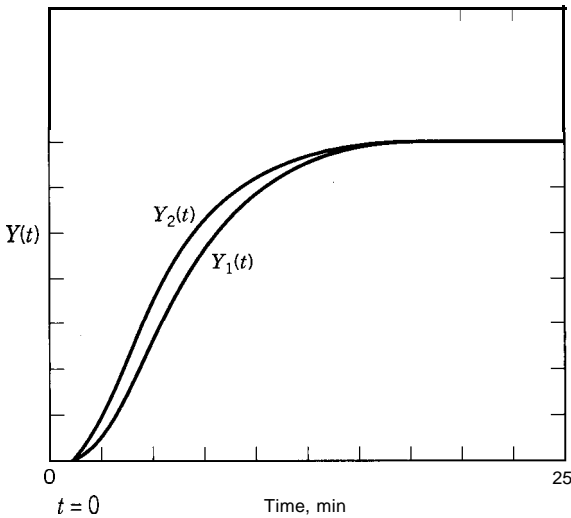
Figure 4-3.2 Comparison of responses of Eqs. 4-3.6 and 4-3.7.

heading of non-self-regulating; this section explains the reason for this term. The third system presented, though of the self-regulating type, has a different response from the systems presented in this and the previous chapter.

4-4.1   Integrating   Processes: Level   Process

Consider the process tank shown in Fig. 4-4.1. An input stream enters the tank freely, whereas the output stream depends on the speed of the pump. The pump speed is regulated by the signal m(t), %. The relation between the output flow and the signal is given by

$$\tau_p \frac{df_o(t)}{dt} + f_o(t) = K_p m(t)$$
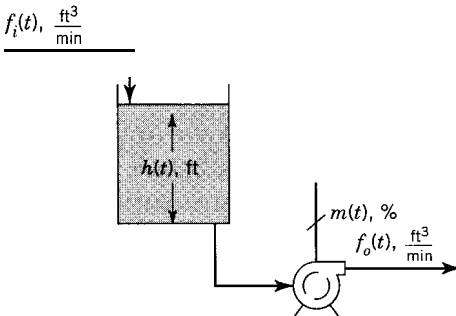
Figure 4-4.1 Process tank with pump manipulating   outlet   flow.

That is, the pump flow does not respond instantaneously to a change in signal but rather as a first-order response with time constant τ_p, min, and gain $K_p \dfrac{ft^3/min}{\%}$. It is important to realize that the pump flow does not depend on the level in the tank but only on the input signal m(t).

Develop the mathematical model, obtain the transfer functions, and draw the block diagrams that relate the tank level, h(t), to the input flow, $f_i(t)$, and the input signal, m(t).

An unsteady-state mass balance around the tank provides the first equation needed:

$$\rho f_i(t) - \rho f_o(t) = \rho A \frac{dh(t)}{dt} \tag{4-4.1}$$

$$1 \text{ eq., } 2 \text{ unk. } [f_o(t), h(t)]$$

where

ρ = liquid density, assumed constant, lb/ft^3
A = cross-sectional area of tank, ft^2

The second equation is provided by the pump:

$$\tau_p \frac{df_o(t)}{dt} + f_o(t) = K_p m(t) \tag{4-4.2}$$

$$2 \text{ eq., } 2 \text{ unk.}$$

Only two equations are required to model this simple process.
Following the usual procedure, we obtain from Eq. 4-4.1

$$H(s) = \frac{1}{As} [F_i(s) - F_o(s)] \tag{4-4.3}$$

where the deviation variables are

$$H(t) = h(t) - \bar{h} \qquad F_i(t) = f_i(t) - \bar{f}_i \qquad F_o(t) = f_o(t) \quad \bar{f}_o$$

From Eq 4-4.2 we obtain

$$F_o(s) = \frac{K_p}{\tau_p s + 1} M(s) \tag{4-4.4}$$

where the new deviation variable is $M(t) = m(t) \quad \bar{m}$.
Substituting Eq. 4-4.4 into Eq. 4-4.3 yields

$$H(s) = \frac{1}{As} F_i(s) - \frac{K_p}{As(\tau_p s + 1)} M(s) \tag{4-4.5}$$

from which we can write the following transfer functions:

$$\frac{H(s)}{F_i(s)} = \frac{1}{As} \qquad (4\text{-}4.6)$$

and

$$\frac{H(s)}{M(s)} = \frac{-K_p}{As(\tau_p s + 1)} \qquad (4\text{-}4.7)$$

These two transfer functions are different from the ones developed so far in this and the previous chapter. The single s term in the denominator indicates the "integrating" nature of the process. Let us develop the response of the system to a change of -B% in the signal m(t). That is,

$$M(t) = -Bu(t)$$

$$M(s) = -\frac{B}{s}$$

Using the techniques learned in Chapter 2, we find that

$$H(s) = \frac{K_p B}{As^2(\tau_p s + 1)}$$

and inverting this equation back to the time domain yields

$$H(t) = \frac{K_p B}{A}(t - \tau_p + \tau_p e^{-t/\tau_p}) \qquad \textbf{(4-4.8)}$$

This equation shows that as time increases, the exponential term decays to zero but the first term continues to increase; this results in a ramp-type level response. In theory, the level should continue to increase, "integrating," without bounds. Realistically, the level will stop increasing when it overflows, an extreme operating condition. If the signal had increased, increasing the pump speed, the analysis would have shown the same type of response but in the opposite direction. That is, the level would have decreased, theoretically without bounds. Realistically, the level would stop decreasing when it reached a very low level, or when the pump starts to cavitate. In practice, however, tanks are usually instrumented with high/low level alarms and switches designed to avoid these extreme operating conditions. These safety controls are required in any well-designed process. Figure 4-4.2 shows the response of the system, and Fig. 4-4.3 shows the block diagram.

The integrating nature of this system develops because the **outlet flow,** $f_o(t)$**,** *is not a function of the level in the tank but is only a function of the signal to the pump as expressed by Eq.* 4-4.2. *That is, there is no "process feedback" to provide regulation.*
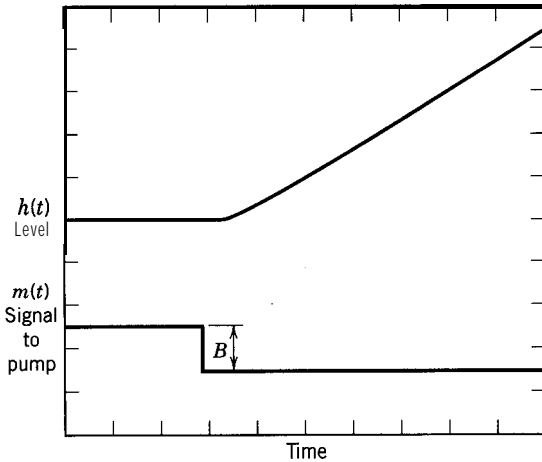
Figure 4-4.2 Response of tank level to a change in signal to pump.

Very often, control valves are used to manipulate the outlet stream. Figure 4-4.4 shows two possible arrangements. Figure 4-4.4*a* is essentially the same as Fig. 4-4.1; that is, the valve's upstream pressure is provided by the pump and therefore is independent of the level in the tank. In Fig. 4-4.4*b*, the upstream pressure is dependent on the level, and therefore the flow is also dependent on the level. This dependence provides the "process feedback" necessary for self-regulation. In this case, the transfer functions would have been

$$\frac{H(s)}{F_i(s)} = \frac{K_1}{\tau s + 1} \tag{4-4.9}$$

and

$$\frac{H(s)}{M(s)} = \frac{-K_2}{(\tau s + 1)(\tau_v s + 1)} \tag{4-4.10}$$

where

τ = time constant of the tank
τ_v = time constant of the valve
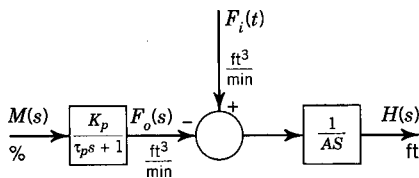
Figure 4-4.3 Block diagram for process tank.
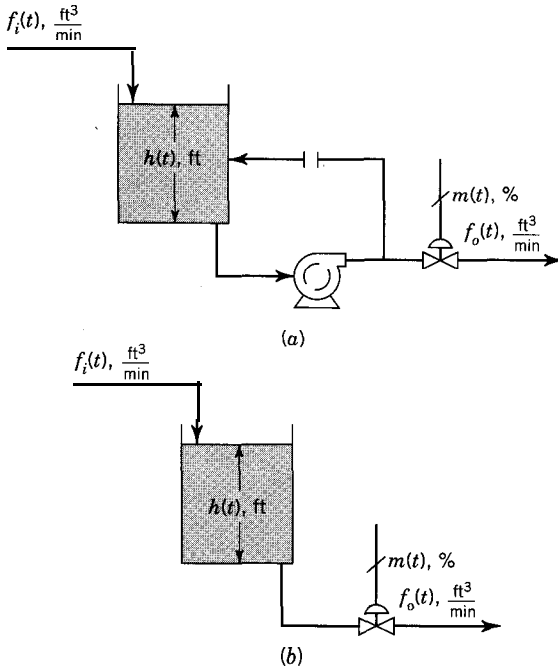
Figure 4-4.4 Process tank with control valve manipulating
outlet   flow.

Sometimes, however, the level does not strongly affect the outlet flow, which results
in a response similar to the one shown in Fig. 4-4.2. To explore this point further,
assume that the signal to the valve changes by $-B\%$, closing the valve some amount.
In this case, the level rises, and in so doing it increases the outlet flow due to the liquid
head. This process continues until the outlet flow balances the inlet flow, reaching a
new steady state. But suppose that the increase in level necessary to reach steady state
puts it above the maximum tank height, thus overflowing the tank. In this case, even
though the process is trying to balance itself, the operation still results in a response
similar to integrating.

Any system described by a transfer function containing an isolated s term in the
denominator is referred to as an integrating system. The response of these systems to
a step change in input is such that, in theory, they will not reach a new steady-state
value, or operating condition. That is, they do not regulate themselves to a new steady-
state condition and thus are sometimes classified as non-self-regulating *systems. In*
practice, as shown in this section, they reach an "extreme" steady-state condition. The
most common example of an integrating system is a level process.

Level control is discussed in more detail in Chapter 7.

4-4.2 Open-Loop Unstable Process: Chemical Reactor

Consider a chemical reactor, shown in Fig. 4-4.5, where the exothermic reaction
A $\rightarrow$ B takes place. To remove the heat of reaction, a jacket surrounds the reactor
where a cooling liquid is maintained at 100°F as a result of a high recirculation rate.
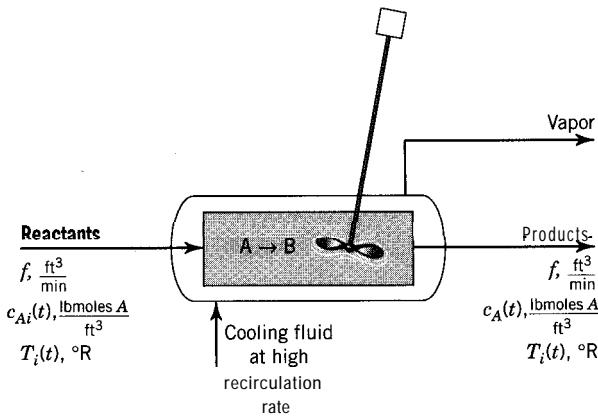
Figure 4-4.5 Chemical reactor.

It is desired to develop the set of equations that describe this process and to write the transfer functions relating the outlet reactor temperature and concentration to the inlet temperature and concentration. Assume that the reactor contents are well mixed, that the reactor is well insulated, and that the heat capacities and densities of the reactant and product are equal to each other. Table 4-4.1 presents all the necessary process information and steady-state values.

We start by writing an unsteady-state mole balance on the reactant component A:

$$fc_{Ai}(t) - Vr_A(t) - fc_A(t) = V \frac{dc_A(t)}{dt} \qquad \text{(4-4.11)}$$

$$1 \text{ eq., } 2 \text{ unk. } [r_A(t), c_A(t)]$$

where

$$r_A(t) = \text{rate of reaction, lbmoles of A reacted/ft}^3\text{-min}$$
$$V = \text{volume of reactor, ft}^3$$

Table 4-4.1 Process Information and Steady-State Values

Process Information

$V = 13.26 \text{ ft}^3$	$A = 36 \text{ ft}^2$
$E = 27{,}820 \text{ Btu/lbmole}$	$R = 1.987 \text{ Btu/lbmole-}°\text{R}$
$\rho = 55 \text{ lbm/ft}^3$	$C_p = 0.88 \text{ Btu/lbm-}°\text{F}$
$\text{AH,.} = 12{,}020 \text{ Btu/lbmole}$	$U = 75 \text{ Btu/(h-ft}^2\text{-}°\text{F)}$
$k_o = 1.73515 \times 10^{13}/\text{min}$	

Steady-State Values

$c_{Ai}(t) = 0.8983 \text{ lbmole/ft}^3$	$T_i(t) = 578 \text{ }°\text{R}$
$T_c = 560.0°\text{R}$	$f = 1.3364 \text{ ft}^3/\text{min}$
$c_A(t) = 0.08023 \text{ lbmole/ft}^3$	$T(t) = 690.0°\text{R}$

The rate of reaction gives

$$r_A(t) = k_o e^{-E/RT(t)} c_A(t) \qquad \textbf{(4-4.12)}$$

2 eq., 3 unk. [T(t)]

An energy balance on the contents of the reactor provides another equation:

$$f\rho C_p T_i(t) - V r_A(t)(\Delta H_r) - UA[T(t) - T_c] - f\rho C_p T(t) = V\rho C_v \frac{dT(t)}{dt} \qquad \textbf{(4-4.13)}$$

3 eq., 3 unk.

where

T_c = temperature of boiling liquid in cooling jacket, °R
ΔH_r = heat of reaction, Btu/lbmole of A reacted

Following the usual procedure, we obtain

$$\frac{C_A(s)}{C_{Ai}(s)} = \frac{\dfrac{K_1}{1 - K_2 K^{24}}(\tau_2 s + 1)}{\dfrac{\tau_1 \tau_2}{1 - K_2 K_4} s^2 + \dfrac{\tau_1 + 72}{1 - K_2 K_4} s + 1} \qquad \textbf{(4-4.14)}$$

$$\frac{C_A(s)}{\Gamma_i(s)} = \frac{\dfrac{-K_2 K_3}{1 - K_2 K_4}}{\dfrac{\tau_1 \tau_2}{1 - K_2 K_4} s^2 + \dfrac{\tau_1 + 72}{1 - K_2 K_4} s + 1} \qquad \textbf{(4-4.15)}$$

$$\frac{\Gamma(s)}{C_{Ai}(s)} = \frac{\dfrac{-K_1 K_4}{1 - K_2 K_4}}{\dfrac{\tau_1 \tau_2}{1 - K_2 K_4} s^2 + \dfrac{\tau_1 + 72}{1 - K_2 K_4} s + 1} \qquad \textbf{(4-4.16)}$$

and

$$\frac{\Gamma(s)}{\Gamma_i(s)} = \frac{\dfrac{K_3}{1 - K_2 K_4}(\tau_1 s + 1)}{\dfrac{\tau_1 \tau_2}{1 - K_2 K_4} s^2 + \dfrac{72}{1 - K_2 K_4} s + 1} \qquad \textbf{(4-4.17)}$$

where

$$K, = \frac{f}{f + Vk_oe^{-E/RT}}, \text{ dimensionless}$$

$$K_2 = \frac{Vk_oE\bar{c}_Ae^{-E/R\bar{T}}}{R\bar{T}^2(f + Vk_oe^{-E/R\bar{T}})}, \frac{\text{lbmoles A}}{\text{ft}^3\text{-}°R}$$

$$K_3 = \frac{f\rho C_p}{\dfrac{Vk_oE\bar{c}_A\Delta H_re^{-E/R\bar{T}}}{R\bar{T}^2} + f\rho C_p + UA}, \text{ dimensionless}$$

$$K_4 = \frac{Vk_o\Delta H_re^{-E/R\bar{T}}}{\dfrac{Vk_oE\bar{c}_A\Delta H_re^{-E/R\bar{T}}}{R\bar{T}^2} + f\rho C_p + UA}, \frac{°R}{\text{lbmoles A/ft}^3}$$

$$\tau_1 = \frac{V}{f + Vk_oe^{-E/R\bar{T}}}, \text{ min}$$

$$\tau_2 = \frac{V\rho C_v}{\dfrac{Vk_oE\bar{c}_A\Delta H_re^{-E/R\bar{T}}}{R\bar{T}^2} + f\rho C + UA}, \text{ min}$$

and the deviation variables are

$$C_A(t) = c_A(t) - \bar{c}_A \qquad C_{Ai}(t) = c_{Ai}(t) - \bar{c}_{Ai}$$

$$\Gamma(t) = T(t) - \bar{T} \qquad \Gamma_i(t) = T_i(t) - \bar{T}_i$$

As discussed in Section 2-3.2 (see Fig. 2-3.1), for a system to be stable, all the roots of the denominator of the transfer function must have negative real parts. Thus for the present chemical reactor, the roots are given by

$$\text{Roots } = \frac{-(\tau_1 + \tau_2) \pm [(\tau_1 + \tau_2)^2 - 4\tau_1\tau_2(1 - K_2K_4)]^{1/2}}{2\tau_1\tau_2} \qquad (4\text{-}4.18)$$

As we have learned in this and the previous chapter, for nonlinear systems the numerical values of the process parameters, τ_1, τ_2, K_2, and K_4, vary as the operating conditions, $\bar{c}_A$ and $\bar{T}$, vary. Thus the location of the roots and the stability itself also vary. Table 4-4.2 shows the roots as the operating conditions change. To generate this information, the energy removed from the reactor by the cooling side was varied by adjusting T_c to obtain the desired operating conditions of $\bar{c}_A$ and $\bar{T}$.

Let us analyze further the response of this chemical reactor. Figure 4-4.6a shows the temperature and concentration responses to a change in $-5°R$ in inlet temperature; these responses are oscillatory around a temperature of $684°R$ and a concentration of 0.102 lbmole/ft³. Table 4-4.2 shows that at this temperature and concentration, the roots are at $-0.0819 \pm 0.2341i$, indicating stable (negative real parts) and oscillatory (imaginary parts) responses. Figure 4-4.6b shows the responses to a change of $-10°F$ in inlet temperature. In this case the temperature starts to decrease, and the concentration to increase, with apparently no bounds. Table 4-4.2 shows that at a temperature of about $668°R$, the roots have positive real parts, indicating unstable behavior, or what is com-

Table 4-4.2 Roots Versus Operating Conditions

c_A, lbmoles/ft^3	T, °R	Roots	c_A, lbmoles/ft^3	T, °R	Roots
0.8905	560.08	− 0.1018, − 0.1577	0.4159	644.00	− 0.0805, + 0.3032
0.8868	566.00	− 0.1026, − 0.1515	0.3499	650.00	− 0.0749, + 0.3275
0.8815	572.00	− 0.1040, − 0.1425	0.2526	660.00	− 0.0564, + 0.3063
0.8738	578.00	− 0.1080, − 0.1279	0.1894	668.00	− 0.1249, + 0.1987
0.8671	582.00	− 0.1135 ± 0.0096i	0.1629	672.00	− 0.0647 ± 0.06691'
0.8588	586.00	-0.1079 ± 0.0165i	0.1020	684.00	− 0.0819 ± 0.23411'
0.8485	590.00	− 0.1012 ± 0.0203i	0.0802	690.00	− 0.1980 ± 0.25021'
0.8287	596.00	− 0.0886 ± 0.021 li	0.0630	696.00	-0.3506 ± 0.17671'
0.7270	614.00	− 0.0882, + 0.03 16	0.0536	700.00	− 0.6491, − 0.3037
0.6764	620.00	− 0.0884, + 0.0871	0.033 1	712.00	− 1.7969, − 0.2095
0.638 1	624.00	− 0.0881, + 0.1265	0.0241	720.00	− 2.8438, − 0.1941
0.5530	632.00	− 0.0864, + 0.2061	0.0176	728.00	− 4.2452, − 0.1859

monly referred to as **an open-loop unstable response.** In theory, this decrease in temperature, and increase in concentration, should continue. However, Fig. **4-4.6b** shows that eventually the temperature stabilizes at **566°R**, and the concentration at 0.8393 lbmole/ft^3. Table 4-4.2 shows that at these conditions, the roots regain their negative real components. At **566°R** the reaction is "quenched," indicating no conversion. At the final temperature of **566°R**, the transfer functions are written as

$$\frac{\Gamma(s)}{\Gamma_i(s)} = \frac{K_3(9.79S + 1)}{(9.75s + 1)(6.6s + 1)} \quad \text{and} \quad \frac{C_A(s)}{\Gamma_i(s)} = \frac{-K_2K_3}{(9.75s + 1)(6.6s + 1)}$$

However, at T = 620°R, which is an unstable operating condition, the transfer functions are written as

$$\frac{\Gamma(s)}{\Gamma_i(s)} = \frac{K_3(7.47S + 1)}{(11.3s + 1)(11.47s - 1)} \quad \text{and} \quad \frac{C_A(s)}{\Gamma_i(s)} = \frac{-K_2K_3}{(11.3s + 1)(11.47s - 1)}$$

Recall that the τ values are the negative reciprocal of the roots.

To explore this behavior further, consider the following transfer function:

$$\frac{Y(s)}{X(s)} = \frac{1}{1 - \tau s}$$

The response to a step change in input, X(s) = l/s, is given in the time domain by

$$Y(t) = 1 - e^{t/\tau}$$

The positive exponential term indicates an unbounded response.

Textbooks in reactor design (Levenspiel, 1972; Fogler, 1992) commonly discuss the important concept of stability from a steady-state point of view; Fig. **4-4.7a** shows a
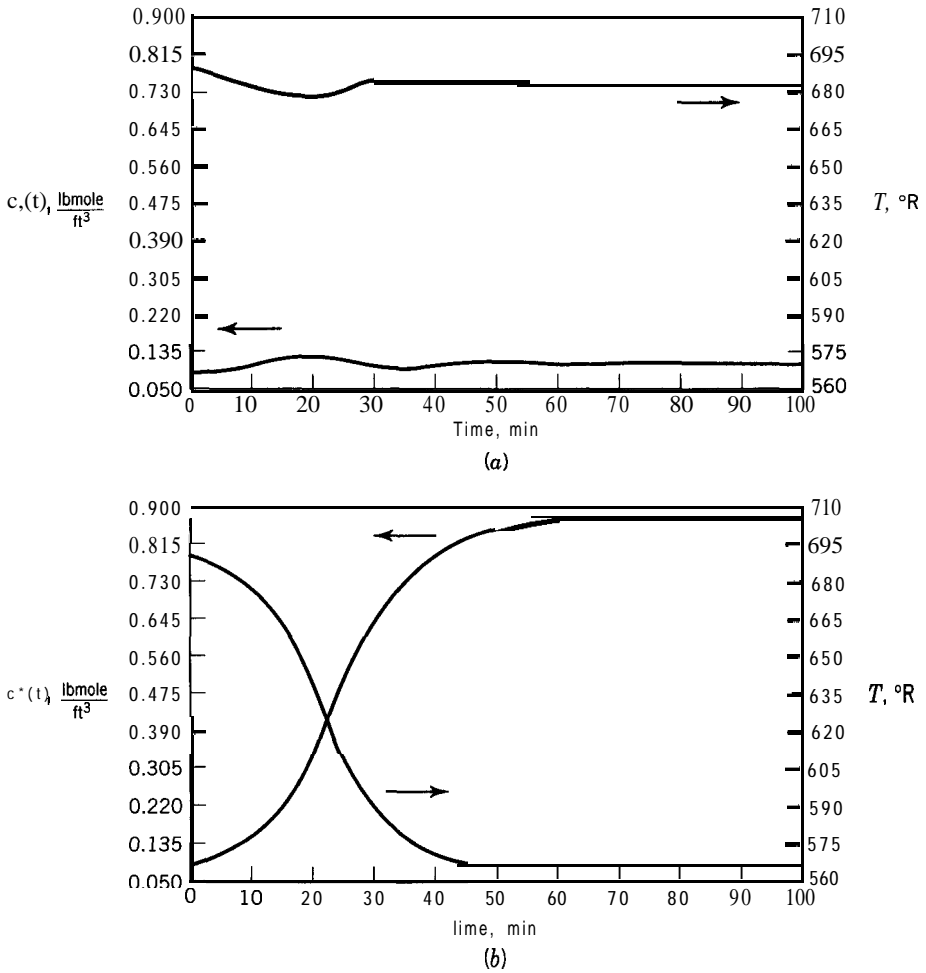
Figure 4-4.6 (a) Response of temperature and concentration to a change of ─ 5°R in inlet temperature. (b) Response of temperature and concentration to a change of 10°R in inlet temperature.

typical graph. The figure, generated using the steady-state values of Table 4-4.1, shows a plot of the heats generated and removed versus the temperature in the reactor. There are three points (A,, B_1, and C,) where the heats are equal (balanced). Only two of these points, A, and C_1, represent stable operating conditions; point B_1 is an unstable condition. At any point to the left of B_1, the heat generated is less than the heat removed, so the temperature in the reactor will decrease until both heats are equal, which occurs at point A,. At any point to the right of B,, the heat generated is greater than the heat removed, so the temperature will increase until both heats are equal, which occurs at point C,. Around point B,, the temperature moves quickly away from that operating condition. Temperatures 573°R and 690°R are the stable conditions. Figure 4-4.7b shows the same graph except that now the inlet temperature has been changed by ─ 5°R
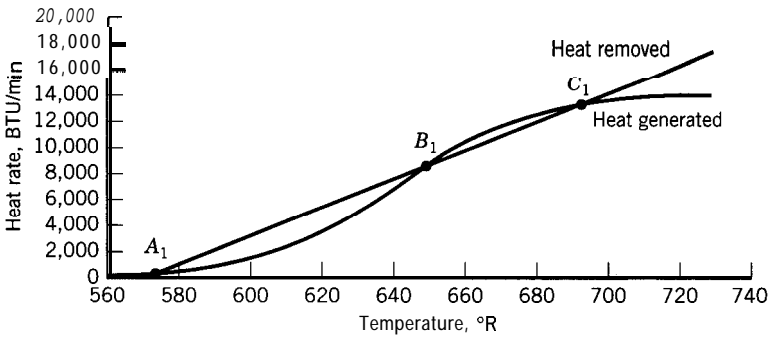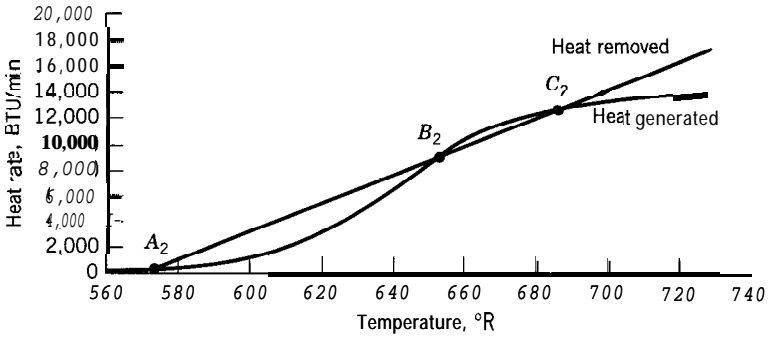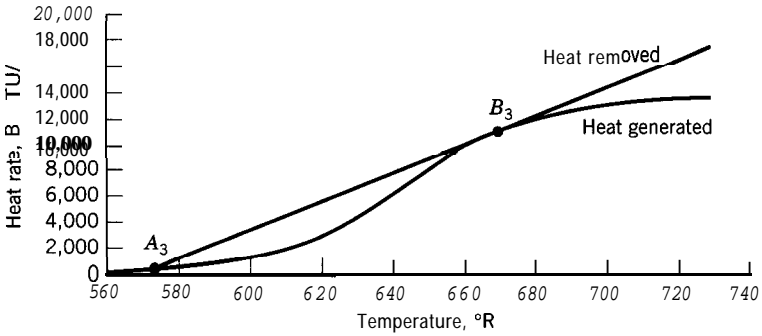
Figure 4-4.7 Steady-state analysis of heat generated and heat removed in chemical reactor. (a) Inlet temperature 578°R. (b) Inlet temperature 573°R. (c) Inlet temperature 568°R.

to 573°R. The figure shows that the new stable conditions are at 569°R and 684°R. Points A, and C_2 are the stable conditions, whereas point B_2 represents the unstable condition; the temperature moves quickly away from point B_2. Figure 4-4.7c shows the graph when the inlet temperature is changed by $-10°R$ to 568°R. In this case there are only two conditions where the heat generated and the heat removed are equal. Point A, represents the only stable condition; at point B_3, the curves just touch each other, which

indicates an unstable condition. The three figures show that for this reactor, the stable operating conditions "move away" from a temperature range from $615°R$ to $670°R$; stable conditions occur below $615°R$ or above $670°R$. This same indication of stability can be obtained by observing the roots given in Table 4-4.2; at any temperature between $615°R$ and $670°R$, there is always one positive root, indicating unstable behavior. The roots also indicate the oscillatory nature of the behavior; this characteristic is not indicated in the steady-state analysis.

In the example at hand, the unstable behavior occurred when the temperature decreased. In this case, the extreme operating condition reached was that of "quenching" the reaction. In some other reactions, however, the unstable behavior may occur when the temperature increases (see Problem 4-17). That is, in these reactions if the positive real part(s) persist(s), then the temperature in the reactor will continue to increase, theoretically without bound, and the concentration to decrease. Very often, engineers refer to this rapid increase in temperature as a run-away reaction. Several things may happen if this occurs. If the temperature increases beyond the maximum safe temperature limit for which the reactor was designed, then an explosion, a melt-down, or the like may occur. *To prevent this unsafe operation, safety overrides must be triggered soon enough and must have the capacity to stop the process.* If the temperature does not reach the maximum safe limit, and it is left alone, it may actually reach a new steady state. This occurs because as the reactants are depleted, the heat generated reaches a limit; it does not continue to increase. At this moment, the heat removed may become equal to the heat generated by the reaction. If so, a new steady-state is obtained; this condition is essentially points C_1 and C_2 in Figs. **4-4.7a** and **b.** A type of reactor where this depletion does not easily occur is a nuclear reactor. The nuclear rods-the fuel-do not deplete easily, and thus there are plenty of reactants available.

Systems described by transfer functions with a $(\tau s - 1)$ or a $(1 - \tau s)$ term in the denominator **are** referred to as **open-loop unstable.** They are also sometimes classified as **non-self-regulating** systems, because as long as the roots with positive real parts persist, these systems will not reach a steady-state condition. The most common example of open-loop unstable behavior is an exothermic reaction.

The design of chemical reactors where exothermic reactions occur is most important, and it affects their controllability. These reactors must have enough cooling capacity to avoid "run-away reactions," and the material of which they are constructed must be able to sustain high temperatures for safe operation. However, other non-safe operations may develop sometimes even before a high temperature is reached. Suppose, for example, that beyond a certain temperature, a new reaction starts to produce a toxic chemical   compound.

4-4.3   Inverse   Response   Processes: Chemical   Reactor

In Section 4-2.3 we considered a nonisothermal chemical reactor. Several transfer functions, Eqs. 4-2.28 through 4-2.42, were developed. Chapter 2 presented the response of transfer functions, similar to Eqs. 4-2.30, 4-2.35, 4-2.37, 4-2.38, and 4-2.40 through 4-2.42, to a step change in forcing function. This section presents and analyzes the response of Eqs. 4-2.29, 4-2.31, and 4-2.39 to the same type of forcing function. That is, we wish to look at the response of systems described by transfer functions with the term $\tau s - 1$ in the numerator.

Figure 4-4.8 shows the response of the temperature (Eq. 4-2.39) and of the concentration (Eq. 4-2.31) to a step change in process flow. It is interesting to note that both initial responses are in the opposite directions from the final ones. That is, the concentration initially starts to increase and then decreases. The temperature first tends to decrease and then increases. This type of response is called *inverse response;* certainly there is an explanation for it. Realizing that the inlet stream is colder than the contents of the reactor, we should not be surprised that when the inlet flow increases, its initial effect is to reduce the temperature in the reactor. Similarly, because the concentration of the inlet stream is greater than that in the reactor, the initial effect of an increase in inlet flow is to increase the concentration in the reactor. The effect of a lower temperature in the reactor is to reduce the rate of reaction, whereas the effect of a higher concentration is to increase it. Thus the lower temperature and the higher concentration represent two opposing effects. The final response is the net result of **these** opposing effects. As the figures show, the temperature eventually increases and the concentration decreases, indicating that the rate of reaction increases until a new steady state is reached.

Mathematically, the inverse response behavior is represented by a positive root in the numerator of the transfer function. As we shall see in Chapter 8, roots of the numerator of transfer functions are called *zeros.* Equation 4-2.39 has a zero at $+ 0.361$, due to $(2.77s - 1)$, and Eq. 4-2.31 has a zero at $+ 0.0223$, due to $(44.75s - 1)$. Similarly, by inspection of Eq. 4-2.29, we realize that the outlet concentration, $c_A(t)$, will exhibit an inverse response when the inlet concentration, $c_{Ai}(t)$, changes.

As we have said, the inverse response can be thought of as the net result of two opposing effects. This phenomenon can be expressed mathematically as two parallel first-order systems with gains of opposite signs; this is shown in Fig. 4-4.9. From the figure, the following transfer function can be obtained:

$$\frac{\Gamma(s)}{F(s)} = \frac{K_2}{\tau_2 s + 1} - \frac{K_1}{\tau_1 s + 1}$$

$$\frac{\Gamma(s)}{F(s)} = \frac{K_2 \tau_1 s + K_2 - K_1 \tau_2 s - K_1}{(\tau_2 s + 1)(\tau_1 s + 1)}$$

$$\frac{\Gamma(s)}{F(s)} = \frac{(K_2 \tau_1 - K_1 \tau_2)s + (K_2 - K_1)}{(\tau_2 s + 1)(\tau_1 s + 1)}$$

This equation provides an inverse response when process 1 reacts faster than process 2, that is, when $\tau_1 < \tau_2$. In addition, the gain of process 2 must be larger than that of process 1, that is, $|K_2| > |K_1|$. Under these conditions, the numerator of the transfer function has its root at

$$s = - \frac{K_2 - K_1}{K_2 \tau_1 - K_1 \tau_2}.$$

Not all chemical reactors exhibit the inverse response behavior. Other common processes that exhibit this type of response are fluidized coal gasifiers, where increased combustion air flow first expands the bed and then consumes the material at a faster
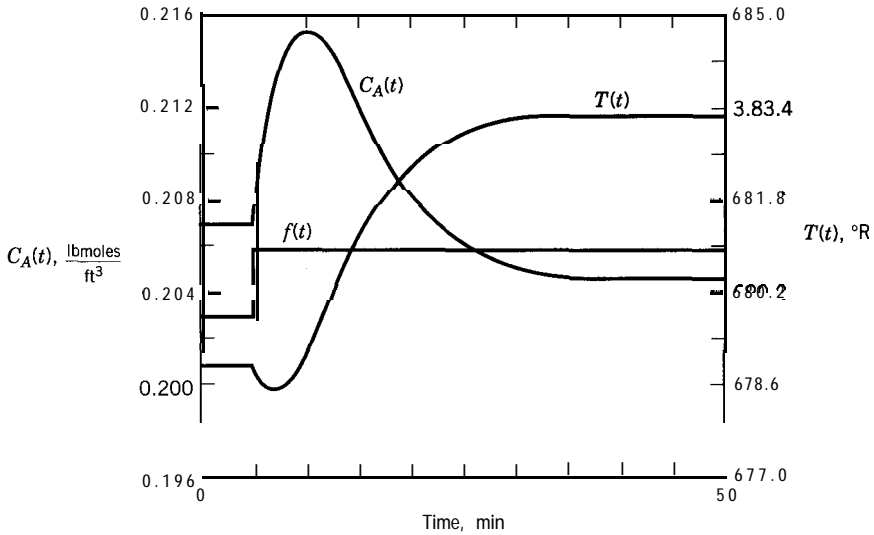
Figure 4-4.8 Inverse response of concentration and temperature to a change in process flow of 0.15 ft³/min.

rate; distillation columns (Buckley *et* al., 1975); and the water level in a boiler drum. The control of processes that exhibit inverse response presents a challenge to the control engineer (Iinoya and Altpeter, 1962). Chapter 12 considers the control of water level in a boiler drum and shows the extra control sometimes used to "counteract" the response and provide the control performance required.

4-5 SUMMARY

This chapter has investigated the steady-state and dynamic characteristics of processes described by higher-order differential equations. It presented the development of the mathematical models, transfer functions, and block diagrams of these processes. We found that most processes are described by higher-order transfer functions. One of the most important facts presented, as far as process control is concerned, is that as the order of the system increases, the apparent dead time also increases. This fact was clearly shown in Section 4-3, Fig 4-3.1. This is one of the two most common reasons
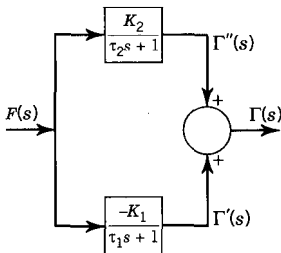
Figure 4-4.9 Explanation of inverse response.

for the occurrence of dead time in processes; the other is transportation delays, as shown in Chapter 3. Figure 4-3.1 also shows an important difference between the response of first-order and higher-order systems. For first-order systems, the steepest slope on the response curve, to a step change in input, is the initial one. For higher-order systems, this is not the case; the steepest slope occurs later on the response curve.

The chapter also introduced and explained the significance of noninteracting and interacting systems.

Another important concept presented was the meaning of the different terms in the transfer functions. The term $(\tau s + 1)$ in the denominator of a transfer function increases the order of the system and slows the response, as shown in Fig. 4-3.1. In this case, we refer to $(\tau s + 1)$ as a first-order lag or simply as a lag. However, when the term $(\tau s + 1)$ is present in the numerator of a transfer function, it indicates a faster response, as shown in Fig. 4-3.2. In this case we refer to $(\tau s + 1)$ as a first-order lead or simply as a lead. The presence of an s term in the denominator of a transfer function indicates an integrating system, as shown in Fig. 4-4.1. When the term $(\tau s - 1)$ is present in the denominator of a transfer function, it indicates an open-loop unstable system, as shown in Section 4-4.2. When the term $(\tau s - 1)$ is present in the numerator of a transfer function, it indicates an inverse response behavior, as shown in Section 4-4.3.

4-6 OVERVIEW OF CHAPTERS 3 AND 4

Chapters 3 and 4 complete our presentation of the types, behavior, and characteristics of processes. All of the characteristic terms, gains, time constants, and dead times were obtained starting from first principles, usually mass and energy balances. Sometimes, however, it is difficult to obtain them as we have done in these two chapters. This difficulty is mainly because of the complexity of the processes and the lack of knowledge or understanding of some physical or chemical properties. In such cases, we must use empirical means to obtain these terms. Some of these methods are presented in Chapters 7, 9, and 14.

As indicated in Chapter 3, we must understand the processes before we can design control systems. Now that we have completed our study of processes, we are ready to control them. Chapter 5 discusses some aspects of the sensor and transmitter combination, control valves, and different types of feedback controllers. Finally, Chapter 6 and subsequent chapters put everything together. Prepare to encounter exciting and challenging material. These chapters show you how to design control systems that will ensure that your processes are safe to operate and, at the same time, produce a product of the desired quality at the design rate.

REFERENCES

1. Buckley, P. S., R. K. Cox, and D. L. Rollins. 1975. "Inverse Response in a Distillation Column." *Chemical Engineering Progress* June.
2. Fogler, H. *S.* 1992. *Elements of Chemical Reaction Engineering.* Englewood Cliffs, N.J.: Prentice-Hall,   Section   8-7.
3. Iinoya, K., and R. J. Altpeter. 1962. "Inverse Response in Process Control." *Industrial and Engineering   Chemistry*   July.
4. Levenspiel, *0.* 1972. *Chemical Reaction Engineering.* New York: Wiley, Chapter 8.
5. Shinskey, F. *G.* 1988. *Process Control Systems.* 3rd ed., New York: McGraw-Hill, p. 47.

PROBLEMS

4-1. Consider the two gas tanks shown in Fig. P4-1. The gas may be assumed to be isothermal and to behave as an ideal gas such that the density in each tank is related to the pressure in that tank by the formula

$$\rho(t) = \frac{Mp(t)}{RT}$$

where

$\rho(t)$ = gas density, lb/ft^3
M = gas molecular weight, $lb/lbmole$
R = ideal gas constant, 10.73 ft^3-psia/lbmole-°R
T = gas temperature, °R
$p(t)$ = pressure in the tank, psia

The gas completely fills the volumes of the tanks, which are constant.

(a) *Critical (choked) flow through* the valves. If the flow (in lb/min) through the valves is assumed to be critical, or "choked," it is proportional to the upstream pressure for each valve.

$$w_1(t) = k_{v_1}p_1(t) \qquad w_2(t) = k_{v_2}p_2(t)$$

where k_{v_1} and k_{v_2} are constant valve coefficients [in (lb/min)/psia], which depend on the valve capacity, gas specific gravity, temperature, and valve design. Choked flow is discussed in detail in Appendix C.

Obtain the transfer functions relating the pressure in each tank to the inlet flow to the first tank. Draw a block diagram showing the pressures. If there were n identical tanks in series, all having the same volumes, and all valves having the same valve coefficients, what would be the transfer function for the pressure in the last tank to the flow into the first tank, $P_n(s)/W_i(s)$?

(b) *Subcritical flows through the valves.* If the flows through the valves are subcritical, then they are given by

$$w_1(t) = k_{v_1}\sqrt{p_1(t)[p_1(t) - p_2(t)]}$$

$$w_2(t) = k_{v_2}\sqrt{p_2(t)[p_2(t) - p_3]}$$

where the valve coefficients k_{v_1} and k_{v_2} are not numerically the same as for critical flow, and the discharge pressure p_3 may be assumed constant.

Obtain the transfer functions relating the pressure in each tank to the inlet flow to the first tank. Draw the block diagram for the tanks, showing the
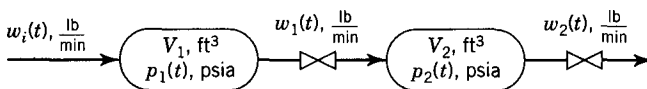
Figure P4-1 Gas tanks for Problem 4-1.
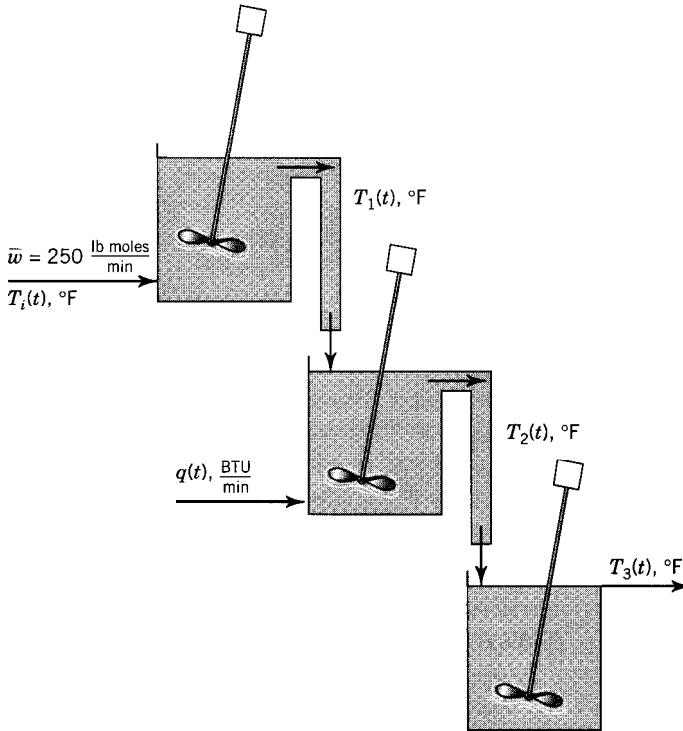
Figure **P4-2** Tanks for Problem 4-2.

transfer function of each block. Write the overall transfer function, $P_2(s)/W_i(s)$, and the formulas for the effective time constants and for the steady-state gains of the transfer functions in terms of the process parameters. *Note:* The effective time constants are defined as the negative reciprocal roots of the denominator of the overall transfer function.

4-2. Consider the process shown in Fig. **P4-2**. The mass flow rate of liquid through the tanks is constant at 250 lb/min. The density of the liquid may be assumed constant at 50 lb/ft^3, and the heat capacity may also be assumed constant at 1.3 Btu/lb-°F. The volume of each tank is 10 ft^3. You may neglect heat losses to the surroundings. It is desired to know how the inlet temperature, $T_i(t)$, and the heat transfer, $q(t)$, affect the outlet temperature, $T_3(t)$. For this process develop the mathematical model, determine the transfer functions relating $T_3(t)$ to $T_i(t)$ and $q(t)$, and draw the block diagram. Give the numerical values and units of each parameter in all transfer functions.

4-3. Consider the process described in Section 3-5 and shown in Fig. 3-5.1. In that problem, the relation expressing the flow provided by the fan and the signal to the fan is algebraic. This means that the fan does not have any dynamics, that is, the fan is instantaneous. In reality this is not the case. Let us assume that the fan has some dynamics such that the flow responds to a change in signal as a first-order response with a time constant of 10 s. Obtain the same information as was required in Section 3-5.
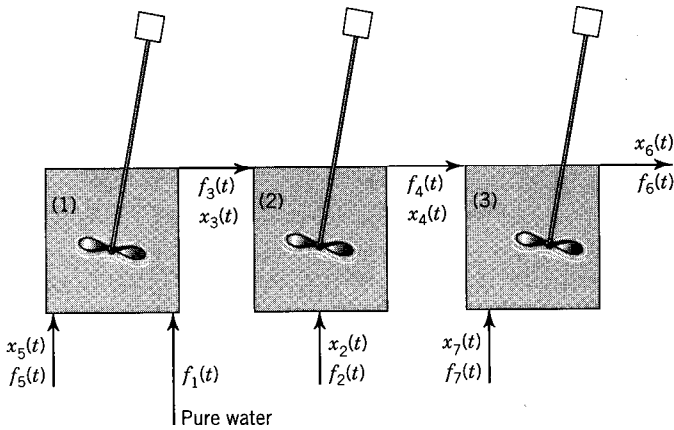
Figure **P4-3** Mixing tanks for Problem 4-4.

4-4. Consider the process shown in Fig. P4-3 where different streams are mixed. Streams 5, 2, and 7 are solutions of water and component A; stream 1 is pure water. The steady-state values for each stream are given in Table P4-1. Determine the following transfer functions, with the numerical values for every term.

$$\frac{X_6(s)}{X_5(s)}, \quad \frac{X_6(s)}{X_2(s)}, \quad \text{and} \quad \frac{X_6(s)}{F_1(s)}$$

4-5. Consider the two stirred tank reactors in series with recycle shown in Fig. P4-4.

Table P4-1 Process Information and Steady-State Values for Problem 4-4

Information

Tank volumes: $V, = V_2 = V_3 = 7000$ gal

Concentration transmitter range: 0.3 to 0.7 mass fraction. This transmitter's dynamics can be described by a dead time of 2 min.

Valve: The flow is proportional to the fraction valve position in the range of 0 to 3800 gpm. The valve dynamics can be considered negligible.

The density of all streams can be considered similar and constant.

Steady-State Values

Stream	Flow, gpm	Mass Fraction
1	1900	0.000
2	1000	0.990
3	2400	0.167
4	3400	0.409
5	500	0.800
6	3900	0.472
7	500	0.900
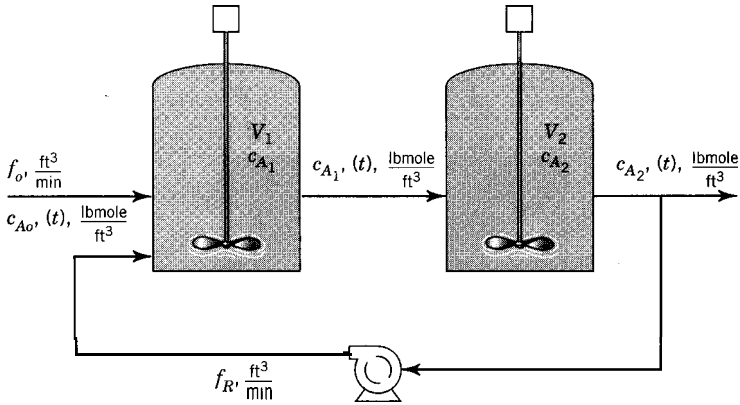
Figure **P4-4** Reactors in series for Problem 4-5.

You may assume the following:

- Each reactor is perfectly mixed and the temperature is constant.
- The reactor volumes, V_1 and V_2, are constant, and so is the density of the reacting mixture.
- The flow into the first reactor, f_o, and the recycle flow, f_R, are constant.
- The chemical reaction is elementary first-order, so its rate of reaction is given by the expression

$$r_A(t) = kc,(t), \text{ lbmoles/ft}^3\text{-min}$$

where

$$c_A(t) = \text{concentration of reactant A, lbmoles/ft}^3$$
$$k = \text{constant reaction rate coefficient, min}^{-1}$$

- The reactors are initially at steady state with an inlet concentration $c_{Ao}(0)$.
- The transportation lag between the reactors and in the recycle line is negligible.

(a) Determine the process transfer functions.

(b) Draw the block diagram for the two reactors.

(c) Use block diagram algebra to determine the transfer function $\dfrac{C_{A_2}(s)}{C_{Ao}(s)}$ for the two reactors.

(d) Determine the gain and the effective time constants of this transfer function, in terms of the parameters of the system: $V, , V_2, f_o, f_R$, and k.

(e) Answer the following questions:

   (i) Can the system be unstable (negative effective time constants)?

   (ii) Can the response of the composition be underdamped (complex conjugate effective time constants)?

   (iii) What do the effective time constants become as the recycle flow, f_R, becomes much larger than the inlet flow, f_o?

4-6. Consider the process shown in Fig. P4-5. The following information is known about the process:
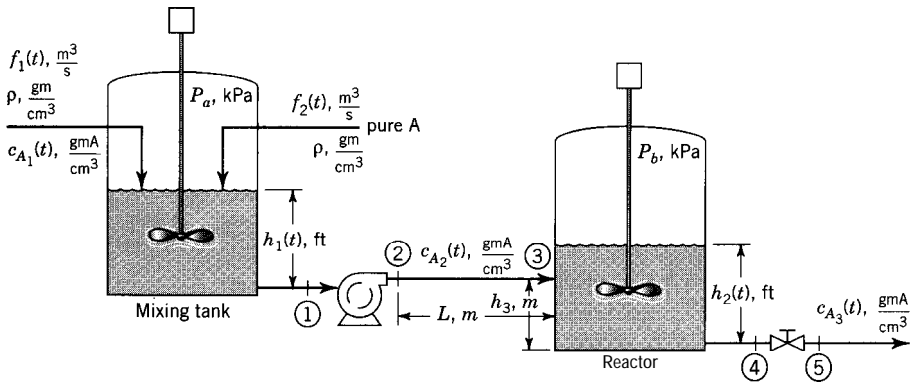
Figure P4-5 Process for Problem 4-6.

- All streams are approximately equal in density.
- The flow through the constant speed pump is given by

$$f(t) = A[1 + B\{p_1(t) - p_2(t)\}^2], \text{ m}^3/\text{s}$$

where A and B are constants.

- The pipe between points 2 and 3 is rather long with a length of L, m. The flow through this pipe is highly turbulent (plug flow). The diameter of the pipe is D, m. The pressure drop between points 2 and 3 is constant; it is ΔP, kPa.
- We may assume that the energy effects associated with the reaction (A → B) are negligible and that, consequently, the reaction occurs at a constant temperature. The rate of reaction is given by

$$r_A(t) = kc_A(t), \text{ kg/m}^3\text{-s}$$

- The flow through the outlet valve is given by

$$f(t) = C_v vp(t) \sqrt{h_2(t)}$$

Develop the mathematical model, and obtain the block diagram that shows the effect of the forcing functions $f_2(t)$, $vp(t)$, and $c_{A_1}(t)$ on the responding variables $h_1(t)$, $h_2(t)$, and $c_{A_3}(t)$, and draw the block diagram for this process.

4-7. Consider the fish tank of Problem 3-9. Given that a power agitator would not only annoy but also endanger the fish in the tank, it seems unreasonable to assume that the water in the tank is perfectly mixed, so we must remove that assumption. One way to do this is to divide the tank into a series of perfectly mixed volumes with recirculation of water between them. For simplicity, assume that the fish tank can be modeled by two perfectly mixed pools, each having half the total volume of the water in the tank, with one above the other (see Fig. P4-6). The recirculation is provided by natural convection and by the aeration pump that bubbles air through the water in the tank. Let the temperature at the bottom be $T_1(t)$, °C, let the temperature at the top be $T_2(t)$, let the recirculation rate be f,
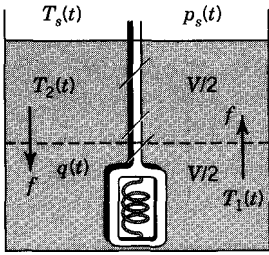
Figure P4-6 Fish tank for
Problem   4-7.

m^3/s, and model the tank under the same assumptions as before, modified as
follows:

- The water in each half of the tank is perfectly mixed.
- The heater is in the bottom half of the tank.
- The vapor pressure of the water is a function of the temperature in the top half,
  $p°[T_2(t)]$.

(a) Derive the equations that represent the response of the temperature in each
    half of the tank to changes in heat input from the electric heater, in the
    surrounding temperature, and in the surrounding water partial pressure. Lin-
    earize the equations and reduce them to two differential equations in devia-
    tions from the initial steady-state conditions.

(b) Derive the transfer functions of the temperatures in the tank to each of the
    three input variables and each other. Express the time constant and the gains
    of each equation in terms of the physical parameters. Draw the block diagram
    for the tank, showing both temperatures, and derive the overall transfer func-
    tions for T,(s) and $T_2(s)$, each in terms of the input variables only. Derive
    the expressions for the roots and the effective time constants of the response
    of the temperature to input changes.

4-8. The following irreversible elementary reaction takes place in the tank shown in
    Fig. P4-7.

$$A + B \rightarrow \text{Product}$$
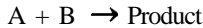
The rate of consumption of reactant A is given by

$$r_A(t) = -kc_A(t)c_B(t)$$

where $r_A(t)$ is the reaction rate, $\text{kmoles/m}^3\text{-min}$.
    The tank may be assumed to be perfectly mixed, and the temperature, volume,
and density of the reacting mixture may be assumed constant. The inlet reactant
concentrations, c_{AO} and c_{BO}, in kmoles/m^3, may also be assumed constant. The
variation of density with concentration may be neglected. Obtain the transfer
functions for the outlet concentrations, C,(s) and C,(s), to changes in the inlet
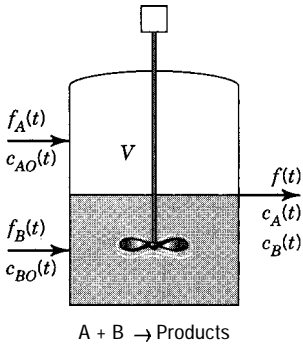
Figure P4-7 Reactor for Problem 4-8.

flows, $f_A(t)$ and $f_B(t)$, in m³/min, and draw the block diagram for the reactor, showing all transfer functions.

Obtain also the overall transfer function for $C_A(s)$ in terms of the input variables only [not including $C_B(s)$, but taking into consideration the interaction between the two concentrations]. Factor the denominator of the transfer function, and obtain the effective time constants (negative reciprocals of the roots) and the steady-state gain.

4-9. The tank shown in Fig. P4-8 is used for continuous extraction of a solute from a liquid solution to a solvent. One way to model the extractor is, as shown in the sketch on the right, by assuming two perfectly mixed phases, the extract and the raffinate, separated by an interface across which the solute diffuses at a rate given by

$$n(t) = K_a V[c_1(t) - c_1^*(t)]$$

where n(t), kmoles/s, is the rate of solute mass transfer across the interface; K_a,
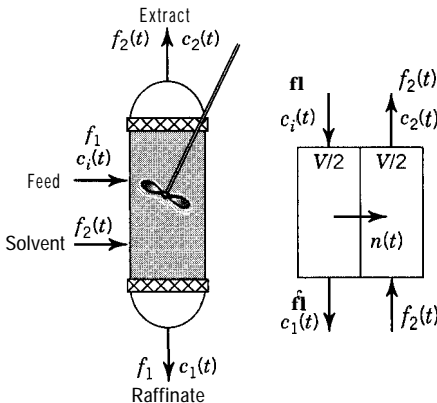
Figure P4-8 Extraction process for Problem 4-9.

s^{-1}, is the coefficient of mass transfer; V, m³, is the contact volume; $c_,(t)$, kmoles/m³, is the solute concentration in the raffinate phase; and $c_1^*(t)$, kmoles/m³, is the raffinate concentration that would be in equilibrium with the extract phase. The equilibrium relationship can be expressed as a straight line:

$$c_1^*(t) = mc_2(t)$$

where m is the slope of the equilibrium line, and $c_2(t)$ is the concentration of the solute in the extract phase. For simplicity, you may assume that the volume of each phase is half the total contact volume and that the feed flow, f_1 , is constant. The two input variables are the feed concentration, $c_i(t)$, and the flow of pure solvent, $f_2(t)$. You may also assume that the variation of the densities of the streams with concentration can be neglected.

Derive the transfer functions of the extractor, draw the block diagram for the extractor, and obtain the overall transfer functions for the composition of each phase in terms of the input variables. Factor the denominator of the overall transfer functions for the extractor, and express the roots in terms of the process parameters. Can the response of the concentrations be oscillatory? Can it be unstable? Justify your answers by analyzing the expressions for the roots.

4-10. A jacketed stirred tank is used to cool a process stream by causing cooling water to flow through the jacket as shown in Fig. P4-9. The process input variables to be considered are the flow of cooling water, $f_c(t)$, m³/min, and the inlet temperature of the process stream, $T_i(t)$, "C. The process output variables of interest are the outlet temperatures of the process and the water, T(t) and $T_c(t)$, "C, respectively.
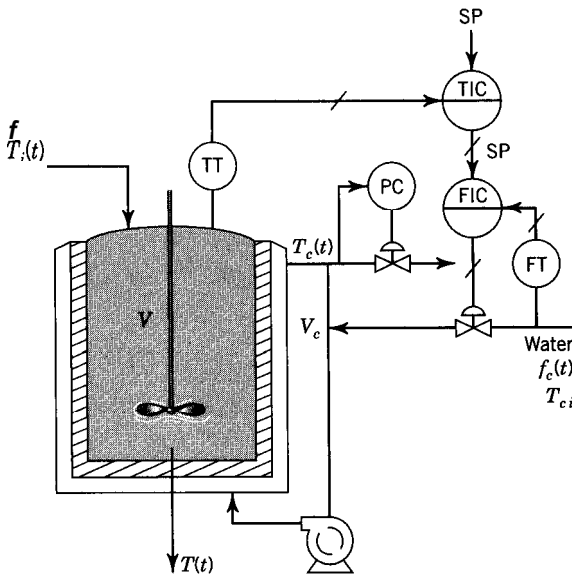
Figure P4-9 Jacketed stirred tank reactor for Problem 4-10.
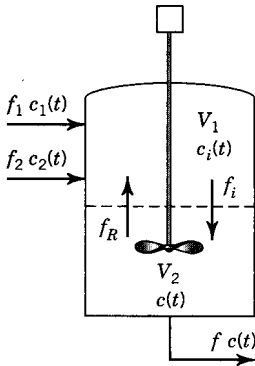
Figure **P4-10** Mixing tank for Problem 4- 11.

(a) List the necessary assumptions, and derive, from basic principles, the following differential equations that represent the dynamic response of the process.

$$\frac{dT(t)}{dt} = \frac{f}{V}[T_i(t) - T(t)] - \frac{UA}{V\rho c_v}[T(t) - T_c(t)]$$

$$\frac{dT_c(t)}{dt} = \frac{f_c(t)}{V_c}[T_{ci} - T_c(t)] + \frac{UA}{V_c\rho_c c_{vc}}[T(t) - T_c(t)]$$

where U, J/min-m²-°C, is the overall heat transfer coefficient and A, m², is the area of heat transfer to the jacket.

(b) Laplace-transform the equations (after linearizing them) and derive the transfer functions of the process. Draw the block diagram and obtain the overall transfer function for the temperature of the fluid leaving the tank. Factor the denominator of the overall transfer function and determine the expressions for the roots as functions of the process parameters. Can the response of the concentrations be oscillatory? Can it be unstable? Justify your answers by analyzing the expressions for the roots.

4-11. One way to model imperfect mixing in a stirred tank is to divide the tank into two or more perfectly mixed sections with recirculation between them. Assume that we divide the blending tank of Problem 3-18 into two perfectly mixed volumes V_1 and V_2 (so that $V_, + V_2 = V$), as shown in Fig. P4-10, where $c_i(t)$ and $c(t)$ are the concentrations of the solute in the two sections, respectively. For simplicity, assume that the inlet flows and the recirculation flow f_R, volumes, and density are constant. Show that the transfer function of the outlet concentration to either inlet concentration (use stream 1 as an example), is given by

$$\frac{C(s)}{C_,(s)} = \frac{K_1}{(\tau_1 s + 1)(\tau_2 s + 1)} \quad \frac{}{K_R}$$

where $K, = \dfrac{f_1}{f + f_R}, \tau_1 = \dfrac{V_1}{f + f_R}, \tau_2 = \dfrac{V_2}{f + f_R},$ and $K_R = \dfrac{f_R}{f + f_R}.$
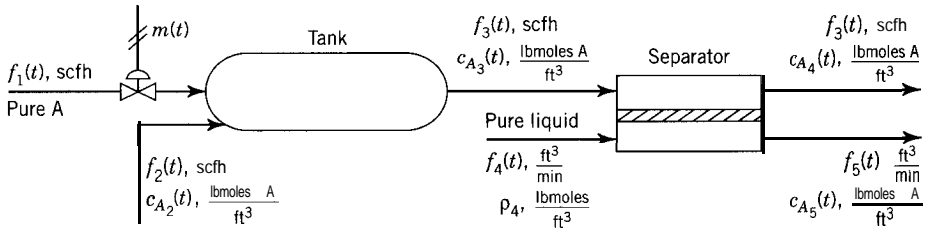
Figure P4-11 Process for Problem 4-12.

Calculate the parameters of the transfer function using the numbers given in Problem 3-18, assuming that the two volumes are equal and that the recirculation flow is (a) zero, (b) f, and (c) $5f$. For each of these cases, calculate the gain and the effective time constants of the transfer function (the negative reciprocals of the roots of the denominator). To what values do the effective time constants go when the recirculation flow becomes very large? How does this result compare to the result of Problem 3-18?

4-12. Consider the process shown in Fig. P4-11. A gas stream $f_2(t)$ enters a tank where it is mixed with another stream, f,(t), which is pure A. From the tank, the gas mixture flows into a separator where component A in the gas diffuses out, through a semipermeable membrane, to a pure liquid. The following may be assumed:

● The pressure drop across the valve is constant. The flow of pure A through this valve is given by

$$f_1(t) = k_v vp(t)$$

where f,(t) is in scfh. The valve position, $vp(t)$, is related to the signal, $m(t)$, by

$$vp(t) = \frac{1}{100} m(t)$$

● The output volumetric flow, $f_3(t)$, from the tank is equal to the sum of the input flows. The gas behaves as an incompressible fluid.
● The gas inside the tank is well mixed.
● The gas side in the separator is assumed to be well mixed. The liquid side is also assumed to be well mixed.
● The rate of mass transfer across the semipermeable membrane is given by

$$N_A(t) = A_X K_A[c_{A_4}(t) - c_{A_3}(t)]$$

where

$N_A(t)$ = rate of mass transfer, lbmoles A/h
$A,$ = cross-sectional area of membrane, ft^2
K_A = overall mass transfer coefficient, ft/h
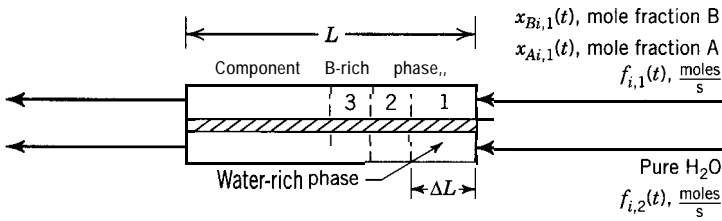
Figure P4-12 Environmental unit for Problem 4-13.

- The amount of component A diffused to the liquid does not significantly affect the gas volumetric flow. Therefore, the gas flow out from the separator can be considered equal to the input flow.
- The amount of component A diffused to the liquid does not significantly affect the liquid volumetric flow. The liquid streams entering and leaving the separator can be assumed to be equal in density.

(a) Write the mathematical model for the tank.

(b) Write the mathematical model for the separator.

(c) Draw the block diagram showing how the output variables, $c_{A_4}(t)$ and $c_{A_5}(t)$, are affected by $m(t), f_2(t)$, and $c_{A_2}(t)$. Also obtain the transfer functions.

4-13. Consider the environmental unit shown in Fig. P4-12. The purpose of this unit is to remove component A from a component B-rich phase. The transfer of A to the water medium occurs across a semipermeable membrane. In this process, the concentration of A is a function of a position along the unit and of time. Thus the equation that describes this concentration is a partial differential equation (PDE) in length and time. Systems described by PDEs are referred to as distributed systems. A common way to "get around" this PDE is to divide the unit into sections, or "pools," and to assume each pool to be well mixed. The dotted lines show the divisions of pools. Using this method, we find that the differential in length, dL, is approximated by ΔL. The smaller the pools, the better the approximation; however, there is a point of diminishing returns.

The mass transfer rate of component A is

$$N_A(t) = Sk_A[x_{An,1}(t) - x^*_{An,2}(t)]$$

where

$N_A(t) = $ moles of A transferred/s

$S = $ surface area, of membrane, across which the transfer takes place, m^2

$k_A = $ mass transfer coefficient, a constant, moles A/m^2-s

$x_{An,1}(t) = $ mole fraction of A in liquid phase 1 (component B-rich phase). The subscript n refers to the pool number.

$x^*_{An,2}(t) = $ mole fraction of A in liquid phase 2 (water-rich phase) that would be in equilibrium with $x_{An,1}(t)$
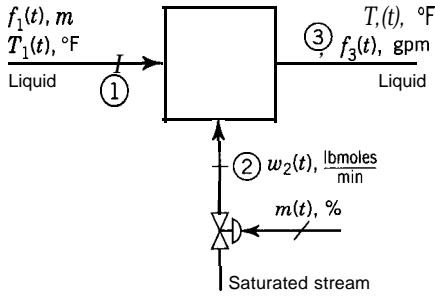
Figure **P4-13** Mixing tank for Problem 4-14.

Assume the equilibrium line is straight with slope m. Then

$$x^*_{An,2}(t) = mx_{An,2}(t)$$

where $x_{An,2}(t)$ = mole fraction of A in liquid phase 2. Component B and water are not transferred across the membrane, and the process occurs isothermally. Assume constant volumes and densities. Develop the mathematical model; determine the transfer functions relating the output variables $x_{A2,1}(t)$ and $x_{A2,2}(t)$ to the forcing functions $x_{Ai,1}(t), f_{i,1}(t)$, and $f_{i,2}(t)$; and draw the block diagram for this process. Only the first two pools are considered in this problem.

4-14. Consider the tank, shown in Fig. **P4-13**, where a fluid is mixed with saturated steam at 1 atm. The steam condenses in the liquid, and the tank is full all the time. The steady-state values and some process information are as follows:

$$\bar{f}_1 = 25 \; gpm \qquad \bar{T}_1 = 60°F \qquad \bar{w}_2 = 3.09 \; \frac{lbm}{min} \qquad \bar{T}_3 = 80°F \qquad \bar{m} = 50\%$$

$$\rho_1 = \rho_3 = 7 \; \frac{lbm}{gal} \qquad C_{p_1} = C_{p_3} = 0.8 \; \frac{Btu}{lbm\text{-}°F} \qquad Volume \; of \; tank = 5 \; gal$$

The flow through the valve is given by

$$w_2(t) = 1.954vp(t) \sqrt{\Delta P}$$

The pressure drop across the valve is a constant 10 psi. The valve position, $vp(t)$, is linearly related to the signal, $m(t)$. As the signal goes between 0 and 100%, the valve position goes between 0 and 1. The valve's dynamics can be described with a first-order time constant of 4 s. Develop the mathematical model, and obtain the transfer functions relating the temperature $T_3(t)$ to $f_i(t)$, $T_i(t)$, and $m(t)$. Be sure to provide the numerical values and units of all gains and the time constants.

4-15. Figure **P4-14** shows the responses of different processes to a step change in input. Give an indication of the form(s) of any possible transfer function(s) for each process.
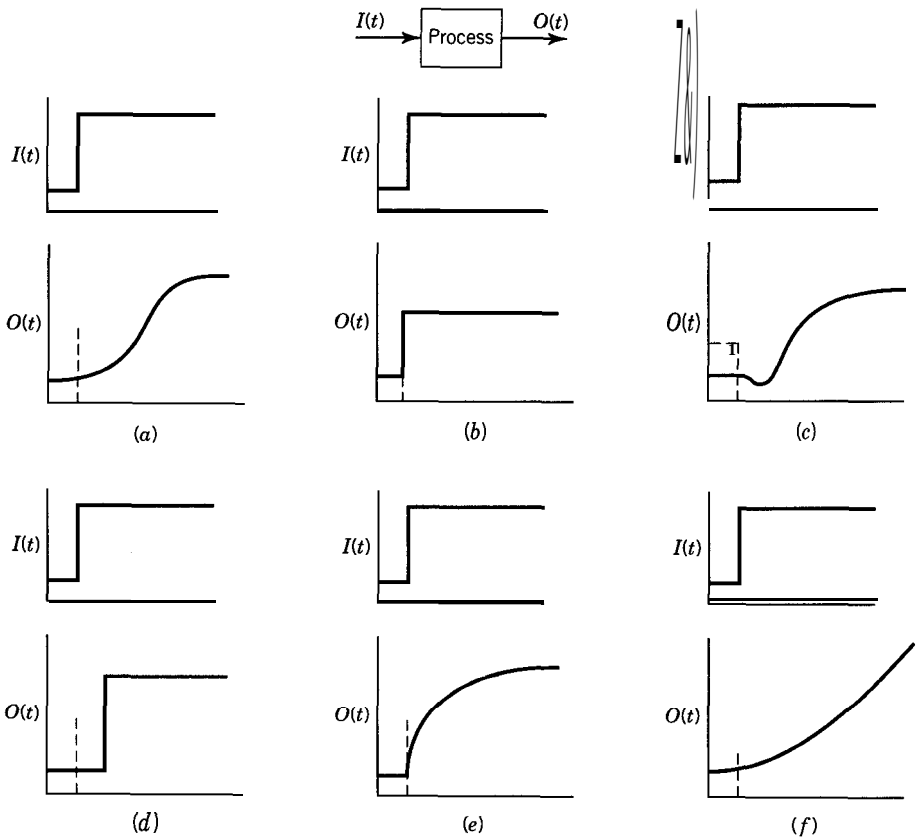
Figure **P4-14** Process responses for Problem 4-15.

Table P4-2 Process Information and Steady-State Values for Problem 4-17

Process Information

$V = 13.26$ ft^3 $\qquad A = 36$ ft^2

$E = 27,820$ Btu/lbmole $\qquad R = 1.987$ Btu/lbmole-°R

$\rho = 55$ lbm/ft^3 $\qquad C_p = 0.88$ Btu/lbm-°F

AH, $= -12,020$ Btu/lbmole $\qquad U = 75$ Btu/(hr-ft^2-°F)

$k_o = 1.73515$ X 10^8/min^{-1} for $0 < T < 200$°F

k_o varies linearly 1.73515 X 10^8/min to $1.735\ 15$ X 10^9/min for 200°F $< T < 205$°F

$k_o = 1.73515$ X 10^8/min for $T > 205$°F

Steady-State Values

$c_{Ai}(t) = 0.4471$ lbmole/ft^3 $\qquad T_i(t) = 635$°R

$T_c = 602.7$°R $\qquad f = 1.3364$ ft^3/min

$c_A(t) = 0.2536$ lbmole/ft^3 $\qquad T(t) = 690.0$°R

4-16. Consider the chemical reactor presented in Section 4-4.2. The section shows, in Fig. 4-4.6, the response of the temperature in the reactor to changes (decreases) in inlet temperature. What would be the response-stable or not, oscillatory or not-to *increases* in inlet temperature?

4-17. Consider again the chemical reactor presented in Section 4-4.2. Table P4-2 presents a new set of process and steady-state information. Obtain the stability information, similar to Table 4-4.2, as the temperature in the reactor changes between $650°R$ and $690°R$.

Chapter 5

Basic Components
of Control Systems

Chapter 1 presented the three basic components of control systems: sensors/transmitters, controllers, and final control elements. In that chapter, we learned that these components perform the three basic operations of every control system: measurement (M), decision (D), and action (A).

The present chapter takes a brief look at the sensor/transmitter combination, the M component, followed by a more detailed study of control valves, the A component, and feedback controllers, the D component. Appendix C presents numerous diagrams, schematics, and other figures to aid in the presentation of different types of sensors, transmitters, and control valves. Thus Appendix C complements this chapter, and the reader is encouraged to read it along with this chapter.

5-1 SENSORS AND TRANSMITTERS

The sensor produces a phenomenon-mechanical, electrical, or the like-related to the process variable it measures. The transmitter in turn converts this phenomenon into a signal that can be transmitted. Thus the purpose of the sensor/transmitter combination is to generate a signal, the transmitter output, that is related to the process variable. Ideally this relationship should be linear; that is, the transmitter output signal should be proportional to the process variable. Often this is the case, as for example with pressure, level, and some temperature transmitters, such as resistance temperature devices (RTDs). In other situations, the transmitter output is a known nonlinear function of the process variable, as for example with thermocouples and orifice flowmeters.

There are three important terms related to the sensor/transmitter combination. The range of the instrument is given by the low and high values of the process variable that is measured. Consider a pressure sensor/transmitter that has been calibrated to measure a process pressure between the values of 20 psig and 50 psig. We say that the range of this sensor/transmitter combination is 20 to 50 psig. The *span* of the instrument is the difference between the high and low values of the range. For the pressure instrument we have described, the span is 30 psi. The low value of the range is often referred to

197
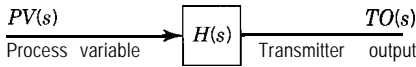
Figure 5-1.1 Block diagram of a sensor/transmitter combination.

as the *zero* of the instrument. This value does not have to be zero in order to be called the zero of the instrument. For our example, the zero of the instrument is 20 psig.

   Appendix C presents some of the most common industrial sensors: pressure, flow, temperature, and level. That appendix also briefly discusses the working principles of an electrical transmitter and of a pneumatic transmitter.

   The transfer function of the sensor/transmitter combination relates its output signal to its input, which is the process variable; this is shown in Fig. 5-1.1. The simplest form of the transfer function is a first-order lag:

$$H(s) = \frac{TO(s)}{PV(s)} = \frac{K_T}{\tau_T s + 1} \tag{51.1}$$

where

K_T = transmitter gain
τ_T = transmitter time constant

   When the relationship between the transmitter output (TO) and the process variable (PV) is linear, the transmitter gain is simple to obtain once the span is known. Consider an electronic pressure transmitter with a range of 0 to 200 psig. Figure 5-1.2 shows the output versus the process variable (input). From the definition of gain in Chapter 3, the gain of a linear transmitter can be obtained by considering the entire change in output over the entire change in input, which is the span of the transmitter.

$$K_T = \frac{(20 - 4)\ \text{mA}}{(200 - 0)\ \text{psig}} = \frac{16\ \text{mA}}{200\ \text{psig}} = 0.08\ \text{mA/psig}$$

or, in percent transmitter output (%TO)

$$K_T = \frac{(100 - 0)\ \%\text{TO}}{(200 - 0)\ \text{psig}} = 0.5\ \%\text{TO/psig}$$
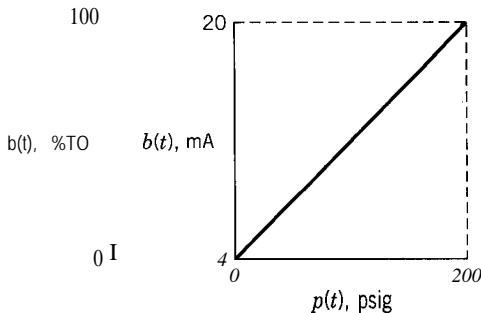
Figure S-l.2 Linear electronic pressure transmitter.

Thus the gain of a sensor/transmitter is the ratio of the span of the output signal to the span of the measured variable.

The preceding example assumed that the gain of the sensor/transmitter is constant over the complete operating range. For most sensor/transmitters this is the case, but there are some instances, such as a differential pressure sensor used to measure flow, when this is not so. A differential pressure sensor measures the differential pressure, h, across an orifice. Ideally, this differential pressure is proportional to the square of the volumetric flow rate, f. That is,

$$h \propto f^2$$

The equation that describes the output signal, in %, from a differential pressure transmitter when used to measure volumetric flow with a range of 0 to f_{max} gpm is

$$b = \frac{100}{(f_{max})^2} f^2$$

where

   b = output signal in %TO
   f = volumetric flow

From this equation, the local gain of the transmitter is obtained as follows:

$$K_T = \frac{db}{df} = \frac{2(100)}{(f_{max})^2} f$$

If the transmitter were linear, its gain would be

$$K'_T = \frac{100}{f_{max}}$$

The expression for K_T shows that the gain is not constant but is rather a function of flow. The greater the flow, the greater the gain. Specifically,

at

$\left(\dfrac{f}{f_{max}} \right)$ 0   0.1   0.5   0.75   1.0

$\left(\dfrac{K_T}{K'_T} \right)$ **0**   **0.2**   1.0   1.50   2.0

Thus the actual gain varies from zero to twice what the gain would be if the transmitter were linear. This fact results in a nonlinearity in flow control systems. Most manufacturers offer differential pressure transmitters with built-in square root extractors yielding a linear transmitter. Also, most modem distributed control systems offer automatic square root extraction of signals. This makes the combination of sensor/transmitter/ square root linear with a gain of $100/f_{max}$. Chapter 11 discusses in more detail the use of square root extractors.

The dynamic parameters are generally obtained empirically using methods similar to

the ones shown in Chapters 6 and 9; some are provided by manufacturers. For an example in which the time constant can be estimated from basic principles, see Problem 3-4. Some analyzer sensor/transmitters, such as **chromatographs,** present a dead time because of their analysis time and sampling operation. Sampled-data systems are discussed in Chapters 14 and 15.

5-2   CONTROL   VALVES

Control valves are the most common final control elements. They perform the action (A) function of the control system by adjusting the flows that affect the controlled variables. This section presents the most important aspects of control valves: the selection of their action and fail position, their capacity and sizing, their flow characteristics, their gain, and their transfer function. Appendix C presents different types of valves and their accessories. The reader is strongly encouraged to read Appendix C along with this section.

A control valve acts as a *variable restriction* in a process pipe. By changing its opening, it changes the resistance to flow and thus the flow itself. Throttling flows is what control valves are all about. The controller output signal positions the valve, determining the valve position that in turn determines the degree of restriction to flow. Therefore, the controller output signal is the *input* to the valve, and the flow is the *output* of the valve.

5-2.1   The Control Valve Actuator

Figure 5-2.1a shows the instrumentation schematic of a control valve. Even with electronic instrumentation, an air pressure actuator is the most common means of adjusting the position of control valves; this is because of the high reliability and low maintenance requirements of air, or pneumatic, actuators. When the signal from the controller is a 4- to 20-mA signal, a *current-to-pressure transducer,* labeled I/P in Fig. 5-2.1a, is
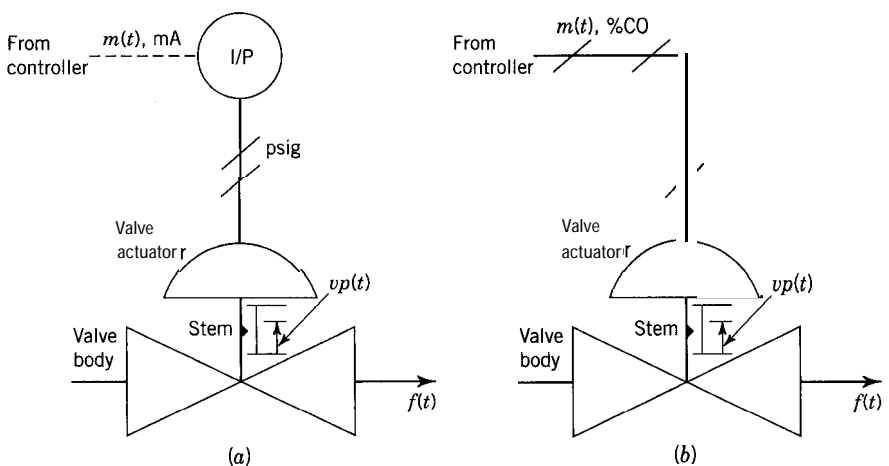
Figure 5-2.1 Instrumentation schematic for a control valve. (a) Detailed. (b) Conceptual simplification.

required to convert the current to an air pressure. The transducer, however, does not change the signal, and it can be omitted in a conceptual diagram such as Fig. 5-2.1*b*. In this diagram, the controller signal $m(t)$ is in *percent controller output* (*%CO*), as opposed to mA or psig.

The control *valve actuator* consists of a diaphragm and a spring, with the diagram attached to the *stem,* which positions the flow restriction in the *valve body.* Figure 52.1 shows these parts of the valve, and Appendix C includes pictures of several control valves with additional details of their parts.

The actuator, as shown in Fig. 5-2.1*b*, converts the controller output, $m(t)$, into the valve position, $vp(t)$. The valve position is usually expressed as a fraction that varies between zero and unity. When the valve position is zero, the valve is closed and the flow is zero. At the other extreme, the valve position is unity, the valve is fully opened, and the flow is maximum. For a full-range valve actuator, a 1% change in controller output results in a 0.01 change in fraction valve position. Most control valves use a full-range   actuator.

Control Valve Action

The first question the engineer must answer when specifying a control valve is "What do I want the valve to do when the energy supply fails?" This question concerns the "fail position" or "action" of the valve. The main consideration in answering this question is, or should be, *safety.* When the safest position of the valve is the closed position, the engineer must specify a fail-closed (FC) valve. Such a valve requires energy to open and is also called an air-to-open (AO) valve. The other possibility is a fail-open (FO) valve. Fail-open valves require energy to close and are called air-to-close (AC) valves.

To illustrate the selection of the action of control valves, let us consider the flash drum shown in Fig. 5-2.2. Steam is condensed in a coil to partially vaporize the liquid feed and separate its components into the vapor and liquid products. There are three
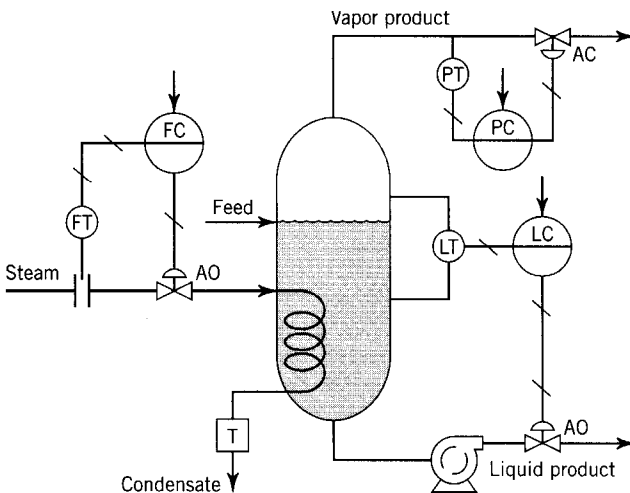
Figure 52.2 Fail positions of control valves on a flash drum.

valves in this example: one on the steam line to the coil and one on each of the liquid and vapor products. The valve on the liquid product controls the level in the tank, and the valve on the vapor product controls the pressure in the tank. The question is what we want each of these valves to do if the electrical power or air supply were to fail. As previously explained, each valve must move to its safest position when either the electric power or the air pressure fails. The safest position for the steam valve is closed, because this prevents a high steam flow that could vaporize all of the liquid and overheat the coil. Therefore, we select a fail-closed, or air-to-open, valve for the steam valve. For the liquid product valve, a fail-closed, or air-to-open, valve would keep the liquid stored in the tank. This action gives the operator time to shut down the feed to the tank and correct the cause of the failure. It is seldom safe for the liquid product to flow uncontrolled to the process downstream. Finally, a fail-open, or air-to-close, valve on the vapor product line would allow the vapor to flow out of the tank and prevent the tank from pressurizing.

In this example we have considered only the safety conditions around the flash drum, but doing so does not necessarily result in the safest operation of the process. The safety of the entire process requires that we also consider the effect of each flow on the downstream and upstream equipment. For example, when the vapor product valve fails opened, an unsafe condition may result in the process that receives the vapor. If this is so, the valve must fail closed. The engineer must then provide a separate pressure relief valve to route the vapors to an appropriate disposal system. The selection of the fail position of control valves is part of the procedure known as Hazard Analysis (HazAn). Such a procedure is performed by teams of engineers at process design time.

It is important to realize that safety is the only consideration in selecting the action of the control valve. As we shall see in the next section, the action of the control valve directly affects the action of the feedback controller.

The action of the valve determines the sign of the gain of the valve. An air-to-open valve has a positive gain, and an air-to-close valve has a negative gain. This is easy to see from the following formulas relating the valve position to the controller output.

$$\text{Air-to-open: } \overline{vp} = \frac{\overline{m}}{100}$$

$$\text{Air-to-close: } \overline{vp} = \frac{100 - \overline{m}}{100}$$

These formulas relate the steady-state values of the variables. They do not consider the dynamic response of the actuator.

5-2.2   Control Valve Capacity and Sizing

The purpose of the control valve is to regulate the manipulated flow in the control system. To regulate flow, the flow capacity of the control valve varies from zero when the valve is closed to a maximum when the valve is fully opened-that is, when the fraction valve position is one. This subsection looks at the formulas provided by valve manufacturers to help the process designer estimate the flow capacity of control valves

and size a valve for a given service. The next subsection presents the dependence of valve capacity on valve position.

Following a convention adopted by all control valve manufacturers, the flow capacity of a control valve is determined by its capacity factor or flow coefficient, C_v, introduced in 1944 by Masoneilan International, Inc. ("Masoneilan Handbook"). By definition, the C_v coefficient is "the flow in U.S. gallons per minute (gpm) of water that flows through a valve at a pressure drop of 1 psi (lb per square inch) across the valve." For example, a valve with a C_v coefficient of 25 can deliver 25 gpm of water when it has a 1-psi pressure drop. Valve catalogs list the C_v coefficients of valves by type and size. Figures C-10.1a through c contain samples of valve catalog entries.

Liquid Service

A control valve is simply an orifice with a variable area of flow. The C_v coefficient and the basic principles that regulate flow through an orifice provide the following formula for the liquid flow through the valve.

$$f = C_v \sqrt{\frac{\Delta p_v}{G_f}}$$

(5-2.1)

where

f = liquid flow, U.S. gpm

Δp_v = pressure drop across the valve, psi

G_f = specific gravity of liquid at flowing conditions

Simple conversion of units in Eq. 5-2.1 gives the mass flow through the valve in lb/h:

$$w = \left(f \frac{gal}{min}\right)\left(60 \frac{min}{h}\right)\left(8.33 G_f \frac{lb}{gal}\right) = 500 C_v \sqrt{G_f \Delta p_v}$$

(5-2.2)

where w is the mass flow in lb/h, and 8.33 lb/gal is the density of water.

There are several other considerations, such as corrections for very viscous fluids, flashing, and cavitation, in determining the flow through control valves for liquid service. These considerations are presented in Appendix C.

Compressible Flow

Different manufacturers have developed different formulas to model the flow of compressible fluids-gases, vapors, and steam-through their control valves. We present the compressible flow formulas of two manufacturers, Masoneilan and Fisher Controls, ("Fisher Catalog 10"), to show the differences in their equations and methods. These are hardly the only two manufacturers of control valves. Several manufacturers produce good valves, including the Crane Company, DeZurik, Foxboro, and Honeywell. We chose Masoneilan and Fisher Controls because their equations and methods are typical of the industry.

Although the equations for compressible flow look quite different from the equation for liquids, it is important to realize that they derive from the equation for liquids. They simply contain the units conversion factors and density corrections for temperature and pressure. It is important to realize that the C_v coefficient of a valve is the same whether the valve is used for liquid or gas service.

Masoneilan proposes the following set of equations. For gas or vapor flow in cubic feet per hour, at the standard conditions of 1 atm and 60°F,

$$f_s = 836 C_v C_f \frac{p_1}{\sqrt{GT}} (y - 0.148 y^3) \qquad (52.3)$$

For gas or vapor mass flow,

$$w = 2.8 C_v C_f p_1 \sqrt{G \frac{520}{T}} (y - 0.148 y^3) \qquad (5\text{-}2.4)$$

For steam flow,

$$w = 1.83 C_v C_f \frac{p_1}{(1 + 0.0007 T_{SH})} (y - 0.148 y^3) \qquad (5\text{-}2.5)$$

where

f_s = gas flow, scfh (scfh = ft³/h at standard conditions of 14.7 psia and 60°F)
G = gas specific gravity with respect to air, calculated by dividing the molecular weight of the gas by 29, the average molecular weight of air
T = temperature at the valve inlet, °R (= °F + 460)
C_f = critical flow factor. The numerical value for this factor ranges between 0.6' and 0.95. Figure C-lo.4 shows this factor for different valve types.
p_1 = pressure at the valve inlet, psia
w = gas flow, lb/h
T_{SH} = degrees of superheat, °F

The term y expresses the compressibility effects on the flow and is defined by

$$y = \frac{1.63}{C_f} \sqrt{\frac{\Delta p_v}{p_1}} \qquad (52.6)$$

where

$\Delta p,$ = $p_1 - p_2$, pressure drop across the valve, psi
p_2 = pressure at valve exit, psia

and y has a maximum value of 1.5.. At low ratios of the pressure drop to the inlet pressure, the gas flow is approximately incompressible and proportional to the square root of the pressure drop across the valve. The formulas reflect this fact, because at low values of y, the function $y - 0.148 y^3 \approx y$. As the ratio of the pressure drop to the inlet pressure increases, the flow through the valve becomes *choked,* because the velocity of the gas approaches the velocity of sound, which is the maximum it can reach. Under

this condition, known as *critical flow,* the flow becomes independent of the exit pressure and of the pressure drop across the valve. The formulas also reflect this fact because, as y approaches its maximum value of 1.5, the function y − 0.148y³ approaches 1.0. When this happens, the flow becomes proportional to the upstream pressure, p_1 .

The critical flow factor C_f is an empirical factor that accounts for the pressure profile in the valve when the flow becomes critical. Note that it cancels out at low ratios of the pressure drop to the inlet pressure when the term $0.148y^3$ becomes negligible. As shown in Fig. C-10.4, the C_f factor depends on the type of valve and even on the direction of flow. This is because the flow patterns in the valve affect the pressure profile and consequently the density of the gas.

Fisher Controls defines two new coefficients for the capacity of valves for compressible fluids; C_g and C_1 . The coefficient C_g determines the gas flow capacity of the valve, whereas the coefficient C_1 , defined as C_g/C_v, is functionally the same as the C_f factor in Masoneilan's formula. The coefficient C, depends on the type of valve, and its values usually range between 33 and 38. The Fisher formula is

$$f_s = C_g \sqrt{\frac{520}{GT}}\, p_1 \sin \left[\left(\frac{59.64}{C_1} \right) \sqrt{\frac{\Delta p_v}{p_1}} \right]_{rad} \qquad (52.7)$$

where all the symbols are the same as for the Masoneilan formulas, and the argument must be limited to $\pi/2$ radians, at which point the flow is critical. To use the argument of the sine function in degrees instead of radians, replace the constant 59.64 radians with its equivalent, $3417°$, in Eq. 5-2.7. When this is done, the argument of the sine wave must be limited to 90". Figures C-lO.lb and C-10.1c give values of C_g and C, for Fisher Controls valves.

The sine function of Eq. 5-2.7 is basically the same function of y that appears in Eqs. 5-2.3 through 5-2.5 expressed differently. Figure 5-2.3 shows that the two functions
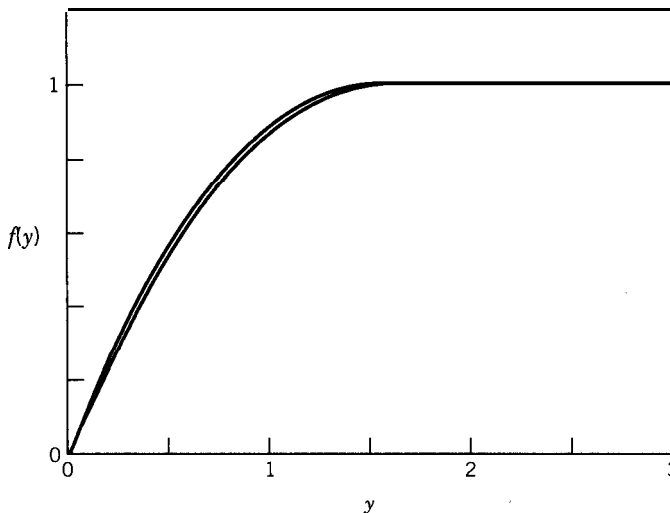
Figure 52.3 Compressible flow functions for Masoneilan and Fisher Controls formulas overlap each other. For Masoneilan, $f(y) = y −$ $0.148y^3$; for Fisher Controls, $f(y) = \sin(y)$, where y is in radians.

overlap. They are empirical functions that model the transition from subcritical flow to critical flow in a valve.

EXAMPLE 5-2.1

From Fig. C-lo-la, a 3-in. Masoneilan valve with full trim has a capacity factor of 110 $gpm/(psi)^{1/2}$ when fully opened. The pressure drop across the valve is 10 psi.

(a) Calculate the flow of a liquid solution with density 0.8 g/cm^3 (the density of water is 1 g/cm^3).
(b) Calculate the flow of gas with average molecular weight of 35 when the valve inlet conditions are 100 psig and 100°F.
(c) Calculate the flow of the gas from part (b) when the inlet pressure is 5 psig. Calculate the flow both in volumetric and in mass rate units, and compare the results for a 3-in. Fisher Controls valve.

SOLUTION

(a) For the liquid solution, using Eq. 5-2.1 yields

$$f = 110 \sqrt{\frac{10}{0.8}} = 389 \text{ gpm}$$

or, in mass units,

$$w = 500(110)\sqrt{(0.8)(10)} = 155,600 \text{ lb/h}$$

(b) For the gas, with $G = 35/29 = 1.207$, inlet pressure $p_1 = 100 + 14.7 = 114.7$ psia, and $T = 100 + 460 = 560°R$, assuming $C_f = 0.9$ and using Eqs. 5-2.6 and 5-2.3, we get

$$y = \frac{1.63}{0.9} \sqrt{\frac{10}{114.7}} = 0.535$$

$$f_s = 836(110)(0.9) \frac{114.7}{\sqrt{(1.207)(560)}} [0.535 - 0.148(0.535)^3]$$

$$= 187,000 \text{ scfh}$$

In mass rate units, using Eq. 5-2.4, we get

$$w = 2.8(110)(0.9)(114.7) \sqrt{1.207 \frac{520}{560}} (0.512)$$

$$= 17,240 \text{ lb/h}$$

(c) When the inlet pressure is $p_1 = 5 + 14.7 = 19.7$ psia, the value of y from Eq. 5-2.6 is 1.290, making the function of y equal to 0.972 (near critical flow), and the flow through the valve becomes

$$f_s = 836(110)(0.9) \frac{19.7}{\sqrt{(1.207)(560)}} (0.972)$$

$$= 61,000 \text{ scfh} \quad (= 5,620 \text{ lb/h})$$

From Fig. C-10.1b, a 3-in. Fisher Controls valve has slightly higher capacity than the corresponding Masoneilan valve: $C_v = 120$, $C_g = 4280$, and C, = 35.7. These values make the flows for this valve about 9% higher than those for the Masoneilan valve. For the liquid solution this is obvious, because the formulas are the same, so the answers are 424 gpm and 169,700 lb/h. For the gas, with inlet pressure of 114.7 psia, using Eq. 5-2.7 yields

$$f_s = 4280 \sqrt{\frac{520}{(1.207)(560)}} (114.7) \sin \left[\frac{59.64}{35.7} \sqrt{\frac{10}{114.7}} \right]$$

$$= 204,000 \text{ scfh}$$

or about 9% higher than for the comparable Masoneilan valve.

When the inlet pressure is 19.7 psia, the argument of the sine function in Eq. 5-2.7 is 1.19 radians and the sine function is 0.928 (near critical flow). With these values the flow through the valve is 68,700 scfh, 13% higher than for the Masoneilan valve. Because this difference is caused largely by the higher capacity of the Fisher valve, we conclude that the formulas for both manufacturers give very similar results.

Sizing of Control Valves

Part of the job of a control engineer is to size control valves for a given service. The formulas presented thus far in this section, although they are useful for estimating the flow through a control valve, were developed for sizing control valves. To size a control valve for liquid service, we must know the flow through the valve, the pressure drop across the valve, and the specific gravity of the liquid. For compressible flow, we also need the inlet pressure and temperature and the average molecular weight of the fluid. With this information, the engineer should use the appropriate formula provided by the specific valve manufacturer to calculate the C_v coefficient. The formulas will be very similar to those presented here (Eqs. 5-2.1 through 5-2.7). Once the C_v (or C,) coefficient is known, the engineer selects, from the manufacturer's catalog, a valve that is large enough for the service. Generally, the calculated C_v falls between two different valve sizes, in which case the larger of the two should be selected. Tables provided by valve manufacturers are very similar to those presented in Fig. C-lo.1 (Appendix C).

When sizing the valve for a brand new service, we obtain the flow from the process steady-state design conditions. This is the flow through the valve at the nominal production rate of the process. We will call this *flow* the *nominal* flow through the valve and denote it $\bar{f}$. The pressure drop across the valve at nominal flow is the one to use in

sizing the valve. For example, the formula to size a valve for liquid service is, from Eq. 5-2.1,

$$\overline{C}_v = \bar{f}\sqrt{\frac{G_f}{\Delta \bar{p}_v}}$$

where $\Delta \bar{p}_v$ is the pressure drop across the valve, in psi, when the flow is the nominal flow, $\bar{f}$, in gpm. Obviously, the valve coefficient must be greater than the one calculated from the foregoing equation. This is because, if the valve is to regulate the flow, it must be able to increase the flow beyond the nominal flow. We call the ratio of the valve coefficient when the valve is fully opened to the valve coefficient at nominal flow, $C_{v,max}/\overline{C}_v$, the *overcapacity factor* of the valve. Typical overcapacity factors are 1.5, for 50% overcapacity, and 2.0, for 100% overcapacity.

Sometimes the control engineer must also choose the pressure drop across the valve at nominal flow, a decision often made in cooperation with the process engineer. The pressure drop across the valve represents an energy loss to the process and should be kept as low as possible, but seldom less than 5 psi. Higher pressure drops are required when the pressure drop in the line and equipment in series with the valve is high, as we shall see when we discuss valve characteristics.

EXAMPLE 5-2.2

A control valve is to regulate the flow of steam into a distillation column reboiler with a design heat transfer rate of 15 million Btu/h. The supply steam is saturated at 20 psig. Size the control valve for a pressure drop of 5 psi and 100% overcapacity.

SOLUTION

From the steam tables, we find that the steam latent heat of condensation is 930 Btu/lb. This means that the nominal flow of steam is 15,000,000/930 = 16,130 lb/h. The valve inlet pressure is 20 + 14.7 = 34.7 psia, and the degrees of superheat is zero (saturated). Assuming a Masoneilan valve with $C_f = 0.8$, Eqs. 5-2.6 and 5-2.5 yield

$$y = \frac{1.63}{0.8}\sqrt{\frac{5}{34.7}} = 0.773$$

$$y - 0.148y^3 = 0.705$$

$$\overline{C}_v = \frac{16,130}{(1.83)(0.8)(34.7)(0.705)} = 450\ \frac{\text{gpm}}{\sqrt{\text{psi}}}$$

For 100% overcapacity, the valve coefficient when fully opened is

$$C_{v,max} = 2.0\,\overline{C}_v = 900\ \frac{\text{gpm}}{\sqrt{\text{psi}}}$$

From Fig. C-10.1a, a 10-in. Masoneilan valve, with a coefficient of 1000, is the smallest valve with enough capacity for this service.

For comparison, we now use the Fisher Controls formula, Eq. 5-2.7, to size this valve. We first find the steam upstream temperature, 250°F, in the steam tables, for saturated steam at 34.7 psia. The molecular weight of steam is 18, and its specific gravity is $G = 18/29 = 0.621$. Next we figure out the nominal volumetric flow at standard conditions as $(16,130)(380)/18 = 341,000$ scfh, where 380 is the volume in scf/lbmole. Assuming a C, of 35, the valve coefficient is

$$\sin\left[\frac{59.64}{35}\sqrt{\frac{5}{34.7}}\right] = \sin(0.647) = 0.603$$

$$\overline{C}_g = \frac{341,000}{\sqrt{\frac{520}{(0.621)(710)}}\,(34.7)(0.603)} = 15,000$$

For 100% overcapacity, the maximum $C_g = 30,000$, which corresponds to a maximum $C_v = C_g/C_1 = 30,000/35 = 856$ gpm/(psi)$^{1/2}$, or about the same as for the Masoneilan valve.

EXAMPLE 5-2.3

Figure 5-2.4 shows a process for transferring an oil from a storage tank to a separation tower. The tank is at atmospheric pressure, and the tower works at 25.9 in. Hg absolute (12.7 psia). Nominal oil flow is 700 gpm, its specific gravity is 0.94, and its vapor pressure at the flowing temperature of 90°F is 13.85 psia. The pipe is S-in. Schedule 40 commercial steel pipe, and the efficiency of the pump is 75%. Size a valve to control the flow of oil. From fluid flow correlations, the frictional pressure drop in the line is found to be 6 psi.
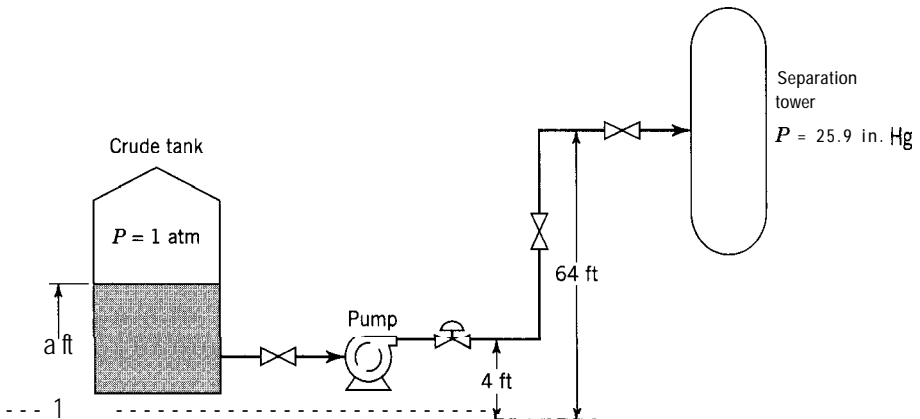
Figure 52.4 Process schematic for Example 5-2.3.

SOLUTION

Before we can size this valve, we must decide where to place it in the line and the pressure drop across the valve at nominal flow. The placement of the valve is important here; there is a possibility that the liquid will flash as its pressure drops through the valve. This would require a larger valve, because the density of the flashing mixture of liquid and vapor will be much less than that of the liquid. Note that if we place the valve at the entrance to the tower, the liquid will flash because the exit pressure, 12.7 psia, is less than the vapor pressure at the flowing temperature, 13.85 psia. A better location for the valve is at the discharge of the pump, where the exit pressure is higher as a result of the hydrostatic pressure of the 60 ft of elevation plus most of the 6 psi of friction drop. The hydrostatic pressure is $(62.3 \text{ lb/ft}^3)(0.94)(60 \text{ ft})/(144 \text{ in}^2/\text{ft}^2) = 24.4$ psi. This means that the pressure at the valve exit will be at least 37.1 psia $(24.4 + 12.7)$, well above the vapor pressure of the oil. There will be no flashing through the valve. The valve should never be placed at the suction of the pump, because there the pressure is lower and flashing would cause cavitation of the pump.

For the pressure drop across the valve, we will use 5 psi, or about the same as the friction drop in the line. To get an idea of the cost of this pressure drop, for an electricity cost of $0.03/kW-h, and 8200 h/year of operation of the pump, the annual cost attributable to the 5-psi drop across the valve is

$$\frac{700 \text{ gal}}{\text{min}} \frac{1 \text{ ft}^3}{7.48 \text{ gal}} \frac{(5)(144) \text{ lbf}}{0.75 \text{ ft}^2} \frac{1 \text{ kW-min}}{44,250 \text{ ft-lbf}} \left(\frac{8,200 \text{ h}}{\text{yr}}\right) \frac{\$0.03}{\text{kW-h}} = \$500/\text{yr}$$

where 0.75 is the efficiency of the pump. This cost may appear insignificant until one considers that a typical process may require several hundred control valves.

The maximum valve coefficient (fully opened) for 100% overcapacity is

$$C_{v,\text{max}} = 2(700)\sqrt{\frac{0.94}{5}} = 607 \frac{\text{gpm}}{\sqrt{\text{psi}}}$$

This requires an 8-in. Masoneilan valve (Fig. C-10.1*a*), which has a C_v of 640. As a comparison, a pressure drop across the valve of 2 psi requires a C_v of 960, corresponding to an 10-in. valve. The annual cost due to a pressure drop of 2 psi is $200 a year. A valve pressure drop of 10 psi requires a C_v of 429, corresponding to an 8-in. valve, and represents an annual cost of $1000 a year.

5-2.3 Control Valve Characteristics

The C_v coefficient of a control valve depends on the valve position. It varies from zero when the valve is closed, $vp = 0$, to a maximum value, $C_{v,\text{max}}$, when the valve is fully opened, that is, when the fraction valve position is unity. It is this variation in the C_v that allows the valve to regulate the flow continuously. The particular function relating the C_v coefficient to the valve position is known as the *inherent valve characteristics.* Valve manufacturers can shape the valve characteristics by arranging the way the area of the valve orifice varies with valve position.

Figure 5-2.5 shows three common valve characteristics: the *quick-opening, linear,*
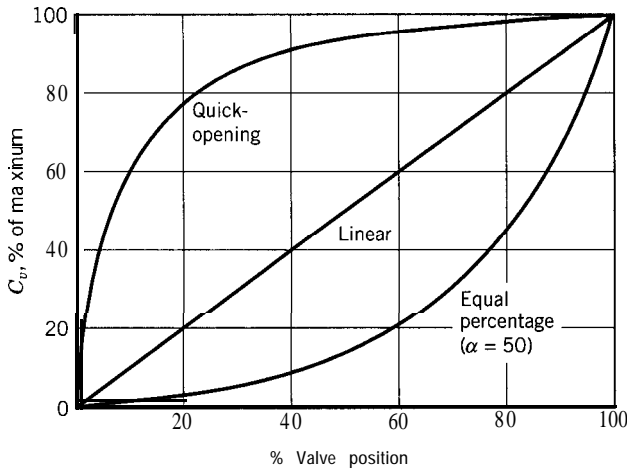
Figure **5-2.5** Inherent valve characteristic curves.

and *equal percentage* characteristics. As is evident from its shape, the quick-opening characteristic is not suitable for regulating flow, because most of the variation in the valve coefficient takes place in the lower third of the valve travel. Very little variation in coefficient takes place for most of the valve travel. Quick-opening valves are appropriate for relief valves and for on-off control systems. Relief valves must allow a large flow as quickly as possible to prevent over-pressuring of process vessels and other equipment. On-off control systems work by providing either full flow or no flow. They do not regulate the flow between the two extremes.

The two characteristics normally used to regulate flow are the linear and equal percentage characteristics. The function for linear characteristics is

$$C_v(vp) = C_{v,\max}vp \qquad (52.8)$$

and that for equal percentage characteristics is

$$C_v(vp) = C_{v,\max}\alpha^{vp-1} \qquad (52.9)$$

where α is the **rangeability parameter,** which has a value of 25, 50, or 100, with 50 being the most common. The actual equal percentage characteristic does not fit Eq. 5-2.9 all the way down to the closed position, because the exponential function cannot predict zero flow at zero valve position. In fact, it predicts a coefficient of $C_{v,\max}/\alpha$ at **vp** = 0. Because of this, the actual characteristic curve deviates from the exponential function in the lower 5% of the travel.

The linear characteristic produces a coefficient proportional to the valve position. At 50% valve position, the flow through the valve is 50% of its maximum flow.

The exponential function has the property that equal increments in valve position result in equal relative or percentage increments in the valve coefficient-hence the name. That is, when the valve position increases by 1% in going from 20% to 21% valve position, the flow increases by the same fraction of its value as when the valve position increases by 1% in going from 60% to 61% position, but the flow has a higher

value at the 60% position than at the 20% position. What makes such a function useful for regulating flow? To achieve uniform control performance, the control loop should have a constant gain. A linear valve characteristic may appear to be the only one that provides a constant gain. However, as we saw in Chapters 3 and 4, most processes are nonlinear in nature, and many exhibit a decrease in gain with increasing load. For such processes, the equal percentage characteristic, having a gain that increases as the valve opens (see Fig. 5-2.5), compensates for the decreasing process gain. As far as the controller is concerned, it is the product of the gains of the valve, the process, and the sensor/transmitter, that must remain constant.

Selecting the correct valve characteristics for a process requires a detailed analysis of the characteristics or "personality" of the process. However, several rules of thumb, based on previous experience, help us in making the decision. Briefly, we can say that valves with the linear flow characteristic are used when the process is linear and the pressure drop across the valve does not vary with flow. Equal percentage valves are probably the most common. They are generally used when the pressure drop across the valve varies with flow and with processes in which the gain decreases when the flow through the valve increases.

Valve Rangeability

Closely associated with the valve characteristics is the valve rangeability, or turn-down ratio. The valve *rangeability* is the ratio of the maximum controllable flow to the minimum controllable flow. It is therefore a measure of the width of operating flows the valve can control. Because the flow must be under control, these flows cannot be determined when the valve is against one of its travel limits. A common way to define the maximum and minimum flows is at the 95% and 5% valve positions-that is,

$$\text{Rangeability} \quad = \quad \frac{\text{Flow at 95\% valve position}}{\text{Flow at 5\% valve position}} \tag{52.10}$$

Another definition uses the 90% and 10% valve positions.

If the pressure drop across the valve is independent of flow, the flow through the valve is proportional to its C_v coefficient. Then we can calculate the valve rangeability from its inherent characteristics. From Eq. 5-2.8, the linear characteristic produces a rangeability of $0.95/0.05 = 19$, and from Eq. 5-2.9, the equal percentage characteristic has an inherent rangeability of $\alpha^{-0.05}/\alpha^{-0.95} = \alpha^{0.90}$, which is 18 for $\alpha = 25$, 34 for $\alpha = 50$, and 63 for $\alpha = 100$. From Fig. 5-2.5 we can see that the rangeability of a quick-opening valve is about 3. This low rangeability is one reason why quick-opening valves are not suitable for regulating flow.

Installed Valve Characteristics

When the pressure drop in the line and equipment in series with a valve is significant compared with the pressure drop across the valve, this pressure drop across the valve varies with the flow through the valve. This variation in pressure drop causes the variation of the flow with valve position to be different from the variation of the C_v coefficient. In other words, the *installed flow characteristics* of the valve are different from
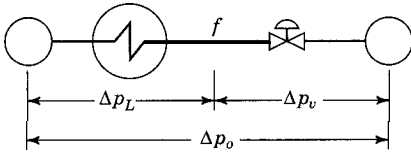
Figure 52.6 Valve in series with a heat
exchanger. The pressure drop across the
valve varies with the flow.

the inherent C_v characteristics. To develop a model for the installed flow characteristics,
consider the piping system shown in Fig. 5-2.6. Although in this system the valve is in
series with a heat exchanger, any flow resistance in series with the valve will cause the
phenomenon we are about to describe and model.

There are two basic assumptions to our model: (1) The pressure drop in the line and
equipment in series with the valve, Δp_L, varies with the square of the flow. (2) There
is a total pressure drop, Δp_0, that is independent of flow. This total pressure drop
provides the total pressure differential available across the valve plus the line and equip-
ment. The first of these assumptions is approximately valid when the flow is turbulent,
which is the most common flow regime in industrial equipment. We can always find
the total pressure drop, Δp_0, by finding the pressure drop across the valve when it is
closed, because then the flow, and consequently the frictional pressure drop in the line
and equipment, are zero.

Let

$$\Delta p_L = k_L G_f f^2 \tag{5-2.11}$$

where

Δp_L = frictional pressure drop across the line, fittings, and equipment in series
    with the control valve, psi
f = flow through the valve and line, gpm
k_L = constant friction coefficient for the line, fittings, and equipment, psi/(gpm)^2
G_f = specific gravity of the liquid (that of water is 1)

The pressure drop across the valve is obtained from Eq. 5-2.1.

$$\Delta p_v = G_f \frac{f^2}{C_v^2} \tag{5-2.12}$$

The total pressure drop is the sum of the two.

$$\Delta p_0 = \Delta p_v + \Delta p_L$$

$$= \left(\frac{1}{C_v^2} + k_L \right) G_f f^2 \tag{5-2.13}$$

Solving for the flow yields

$$f = \frac{C_v}{\sqrt{1 + k_L C_v^2}} \sqrt{\frac{\Delta p_0}{G_f}}$$

(52.14)

This formula constitutes the model of the installed characteristics for any valve in liquid service. Note that if the line pressure drop is negligible, then $k_L = 0$, $\Delta p_0 = \mathbf{Ap,,}$ and Eq. 5-2.14 becomes the same as Eq. 5-2.1. In this case, the installed characteristics are the same as the inherent characteristics because the pressure drop across the valve is constant. The friction coefficient is calculated from the line pressure drop at nominal flow. From Eq. 5-2.11,

$$k_L = \frac{\Delta \bar{p}_L}{G_f \bar{f}^2}$$

(52.15)

To obtain the installed characteristics as a fraction of maximum flow, we first obtain the maximum flow through the valve by substituting the maximum C_v in Eq. 5-2.14.

$$f_{max} = \frac{C_{v,max}}{\sqrt{1 + k_L C_{v,max}^2}} \sqrt{\frac{\Delta p_0}{G_f}}$$

(52.16)

Then we divide Eq. 5-2.14 by Eq. 5-2.16 to obtain

$$\frac{f}{f_{max}} = \frac{C_v}{C_{v,max}} \sqrt{\frac{1 + k_L C_{v,max}^2}{1 + k_L C_v^2}}$$

(52.17)

Note that the maximum flow through the valve, f_{max}, is independent of the valve characteristics, whereas the normalized installed characteristics (Eq. 5-2.17) are independent of the total pressure drop, $\mathbf{Ap,.}$ In fact, for a valve with a given capacity, the normalized flow characteristics, and consequently its rangeability, depend only on the friction coefficient of the line, k_L, and on the inherent characteristics of the valve.

The model that results in Eqs. 5-2.14, 5-2.16, and 5-2.17 applies only to liquid flow through the valve without flashing. We could develop a similar model for gas flow through the valve. However, such a model must differentiate between the pressure drop in the line upstream of the valve and the pressure drop downstream of the valve. It must also consider whether the flow through the valve is critical or subcritical. Therefore, such a model would be represented not by simple formulas but by a computer program or spreadsheet. The use of the formulas for installed characteristics in liquid flow is illustrated in the following example.

EXAMPLE 5-2.4

For the valve of Example 5-2.3, find the maximum flow through the valve, the installed flow characteristics, and the rangeability of the valve. Assume both linear and equal

percentage characteristics with rangeability parameter of 50. Analyze the effect of varying the pressure drop across the valve at nominal flow.

SOLUTION

Although the pressure rise through the pump of Fig. 5-2.4 is also variable, we will assume that it is constant for simplicity. Alternatively, we could handle the variation of the pressure rise by adding the difference between the pressure rise at zero flow and the pressure rise at nominal flow to the 6-psi pressure drop in the line at nominal flow. In Example 5-2.3, we figured out that for a 5-psi drop across the valve, specific gravity of 0.94, and nominal flow of 700 gpm, the required valve coefficient for 100% over-capacity is 607 gpm/(psi)$^{1/2}$. However, the smallest valve with this capacity is an 8-in. valve with $C_{v,max}$ = 640 gpm/(psi)$^{1/2}$; we will use this value. The line friction coefficient is

$$k_L = \frac{6 \text{ psi}}{(0.94)(700 \text{ gpm})^2} = 13.0 \times 10^{-6} \frac{\text{psi}}{(\text{gpm})^2}$$

and the total (constant) flow-dependent pressure drop is

$$\Delta p_0 = \Delta p_v + \Delta p_L = 5 + 6 = 11 \text{ psi}$$

The maximum flow is, from Eq. (5-2.16),

$$f_{max} = \frac{640}{\sqrt{1 + (13.0 \times 10^{-6})(640)^2}} \sqrt{\frac{11}{0.94}} = 870 \text{ gpm}$$

Had we used the calculated $C_{v,max}$ of 607, we would have gotten a maximum flow of 862 gpm. Either way, the maximum flow is much less than twice the nominal flow, 1400 gpm, although the valve was sized for 100% overcapacity. This is because the line resistance limits the flow as the valve opens. It is not possible to select a valve big enough to deliver twice the nominal flow, because even if the entire 11 psi were across the line, the flow would be (1 1/0.94 x 13.0 X 10^{-6})$^{1/2}$ = 947 gpm.

To obtain the valve rangeability, calculate the flow at 95% valve position and at 5% valve position using Eq. 5-2.14. For linear characteristics, at vp = 0.05, C_v = $C_{v,max}vp$ = (640)(0.05) = 32, and from Eq. (5-2.14),

$$f_{0.05} = \frac{32}{\sqrt{1 + (13.0 \times 10^{-6})(32)^2}} \sqrt{\frac{11}{0.94}} = 109 \text{ gpm}$$

Similarly, at vp = 0.95, C_v = 608, and $f_{0.95}$ = 862 gpm. The rangeability is then (862)/(109) = 7.9, which is much less than the nominal rangeability of 19. For the equal percentage characteristics with rangeability parameter α = 50, the flows are

At vp = 0.05, C_v = $C_{v,max}\alpha^{vp-1}$ = (640)(50)$^{0.05-1}$ = 15.6 f= 53.2 gpm

Table 52.1 Results for Example 5-2.4

	Valve Pressure Drop, psi		
	2	5	10
Total pressure drop, psi	8	11	16
Calculated $C_{v,max}$	960	607	429
Required valve size[a]	1 0-in.	8-in.	8-in.
Actual $C_{v,max}$[a]	1000	640	640
Maximum flow, gpm	779	870	1049
Linear rangeability	5.4	7.9	7.9
Equal % rangeability	10.8	15.8	15.8

[a] From Fig. C-10.1a, Appendix C.

Similarly, at $vp = 0.95$, $C_v = 526, f = 839$ gpm, and the installed rangeability is $839/53.2 = 15.8$, also much lower than the inherent rangeability of 34.8, but about twice the rangeability of the linear valve for the same service.

Table 5-2.1 summarizes the results for pressure drops across the valve of 2 psi, 5 psi, and 10 psi. In each case, it is assumed that the line pressure drop does not change but that the total available pressure drop is the sum of the valve and line pressure drops at design flow. This is realistic, because at process design time, the valve pressure drop is decided on and then the pump is sized to provide the necessary total pressure drop. The table shows that the maximum flow increases with pressure drop even when the valve size decreases. Note also that the rangeability of the valve does not change when the valve size and the line pressure drop remain the same, even when the total available pressure drop increases.

Fig. 5-2.7 shows plots of the normalized installed characteristics that correspond to the three pressure drops. These characteristics were computed using Eq. 5-2.17. For comparison, Fig. 5-2.7 also shows the inherent characteristics of each valve; these are the flow characteristics when there is no pressure drop in the line (constant pressure drop across the valve). The characteristics for valve pressure drops of 5 and 10 psi overlap because the valve size does not change. Note how the installed characteristics for the linear valve turn into quick-opening characteristics-more so for the larger valve. By contrast, the installed characteristics for the equal percentage valve remain more linear, although they too flatten out at the high flows.

5-2.4 Control Valve Gain and Transfer Function

The gain of the valve, like that of any other device, is the steady-state change in output divided by the change in input. The valve schematic diagram of Fig. 5-2.1 shows that the output of the valve is the flow and that its input is the controller output signal in percent controller output (%CO). The gain of the valve is therefore defined by

$$K_v = \frac{df}{dm'} \frac{\text{gpm}}{\text{%CO}}$$

(52.18)

The valve gain can also be defined in other units, such as (lb/h)/(%CO) and scfh/(%CO).

(a)

(b)
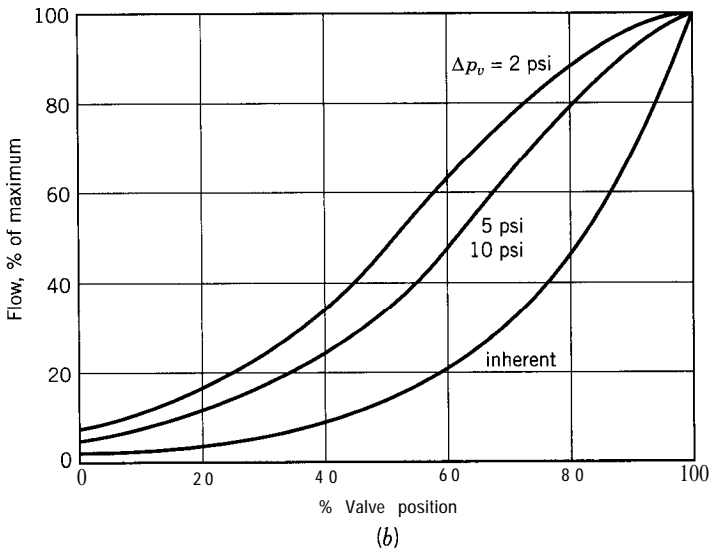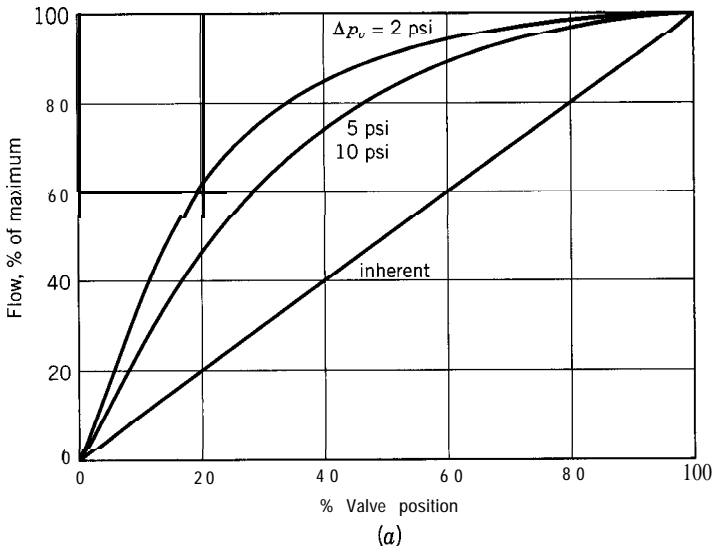
Figure 5-2.7 Installed flow characteristics of the control valve of Example
5-2.4. (a) Linear inherent characteristics. (b) Equal percentage character-
istics with $\alpha = 50$.

Using the chain rule of differentiation, we can show the valve gain as the product of
three terms relating the dependence of the valve position on the controller output, the
dependence of the C_v on the valve position, and the dependence of the flow on the C,.

$$K_v = \left(\frac{dvp}{dm}\right)\left(\frac{dC_v}{dvp}\right)\left(\frac{df}{dC_v}\right) \tag{52.19}$$

The dependence of the valve position is simply the conversion of percent controller output to fraction valve position, but the sign depends on whether the valve fails closed or opened.

$$\frac{dvp}{dm} = \pm \frac{1}{100}, \frac{\text{fraction } vp}{\%\text{CO}} \qquad (5\text{-}2.20)$$

where the plus sign is used if the valve fails closed (air-to-open), the minus sign if the valve fails opened (air-to-close).

The dependence of the C_v on the valve position depends on the valve characteristics. From Eqs. 5-2.8 and 5-2.9,

Linear: $$\frac{dC_v}{dvp} = C_{v,\max} \qquad (5\text{-}2.21)$$

Equal   percentage: $$\frac{d\overline{C}_v}{dvp} = (\ln \alpha) C_{v,\max} \alpha^{\overline{vp}-1} = (\ln \alpha)\overline{C}_v \qquad (5\text{-}2.22)$$

where the nonlinear exponential function has been linearized.

Finally, the dependence of flow on the C_v is a function of the installed characteristics of the control valve. We will consider first the simpler case of constant pressure drop across the valve and then the more complex case, which includes the pressure drop in the line in series with the valve.

Constant Valve Pressure Drop

When the pressure drop in the line in series with the valve is negligible, the inlet and outlet pressures, and thus the valve pressure drop, remain constant. For liquid service, from Eq. 5-2.1, the dependence of flow on the C_v coefficient is

$$\frac{df}{dC_v} = \sqrt{\frac{\Delta p_v}{G_f}} \qquad (5\text{-}2.23)$$

The gain of a valve with linear characteristics is now obtained by substituting Eqs. 5-2.20, 5-2.21, and 5-2.23 into Eq. 5-2.19.

$$\boxed{K_v = \pm \frac{1}{100} C_{v,\max} \sqrt{\frac{\Delta p_v}{G_f}} = \pm \frac{f_{\max}}{100}, \frac{\text{gpm}}{\%\text{CO}}} \qquad (52.24)$$

where $f_{\max}$ is the flow through the valve when it is fully opened. Note that the gain of the linear valve is constant when the pressure drop across the valve is constant. It can

be similarly shown (and is left as an exercise), that the gain for either liquid or gas flow in mass units is

$$K_v = \pm \frac{w_{max}}{100}, \frac{\text{lb/h}}{\%CO}$$

(52.25)

when the valve inlet and outlet pressures do not vary with flow.

The gain of a valve with equal percentage characteristics is similarly obtained by substituting Eqs. 5-2.20, 5-2.22, and 5-2.23 into Eq. 5-2.19.

$$K_v = \pm \frac{1}{100} (\ln \alpha) \overline{C}_v, \quad \frac{\&}{G_{f-}} = \pm \frac{\ln \alpha}{100} \overline{f}, \frac{\text{gpm}}{\%CO}$$

(52.26)

This formula shows that the gain of an equal percentage valve is proportional to the flow when the pressure drop across the valve is constant. The gain for either liquid or gas flow in mass units is

$$K_v = \pm \frac{\ln \alpha}{100} \overline{w}, \frac{\text{lb/h}}{\%CO}$$

(52.27)

when the valve inlet and outlet pressures do not vary with flow.

Variable Pressure Drop Across the Valve

To obtain the dependence of the flow on the C_v coefficient for liquid flow when the pressure drop across the valve is variable, differentiate Eq. 5-2.14 with respect to the C_v using the rules of differential calculus to obtain

$$\frac{df}{dC_v} = \frac{\sqrt{1 + k_L \overline{C}_v^2} - \overline{C}_v (1 + k_L \overline{C}_v^2)^{-1/2} k_L \overline{C}_v}{1 + k_L \overline{C}_v^2} \sqrt{\frac{\Delta p_0}{G_f}}$$

$$= (1 + k_L \overline{C}_v^2)^{-3/2} \sqrt{\frac{\Delta p_0}{G_f}}$$

(52.28)

For a valve with linear characteristics, the expression in Eq. 5-2.28 is multiplied by the constant $\pm C_{v,max}/100$ to obtain the valve gain. It is easy to see that the gain of the linear valve decreases as the valve opens, because of the increase in $\overline{C}_v$. For the equal per-

centage valve, the gain is obtained by substituting Eqs. 5-2.20, 5-2.22, and 5-2.28 into Eq. 5-2.19.

$$K_v = \pm \frac{\ln \alpha}{100} \frac{\overline{C}_v}{(1 + k_L \overline{C}_v^2)^{3/2}} \sqrt{\frac{\Delta p_0}{G_f}}$$

$$= \pm \frac{\ln \alpha}{100} \frac{\bar{f}}{1 + k_L \overline{C}_v^2} \qquad (5\text{-}2.29)$$

where we have substituted Eq. 5-2.14. Note that this gain is less variable with valve opening, because the flow term in the numerator tends to cancel some of the effect of the C_v term in the denominator, at least until the valve is close to fully opened. This near linearity of the installed characteristics of the equal percentage valve can also be observed in the plots of Fig. 5-2.7b.

EXAMPLE 5-2.5

Find the gain and the valve position at design conditions for the steam valve of Example 5-2.2. Assume that the 10-in. valve with $C_{v,max} = 1000$ is selected and that the pressures around the valve are independent of flow. Consider both a valve with linear characteristics and an equal percentage valve with rangeability parameter of 50. For the latter, find the gain at the nominal flow of 16,130 lb/h.

SOLUTION

This is a steam valve, so we will assume that it fails closed to prevent overheating of the reboiler. Then, as the controller signal opens the valve, the valve gain is positive.

For the linear valve, the valve position at design flow is found from Eq. 5-2.8.

$$\overline{vp} = \frac{\overline{C}_v}{C_{v,max}} = \frac{450}{1000} = 0.450$$

The gain is obtained from Eq. 5-2.25.

$$K_v = + \frac{w_{max}}{100} = \$16,130 \frac{1000}{450} = 358, \frac{lb/h}{\%CO}$$

where, because the pressures are constant, we have used the ratio of the C_v values to estimate the maximum flow.

For the equal percentage valve with $\alpha = 50$, the valve position is calculated by using Eq. 5-2.9.

$$\alpha^{\overline{vp}-1} = \frac{\overline{C}_v}{C_{v,max}} = \frac{450}{1000} = 0.450$$

$$vp = \frac{\ln(0.450)}{\ln(50)} + 1 = 0.80$$

The gain is obtained by using Eq. 5-2.27.

$$K_v = + \frac{\ln \alpha}{100} \bar{w} = \frac{\ln(50)}{100} \, 16{,}130 = 630 \, \frac{\text{lb/h}}{\%\text{CO}}$$

As expected, the gain of the equal percentage valve at design flow is greater than the constant gain of the linear valve.

EXAMPLE 5-2.6

Calculate the gain of the valve in Example 5-2.4 at nominal flow. Consider both a linear valve and an equal percentage valve with rangeability parameter $\alpha = 50$.

SOLUTION

From Examples 5-2.3 and 5-2.4, we know that $\Delta p_0 = 11$ psi, $\bar{f} = 700$ gpm, $K_L = 13.0 \times 10^{-6}$ psi/(gpm)^2, $C_v = 303$, and $C_{v,\text{max}} = 640$ gpm/(psi)$^{1/2}$. The valve feeds a distillation column, so let us assume that it fails closed. Thus its gain is positive because the controller signal opens it. The gain of the linear valve with variable pressure drop is obtained by substituting Eqs. 5-2.20, 5-2.21, and 5-2.28 into Eq. 5-2.19.

$$K_v = + \frac{1}{100} \frac{C_{v,\text{max}}}{(1 + k_L \bar{C}_v^2)^{3/2}} \sqrt{\frac{\Delta p_0}{G_f}}$$

$$= \frac{640}{(100)[1 + 13.0 \times 10^{-6}(303)^2]^{3/2}} \sqrt{\frac{11}{0.94}}$$

$$= 6.7, \frac{\text{gpm}}{\%\text{CO}}$$

This gain is less than half the gain of 15 gpm/%CO that the valve would have if the pressure drop of 5 psi across the valve remained constant.

The gain of the equal percentage valve is obtained from Eq. 5-2.29.

$$K_v = + \frac{\ln(50)}{100} \frac{700}{1 + 13\,0 \times 10^{-6}(303)^2} = 12.5 \, \frac{\text{gpm}}{\%\text{CO}}$$

This gain is about half the gain of 27 gpm/%CO that the valve would have if the pressure drop across the valve were independent of flow.

Valve Transfer Function

Figure 5-2.8 shows the block diagram for a control valve. It is usually sufficient to model the valve as a first-order lag, which results in the transfer function

$M(s), \%CO \longrightarrow \boxed{G_v(s)} \longrightarrow F(s), \text{gpm}$
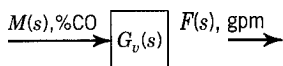
Figure 5-2.8 Block diagram
of a control valve.

$$G_v(s) = \frac{K_v}{\tau_v s + 1} \qquad (52.30)$$

where

K_v = valve gain, gpm/%CO or (lb/h)/%CO or scfh/%CO
τ_v = time constant of valve actuator, s

The actuator time constant is usually of the order of a few seconds and can be neglected when the process time constants are of the order of minutes.

The block diagram of Fig. 5-2.8 assumes that the pressure drop across the valve either is constant or is a function of flow only. When the pressure drop across the valve is a function of other process variables, as in the control of level or gas pressure, the block diagram must include the effect of these variables on the flow through the valve. Chapters 3 and 4 show examples of block diagrams in which level and pressure variables affect the flow through the valve.

5-2.5 Control Valve Summary

This section has presented some important considerations in the modeling and sizing of control valves. Although there are other considerations that must be taken into account when specifying a control valve, the formulas presented here allow the modeling of control valves for the purposes of designing and analyzing the complete control system. The reader who wants more details on the complete specification of control valves should see the references given at the end of this chapter.

5-3 FEEDBACK CONTROLLERS

This section presents the most important types of industrial controllers. Specifically, we will consider the different types of algorithms used in analog controllers and the most common ones used in distributed control systems (DCSs) and in "stand-alone controllers," which are also sometimes referred to as single-loop controllers or simply as loop controllers. As presented in Chapter 1, the DCSs and the "stand-alone" controllers are computer-based, so they process the signals not on a continuous basis but rather in a discrete fashion. However, the sampling time for these systems is rather fast, usually ranging from 10 times a second to about once a second. Thus for all practical purposes, these controllers appear to be continuous. Chapter 15 presents other details related to discrete controllers, such as how the algorithms are written for discrete operation and the effect of sampling time.

Briefly, the controller is the "brain" of the control loop. As we noted in Chapter 1, it is the device that performs the decision (D) operation in the control system. To do this, the controller

1. Compares the process signal it receives, the controlled variable, with the set point. The set point is the desired value of the process signal.
2. Sends an appropriate output signal to the control valve, or any other final control element, in order to maintain the controlled variable at its set point.

Figure 5-3.1 shows different types of controllers. Figs. **5-3.1a** and *b* show some stand-alone controllers. These controllers have a series of buttons/windows that make it possible to adjust the set point, read the value of the controlled variable, transfer between the automatic and manual modes, read the output signal from the controller, and adjust the output signal when in the manual mode. Most loop controllers have these options on the front panel for ease of operation. Fig. **5-3.1c** shows what is known as a *distributed control system (DCS)*.

The auto/manual button determines the operation of the controller. When this button is in the auto (automatic) position, the controller decides on the appropriate signal and outputs it to the final control element to maintain the controlled variable at the set point. In the manual position, the controller stops deciding and allows operating personnel to change the output manually. In this mode, the controller just provides a convenient (and expensive) way to adjust the final control element. In the auto mode, information from the manual adjustment is ignored, or disabled; only the set point influences the output. In the manual mode, on the other hand, the set point has no effect on the controller output; only the manual output influences the output. When a controller is set in manual, there is not much need for the controller. *Only when the controller is in automatic are the benefits* **of** *automatic process control obtained.*

5-3.1 Actions of Controllers

The selection of the controller action is critical. *If the action is not correctly selected, the controller will not control.* Let us see how to select the action and what it means.

Consider the heat exchanger control loop shown in Fig. 5-3.2; the process is at steady state, and the set point is constant. Assume that the signal from the temperature transmitter increases, indicating that the outlet temperature has increased above set point. To return this temperature to set point, the controller must close the steam valve by some amount. Because the valve is fail-closed (FC), the controller must reduce its output signal to the valve (see the arrows in the figure). When an increase in the process variable requires a decrease in controller output, the controller must be set to **reverse** *action.* Often the term *increase/decrease* (as the input signal to the controller increases, the output signal from the controller must decrease), or simply *decrease,* is also used.

Alternatively, consider the level control loop shown in Fig. 5-3.3; the process is at steady state, and the set point is constant. Assume that the signal from the level transmitter increases, indicating that the level has increased above the set point. To return this level to set point, the controller must open the valve by some amount. Because the valve is fail-closed (FC), the controller must increase its output signal to the valve (see the arrows in the figure). To make this decision, the controller must be set to *direct action.* Often the term *increase/increase* (as the input signal to the controller increases, the output signal from the controller must also increase), or simply *increase,* is also used.

In summary, to determine the action of a controller, the engineer must know

1. The process requirements for control.
2. The fail-safe action of the control valve or other final control element.

(a)
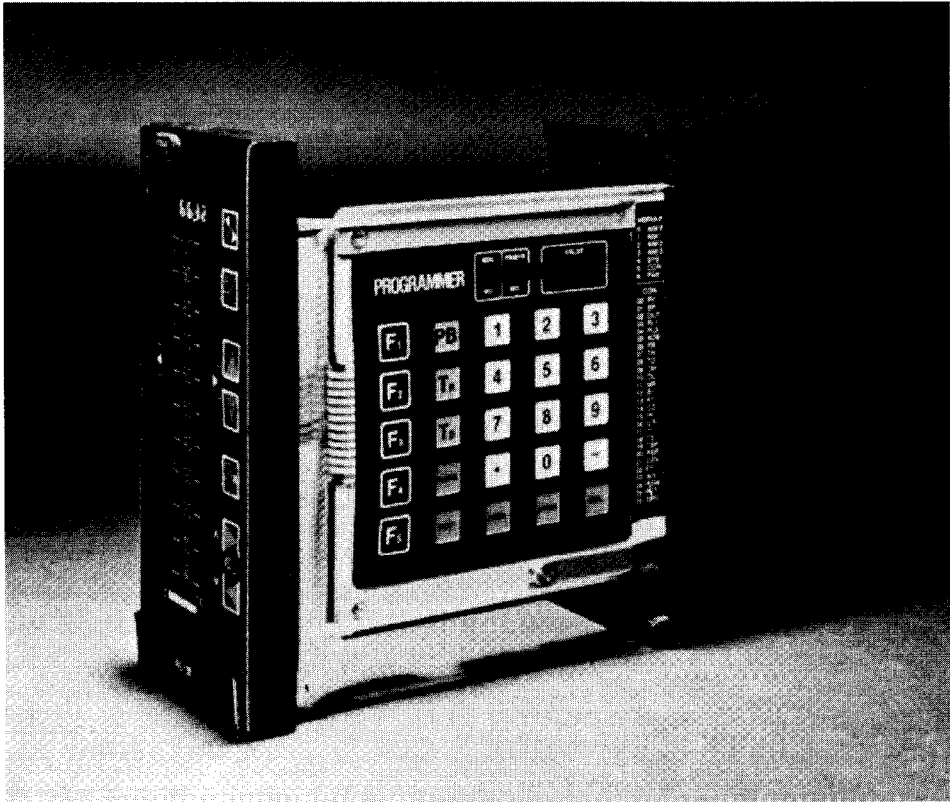
(b)

Figure 5-3.1 Controllers. (a) Loop controllers. (Courtesy of Johnson-Yokagawa.) (b) Loop con-
troller. (Courtesy of Fischer & Porter.)
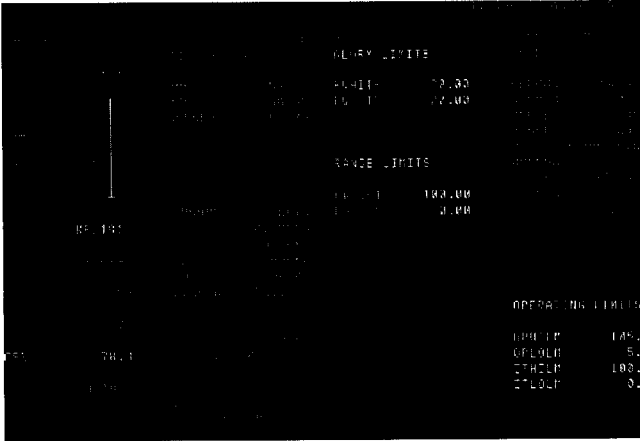
(c)

Figure 53.1 *(Continued)* (c) Distributed Control System. (Courtesy of Honeywell).

Both things must be taken into consideration. What should be the action of the level controller if a fail-open (FO) valve is used? And what should it be if the level is controlled with the inlet flow instead of the outlet flow? In the first case the control valve action changes, whereas in the second case the process requirements for control change.

The controller action is set by a switch or by a configuration bit on most controllers.

53.2   Types of Feedback Controllers

The way feedback controllers make a decision is by solving an equation based on the difference between the controlled variable and the set point. In this section, we examine
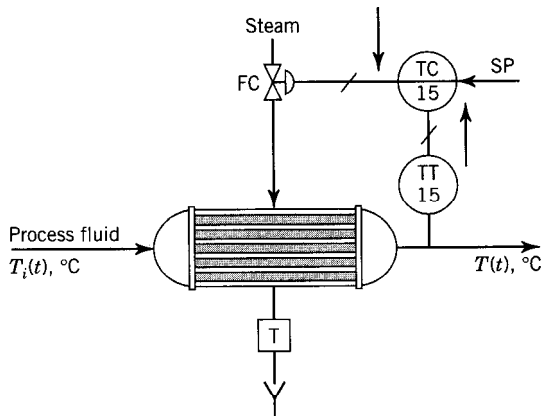
Figure 53.2 Heat exchanger control loop.
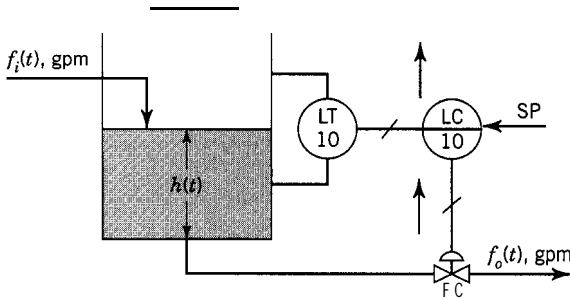
Figure 53.3 Liquid level control loop.

the most common types of controllers by looking at the equations that describe their operation.

As we saw in Chapter 1, the signals entering and exiting the controllers are either electrical or pneumatic. Even in computer systems, the signals entering from the field are electrical before they are converted, by an analog-to-digital (A/D) converter, to digital signals. Likewise, the signal the computer system sends back to the field is an electrical signal. To help simplify the presentation that follows, we will use all signals in percent. That is, we will speak of 0 to 100% rather than 4 to 20 mA, 3 to 15 psig, or any other type of signal.

As we have said, feedback controllers decide what to do to maintain the controlled variable at set point by solving an equation based on the difference between the set point and the controlled variable. This difference, or error, is computed as

$$e(t) = r(t) - c(t) \tag{5-3.1}$$

where

$c(t)$ = controlled variable. Most often, the controlled variable is given by the transmitter output (TO) and consequently has units of %TO.

$r(t)$ = set point. This is the desired value of the controlled variable and thus has units of %TO.

$e(t)$ = error in %TO.

The error could also have been computed as $e(t) = c(t) - r(t)$. However, Eq. 5-3.1 will be the convention used in this book.

Equation 5-3.1 is written in deviation variable form as

$$E(t) = R(t) - C(t) \tag{5-3.2}$$

where

$E(t)$ = the error in deviation form. Assuming that the error at the initial steady-state is zero, which is the convention used in this book, $E(t) = e(r)$ 0.

$R(t)$ = the set point in deviation variable form. It is defined as $R(t) = r(t) - \bar{r}$, where $\bar{r}$ is the initial steady-state value of the set point.

C(t) = the controlled variable in deviation form. It is defined as $C(t) = c(t) - \bar{c}$, where $\bar{c}$ is the initial steady-state value of the controlled variable.

Taking the Laplace transform of Eq. 5-3.2 yields

$$E(s) = R(s) - C(s) \tag{53.3}$$

The conventional block diagram representation for the controller is shown in Fig. 5-3.4. $M(s)$ is the Laplace variable used to denote the controller output, so it has units of percent controller output (%CO). G,(s) is the transfer function that describes how the controller acts upon an error. The following paragraphs present the different controllers along with their transfer functions.

Proportional Controller (P)

The proportional controller is the simplest type of controller we will discuss. The equation that describes its operation is

$$\boxed{m(t) = \bar{m} + K_c e(t)} \tag{53.4}$$

where

m(t) = controller output, %CO. The term m(t) is used to stress that as far as the controller is concerned, this output is the manipulated variable.

K_c = controller gain, $\dfrac{\%CO}{\%TO}$

$\bar{m}$ = bias value, %CO. This is the output from the controller when the error is zero. The value $\bar{m}$ is a constant and is also the output when the controller is switched to manual. It is very often initially set at mid-scale, 50 %CO.

Note that because the controlled variable is the signal from the transmitter with units of %TO, the set point must also have units of %TO. As the set point is entered in engineering units of the process variable, it is converted by the control system (controller) into %TO. This conversion is done using the transmitter range.

Equation 5-3.4 shows that the output of the controller is proportional to the error between the set point and the controlled variable. The proportionality is given by the controller gain, K_c. As a result of our definition of error, when K_c is positive, an increase in the controlled variable, c(t), results in a decrease in controller output, m(t). Thus a

R(s), %TO   E(s), %TO   M(s), %CO
                          $G_c(s)$
        +
        −
    CM, %TO
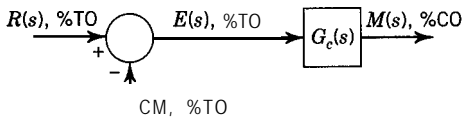
Figure 53.4 Block diagram representation of controller.

positive K_c results in a reverse-acting controller. To obtain a direct-acting controller, we must either use a negative K_c or reverse the definition of the error, that is, e(t) = c(t) — r(t). In this text, we will use the definition of the error as in Eq. 5-3.1 and use a negative K_c when a direct-acting controller is required. Most industrial feedback controllers, however, do not allow negative gains; in such cases, the error computation is reversed. This change in error computation is done internally by the controller. The user does not have to do anything but select the correct action. Note that whatever definition is used, the effect of the set point on the output is opposite to the effect of the controlled variable.

The controller gain determines how much the output from the controller changes for a given change in error; this is illustrated graphically in Fig. 5-3.5. The figure shows that the larger the K_c value, the more the controller output changes for a given error. Thus K_c establishes the sensitivity of the controller to an error, that is, how much the controller output changes per unit error.

Proportional'controllers offer the advantage of having only one adjustable, or tuning, parameter, K_c. However, they suffer a major disadvantage: the controlled variable is operated with an **offset.** Offset can be described as a **steady-state deviation of the controlled variable from set point,** or simply as a **steady-state error.** To examine the mean-ing of offset, consider the liquid level control loop shown in Fig. 5-3.3. The design operating conditions are $\bar{f}_i = \bar{f}_o$ = 150 gpm and h = 6 ft. Let us also assume that in order for the outlet valve to deliver 150 gpm, the signal to it must be 50 %CO. If the inlet flow, $f_i(t)$, increases, then the response of the system with a proportional controller looks like Fig. 5-3.6. The controller returns the controlled variable to a steady value,
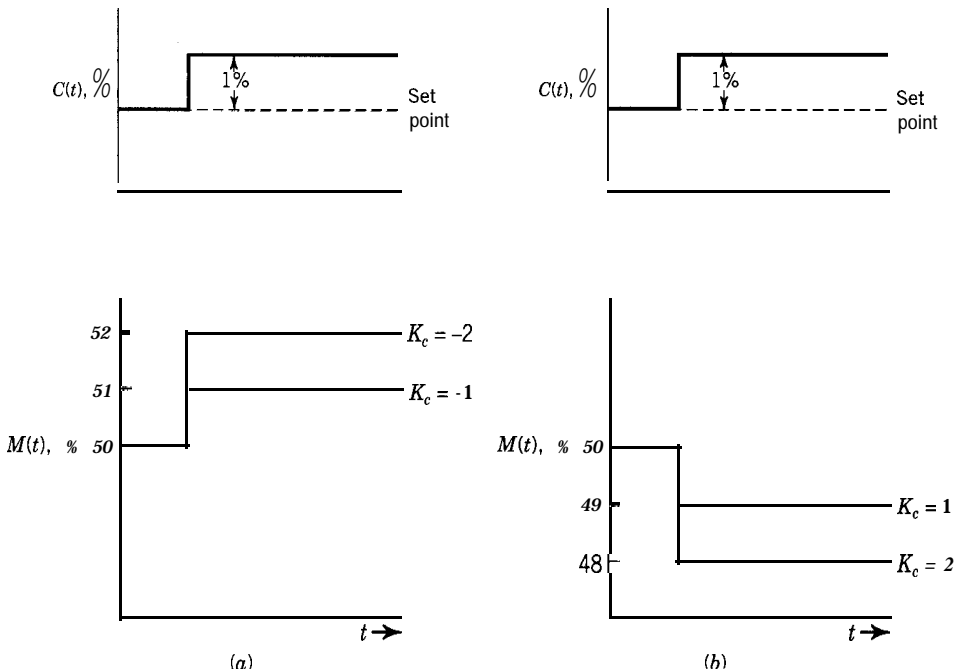
Figure **5-3.5** Effect of controller gain on output of controller. (a) Direct-acting controller. (b) Reverse-acting controller.
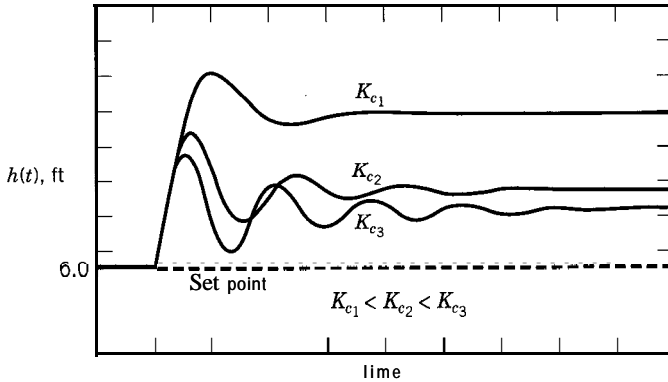
Figure 53.6 Response of liquid level process.

but not to the required set point. The difference between the set point and the new steady-state is the offset. The proportional controller is not "intelligent enough" to drive the controlled variable back to set point. The new steady-state value satisfies the controller.

Fig. 5-3.6 shows three response curves corresponding to three different values of K_c. This figure shows that the larger the value of K_c the smaller the offset. Why not, then, set a maximum gain to eliminate the offset? Fig. 5-3.6 also shows that although the larger K_c reduces the offset, the process becomes more oscillatory. For most processes, there is a maximum value of K_c beyond which the process goes unstable. Thus there is a limit to the value at which we can set K_c while at the same time maintaining stability. Consequently, the offset cannot be completely eliminated. The calculation of this maximum value of the controller gain, referred to as the ultimate gain, K_{cu}, is presented in Chapters 6 and 7.

Let us now look at a simple explanation why offset exists; a more rigorous proof is given in Chapter 6. Consider the liquid level control system shown in Fig. 5-3.3 with the same operating conditions previously given: $\bar{f}_i = \bar{f}_o = 150$ gpm and $\bar{h} = 6$ ft. Recall that the proportional controller, direct acting $(-K_c)$, solves the equation

$$m(t) = 50\% + (-K_c)e(t) \tag{5-3.5}$$

Assume now that the inlet flow increases to 170 gpm. When this happens, the liquid level increases and the controller in turn increases its output to open the valve and bring the level back down. In order to reach a steady operation, the outlet flow must now be 170 gpm. To deliver this new flow, the outlet valve must be open more than before, when it needed to deliver 150 gpm. This is a fail-closed valve, so let us assume that the new required signal to the valve to deliver 170 gpm is 60%. That is, the output from the controller must be 60%. Looking back at Eq. 5-3.5, we note that the only way for the controller output to be 60% is for the second term of the right-hand side to have a value of $+ 10\%$ and for this to be so, the error term cannot be zero at steady state. **This required steady-state error is the offset!** Note that a negative error means that the controlled variable is greater than the set point. The actual level in feet can be calculated from the calibration of the level transmitter for each controller gain.

Two points need to be stressed in this example. First, the magnitude of the offset depends on the value of the controller gain. Because the total term must have a value of $+10$ %CO,

K_c	$e(\infty)$, Offset, %TO
1	10
2	5.0
4	2.5

As previously mentioned, the larger the gain, the smaller the offset. The reader must remember that above a certain K_c, most processes go unstable. The controller equation does not show this, however; it will be discussed in Chapter 6.

Second, it seems that all a proportional controller is doing is reaching a steady-state operating condition. Once a steady state is reached, the controller is satisfied. The amount of deviation from set point, or offset, depends on the controller gain.

Many controller manufacturers do not use the term K_c for the tuning parameter; they use the term *proportional band* (PB). The relationship between gain and proportional band is given by

$$PB = \frac{100}{K_c} \qquad (5\text{-}3.6)$$

In these cases, the equation that describes the proportional controller is written as

$$m(t) = \overline{m} + \frac{100}{PB}\, e(t) \qquad (5\text{-}3.7)$$

PB is usually referred to as percent proportional band.

Equation 5-3.6 presents a most important fact. A large controller gain is the same as a low, or narrow, proportional band, and a low controller gain is the same as a large, or wide, proportional band. An increase in PB is similar to a decrease in K_c, resulting in a controller less sensitive to an error. A decrease in PB is similar to an increase in K_c, resulting in a more sensitive controller. K_c and PB are reciprocals, so care must be taken when tuning the controller.

Let us offer another definition of this term. *The proportional band is the error (expressed in percentage of the range of the controlled variable) required to move the output **of** the controller from its lowest to its highest value.* Consider the heat exchanger control loop shown in Fig. 5-3.2. The temperature transmitter has a range from 100°C to 300°C, and the set point of the controller is at 200°C. Figure 5-3.7 gives a graphical explanation of this definition of PB. The figure shows that a 100% PB means that as the controlled variable varies by 100% of its range, the controller output varies by 100% of its range. A 50% PB means that as the controlled variable varies by 50% of its range, the controller output varies by 100% of its range. Also note that a proportional only controller with a 200% PB will not move its output the entire range. A 200% PB means a very small controller gain or very little sensitivity to errors.
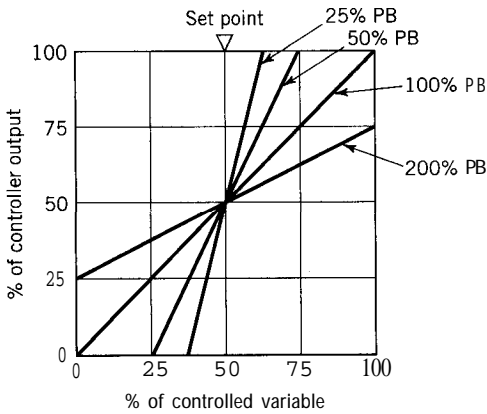
Figure 53.7 Definition of proportional band.

To obtain the transfer function for the proportional controller, we can write Eq. 53.1 as

$$m(t) - \overline{m} = K_c e(t)$$

or, in deviation variable form,

$$M(t) = K_c E(t)$$

where $M(t) = m(t)$ $\overline{m}$ and $E(t)$ is as previously defined. Taking the **Laplace** transform yields the transfer function

$$G_c(s) = \frac{M(s)}{E(s)} = K_c \qquad (5\text{-}3.8)$$

Equation 5-3.8 is the transfer function of a proportional controller and the one to apply in Fig. 5-3.4 when this controller is used.

To briefly summarize, proportional controllers are the simplest controllers and offer the advantage of having only one tuning parameter, K_c or **PB.** The disadvantage of these controllers is their operation with an offset in the controlled variable. In some processes, such as the level in a surge tank, the cruise control in a car, or a thermostat in a house, this offset may not be of any major consequence. In cases in which the process can be controlled within a band from set point, proportional controllers are sufficient. However, when the process variable must be controlled at the set point, not near it, proportional controllers do not provide the required control.

Proportional-Integral   Controller   (PI)

Most processes cannot be controlled with an offset; that is, they must be controlled at the set point. In these instances, an extra amount of intelligence must be added to the

proportional controller to remove the offset. This new intelligence, or new mode of control, is the integral, or reset, action; consequently, the controller becomes a proportional-integral (PI) controller. The describing equation is

$$m(t) = \bar{m} + K_c e(t) + \frac{K_c}{\tau_I} \int e(t)\, dt \qquad (5\text{-}3.9)$$

where τ_I = *integral (or reset) time.* Most often, the time unit used is minutes; less often, seconds are used. The unit used depends on the manufacturer. Therefore, the PI controller has two parameters, K_c and τ_I, both of which must be adjusted (tuned) to obtain satisfactory control.

To understand the physical significance of the reset time, τ_I, consider the hypothetical example shown in Fig. 5-3.8. At some time, $t = 0$, a constant error of 1% in magnitude is introduced in the controller. At this moment, the PI controller solves the following equation:

$$m(t) = 50\% + K_c(1) + \frac{K_c}{\tau_I} \int_0^t (1)\, dt$$

or

$$m(t) = 50\% + K_c + \frac{K_c}{\tau_I} t$$

When the error is introduced at $t = 0$, the controller output changes immediately by an amount equal to K_c; this is the response due to the proportional mode. As time increases, the output also increases in a ramp fashion, as expressed by the equation and shown in the figure. Note that when $t = \tau_I$ the controller's output becomes

$$m(t) = 50\% + K_c + K_c$$
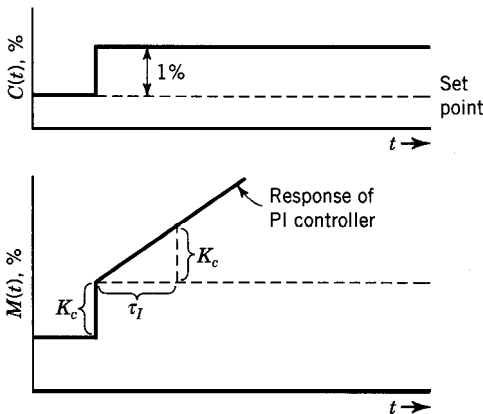
Figure 5-3.8 Response of PI controller (direct action) to a step change in error.
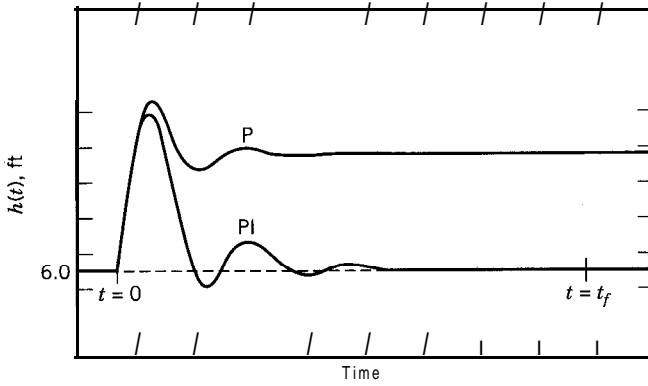
Figure 5-3.9 Response of liquid level process under P and PI controllers.

Thus, in an amount of time equal to τ_I, the integral mode repeats the immediate action taken by the proportional mode. The smaller the value of τ_I, the faster the controller integrates. Realize that the smaller the value of τ_I, the larger the term in front of the integral, K_c/τ_I, and consequently, the more weight given to the integral term.

To understand why the PI controller removes the offset, consider the level control system previously used to explain the offset required by a P controller. Figure 5-3.9 shows the response of the level under P and PI controllers to a change in inlet flow from 150 gpm to 170 gpm. The response with a P controller shows the offset, whereas the response with a PI controller shows that the level returns to set point, with no offset. Under PI control, as long as the error is present, the controller keeps changing its output (integrating the error). Once the error disappears (goes to zero), the controller does not change its output anymore (it integrates a function with a value of zero). As shown in the figure, at time t_f the error disappears. The signal to the valve must still be 60%, requiring the valve to deliver 170 gpm. Let us look at the PI equation at the moment the steady state is reached.

$$m(t) = 50\% + K_c(0) + \frac{K_c}{\tau_I} \int (0) \; dt$$

or

$$m(t) = 50\% + 0\% + 10\% = 60\%$$

The equation shows that even with a "zero" error, the integral term is not zero but rather 10%, which provides the required output of 60%. The fact that the error is zero does not mean that the value of the integral term is zero. It means that the integral term remains constant at the last value! Integration means area under the curve, and even though the level is the same at $t = 0$ and at $t = t_f$, the value of the integral is different (a different area under the curve) at these two times. The value of the integral term times K_c/τ_I is equal to 10%. Once the level returns to set point, the error disappears and the integral term remains constant. *Integration is the mode that removes the offset!*

This has been a brief explanation of why reset action removes the offset; Chapter 6 provides a more rigorous proof.

Some manufacturers do not use the term reset time, τ_I, for their tuning parameter. They use the reciprocal of reset time, which we shall refer to as *reset rate, τ_I^R*; that is,

$$\tau_I^R = \frac{1}{\tau_I} \qquad (5\text{-}3.10)$$

The unit of τ_I^R is therefore l/time, or simply (time)) [1]. Note that when τ_I is used and faster integration is desired, a smaller value must be used in the controller. However, when τ_I^R is used, a larger value must be used. Therefore, before tuning the reset term, the user must know whether the controller uses reset time (time) or reset rate (time-'). τ_I *and τ_I^R are reciprocals, so their effects are opposite.*

As we learned in the previous section, two terms are used for the proportional mode (K_c and *PB*), and we have just learned that there are also two terms for the integral mode (τ_I and τ_I^R). This can be confusing, so it is important to keep the differences in mind when tuning a controller. Equations 5-3.9, 5-3.11, 5-3.12, and 5-3.13 show four possible combinations of tuning parameters (Eq. 5-3.24 in Section 5-3.3 presents still another combination); we refer to Eq. 5-3.9 as the *classical controller.*

$$m(t) = \overline{m} + \frac{100}{PB} e(t) + \frac{100}{PB \cdot \tau_I} \int e(t)\ dt \qquad (5\text{-}3.11)$$

$$m(t) = \overline{m} + \frac{100}{PB} e(t) + \frac{100\tau_I^R}{PB} \int e(t)\ dt \qquad (5\text{-}3.12)$$

$$m(t) = \overline{m} + K_c e(t) + K_c \tau_I^R \int e(t)\ dt \qquad (5\text{-}3.13)$$

Using the same procedure we followed for the proportional controller, we obtain the transfer function for the PI controller from Eq. 5-3.9.

$$\boxed{G_c(s) = \frac{M(s)}{E(s)} = K_c \left(1 + \frac{1}{\tau_I s}\right) = K_c \left(\frac{\tau_I s + 1}{\tau_I s}\right)} \qquad (5\text{-}3.14)$$

To summarize, proportional-integral controllers have two tuning parameters: the gain or proportional band and the reset time or reset rate. *Their advantage is that the integration removes the offset.* Close to 85% of all controllers in use are of this type.

Proportional-Integral-Derivative   Controller   (PID)

Sometimes another mode of control is added to the PI controller. This new mode of control is the derivative action, which is also called the rate action, or preact. Its purpose is to anticipate where the process is heading by looking at the time rate of change of the error, its derivative. The describing equation is

$$m(t) = \bar{m} + K_c e(t) + \frac{K_c}{\tau_I} \int e(t) \, dt + K_c \tau_D \frac{de(t)}{dt} \qquad (5\text{-}3.15)$$

where τ_D = derivative (or rate) time. Most often the time unit is minutes, but some manufacturers use seconds.

The PID controller has three terms, K_c or **PB**, τ_I or τ_I^R, and τ_D, that must be adjusted (tuned) to obtain satisfactory control. The derivative action gives the controller the capability to anticipate where the process is heading-that is, to "look ahead"-by calculating the derivative of the error. The amount of "anticipation" is decided by the value of the tuning parameter, τ_D.

Let us consider the heat exchanger shown in Fig. 5-3.2 and use it to clarify what is meant by anticipation. Assume that the inlet process temperature decreases by some amount and the outlet temperature starts to decrease correspondingly, as shown in Fig. 5-3.10. At time t_a, the amount of the error is positive and small. Consequently, the amount of control correction provided by the proportional and integral modes is small. However, the derivative of this error, the slope of the error curve, is large and positive, making the control correction provided by the derivative mode huge. By looking at the derivative of the error, the controller knows that the controlled variable is heading away from set point rather fast, and it uses this fact to help in controlling. At time t_b, the error is still positive and is larger than before. The amount of control correction provided
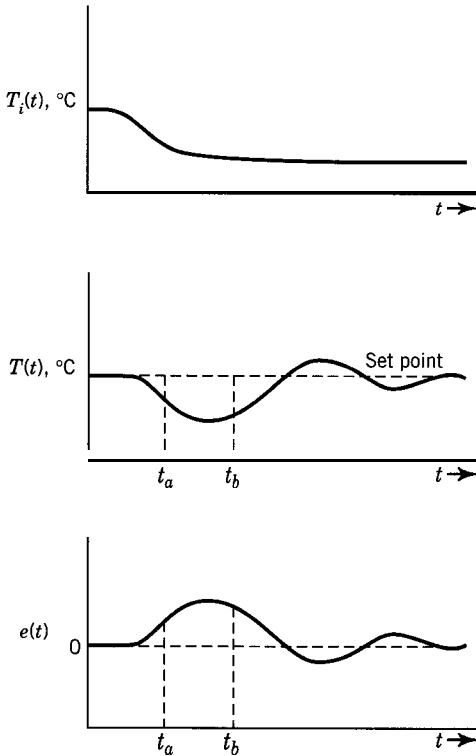
Figure 5-3.10 Heat exchanger control.

by the proportional and integral modes is also larger than before and is still adding to the output of the controller to open the steam valve further. However, the derivative of the error at this time is negative, signifying that the error is decreasing; the controlled variable has started to come back to set point. Using this fact, the derivative mode starts to subtract from the other two modes, because it recognizes that the error is decreasing. This algorithm results in reduced overshoot and decreases oscillations around set point.

PID controllers are recommended for use in slow processes (processes with multiple time constants or dead time) such as temperature loops, which are usually free of noise. Fast processes (processes with short time constants) are easily susceptible to process noise. Typical of these fast processes are flow loops and liquid pressure loops. Consider the recording of a flow shown in Fig. 5-3.11. The application of the derivative mode will only result in the amplification of the noise, because the derivative of the fast-changing noise is a large value. Processes with long time constants are usually damped and, consequently, are less susceptible to noise. In the case of a slow process with a noisy transmitter, however, the transmitter must be fixed or the noise filtered before the PID controller is used.

The transfer function of a PID controller is obtained by using the same procedure followed for a P and a PI controller:

$$G_c(s) = \frac{M(s)}{E(s)} = K_c \left(1 + \frac{1}{\tau_I s} + \tau_D s \right) \tag{5-3.16}$$

Actually, when the PID controller is implemented with Eq. 5-3.16, it does not work very well. To improve the performance of the derivative mode, the algorithm is slightly changed to

$$G_I(s) = \frac{M(s)}{E(s)} = K_c \left[1 + \frac{1}{\tau_I s} + \frac{\tau_D s}{\alpha \tau_D s + 1} \right] \tag{53.17}$$

The equation shows that the derivative portion is multiplied by the term $1/(\alpha \tau_D s + 1)$. This term, which can be recognized as the transfer function of a first-order system with gain of unity and a time constant equal to $\alpha \tau_D$, is referred as a filter. The filter does not
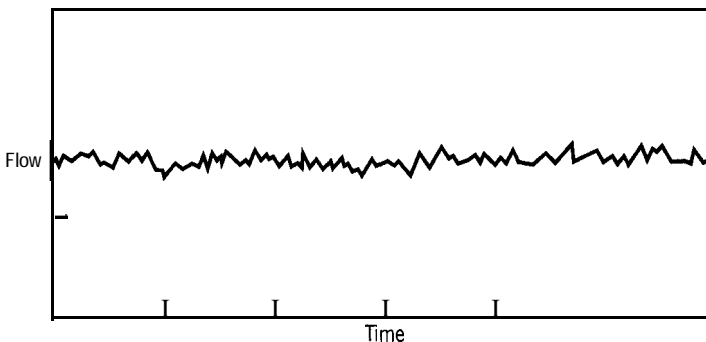
Figure 53.11 Recording of flow.

usually affect the performance of the controller because its time constant, $\alpha\tau_D$, is small. Typical values of α range between 0.05 and 0.2, depending on the manufacturer.

Equation 5-3.17 can be algebraically rearranged into

$$G_c(s) = \frac{M(s)}{E(s)} = K_c \left[\frac{(a + 1)\tau_D s + 1}{\alpha\tau_D s + 1} + \frac{1}{\tau_I s} \right] \tag{5-3.18}$$

The term $\dfrac{(\alpha + 1)\tau_D s + 1}{(\alpha\tau_D s + 1)}$ is a lead/lag unit which was introduced in Chapter 2 and is further discussed in Chapter 12. This transfer function shows the PID controller as a lead/lag unit in parallel with an integration. The net lead of the lead/lag unit is the derivative time.

In analog controllers and many computer-based controllers, the describing transfer function for the PID controllers used is

$$\boxed{G_c(s) = \frac{M(s)}{E(s)} = K_c' \left(1 + \frac{1}{\tau_I' s} \right) \left(\frac{\tau_D' s + 1}{\alpha\tau_D' s + 1} \right)} \tag{53.19}$$

Figure 5-3.12 shows the block diagram of Eq. 5-3.19. The diagram shows that this PID controller can be considered as a lead/lag unit in series with a PI controller, sometimes referred to as a "series PID" or "rate-before-reset."

In Eq. 5-3.19 the prime notation has been used to indicate that the tuning parameters are not the same as those in Eq. 5-3.16 or Eq. 5-3.17. Using algebraic manipulations with Eq. 5-3.16 and 5-3.19, the following relations can be obtained:

$$K_c' = K_c \left(0.5 + \sqrt{0.25 - \frac{\tau_D}{\tau_I}} \right)$$

$$\tau_I' = \tau_I \left(0.5 + \sqrt{0.25 - \frac{\tau_D}{\tau_I}} \right) \tag{5-3.20}$$

$$\tau_D' = \frac{\tau_D}{0.5 + \sqrt{0.25 - \frac{\tau_D}{\tau_I}}}$$

Chapter 6 shows how to obtain the tuning parameters K_c, τ_I, τ_D, K_c', τ_I', and τ_D'.

The controller described by Eq. 5-3.16 is sometimes referred to as an ideal PID, whereas the controller described by Eq. 5-3.19 is referred to as an actual PID.
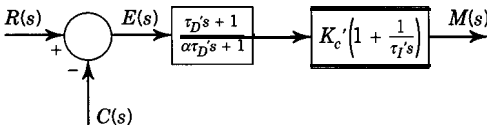
Figure 53.12 Block diagram of PID controller—
Eq. 5-3.19.

To summarize, PID controllers have three tuning parameters: the gain or proportional band, the reset time or reset rate, and the rate time. PID controllers are recommended for processes that are free of noise. The advantage of the derivative mode is that it provides   anticipation,

Proportional-Derivative Controller (PD)

This controller is used in processes where a proportional controller can be used, where steady-state offset is acceptable, but where some amount of anticipation is desired and no noise is present. The describing equation is

$$m(t) = \overline{m} + K_c e(t) + K_c \tau_D \frac{de(t)}{dt} \qquad (53.21)$$

and the "ideal" transfer function is

$$G_c(s) = \frac{M(s)}{E(s)} = K_c(1 + \tau_D s) \qquad (5\text{-}3.22)$$

whereas the "actual," or implemented, transfer function is

$$G_c(s) = \frac{M(s)}{E(s)} = K_c \left[\frac{(1 + \alpha)\tau_D s + 1}{\alpha\tau_D s + 1} \right] \qquad (5\text{-}3.23)$$

5-3.3   Modifications to the PID Controller and Additional Comments

Section 5-3.2 pointed out the differences in tuning parameters: K_c versus PB, and τ_I versus τ_I^R. It is unfortunate that there is no one single set, but it is a fact, and the engineer must be aware of the differences. There is yet another set of parameters used by some manufacturers, shown as

$$m(t) = \overline{m} + K_c e(t) + K_I \int e(t)\, dt + K_D \frac{de(t)}{dt} \qquad (5\text{-}3.24)$$

The three tuning parameters are in this case K_c, K_I, and K_D .

There are other common modifications found in some controllers. Figure 5-3.13a shows a common way to introduce a set point change. When this takes place, a step change in error is also introduced, as shown in Fig. 5-3.13b. Because the derivative calculation is based on the error, this calculation results in a drastic change in controller output. Such a change is unnecessary and is often detrimental to the process operation.
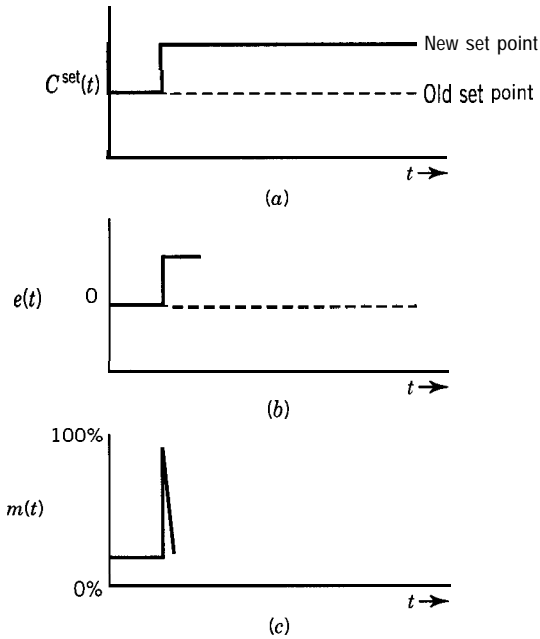
Figure 53.13 Effect of set point changes in PID control-
lers.

The most common way to avoid this problem is to use the negative of the derivative
of the controlled, or process, variable, $-\dfrac{dc(t)}{dt}$, instead of the derivative of the error.
That is,

$$m(t) = \overline{m} + K_c e(t) + \frac{K_c}{\tau_I} \int e(t)\ dt \quad K_c \tau_D \frac{dc(t)}{dt} \tag{53.25}$$

The response of both derivatives is the same when the set point is constant.

$$\frac{de(t)}{dt} = \frac{d[r(t) - c(t)]}{dt} = \frac{dr(t)}{dt} - \frac{dc(t)}{dt}$$

Under constant set point, the first derivative term on the right is zero, and thus

$$\frac{de(t)}{dt} = -\frac{dc(t)}{dt}$$

At the moment the set point change is introduced, the "new" derivative does not produce

the drastic response. Shortly after, the responses of the two derivatives become the same again. The Laplace transform of Eq. 5-3.25 is given by Eq. 5-3.26.

$$M(s) = K_c\left[\left(1 + \frac{1}{\tau_I s}\right)E(s) - \tau_D C(s)\right] \qquad (53.26)$$

or

$$M(s) = K_c\left[\left(1 + \frac{1}{\tau_I s}\right)E(s) - \frac{\tau_D s}{\alpha\tau_D s + 1}C(s)\right] \qquad (53.27)$$

or

$$M(s) = K_c'\left(1 + \frac{1}{\tau_I' s}\right)\left[E(s) - \frac{\tau_D' s + 1}{\alpha\tau_D' s + 1}C(s)\right] \qquad (53.28)$$

Figure 5-3.14 shows the block diagrams of Eqs. 5-3.27 and 5-3.28. This modification is commonly referred to as a derivative-on-process variable.

The algorithm given in Eq. 5-3.25 drastically reduces the undesirable effect of set point changes on the response of the algorithm. However, the proportional term, $K_c e(t) = K_c[r(t)$ $c(t)]$, still provides a sudden response when the set point is changed. This sudden change in response due to the proportional mode is referred to as *proportional kick*. Under some circumstances, such as large values of K_c, this response may also be detrimental to the process operation. The following modification is sometimes proposed.

$$m(t) = \overline{m} \quad K_c c(t) + \frac{K_c}{\tau_I}\int e(t)\,dt \quad K_c\tau_D\frac{dc(t)}{dt} \qquad (53.29)$$

The algorithms shown in Eqs. 5-3.15, 5-3.25, and 5-3.29 are different in their response to set point changes; however, their responses are the same for disturbances.

Another modification to the basic PID algorithm is one in which the control calculation is based on the square of the error, or

$$m(t) = \overline{m} + K_c|e(t)|\left[e(t) + \frac{1}{\tau_I}\int e(t)\,dt + \tau_D\frac{de(t)}{dt}\right] \qquad (53.30)$$

(a)   (b)
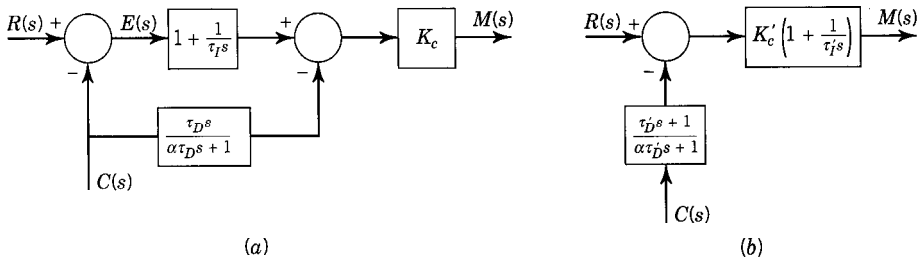
Figure 5-3.14 Block diagrams of derivative-on-process variable PID controller. (a) Eq. 5-3.27. (b) Eq. 5-3.28.

The basic idea is that when the error is small, not much corrective action is needed. When e(t) is small, $|e(t)|e(t)$ is smaller, and not much action is obtained. However, when e(t) is large, significant corrective action is needed to return to set point; in this case, $|e(t)|e(t)$ is larger and provides the required action. Although error-squared controllers are usually difficult to tune, they have shown some advantages in controlling integrating processes such as level loops.

The PID-gap controller, also referred as the dead-band controller, is not really a modification to the basic PID algorithm. In this controller, as long as the controlled variable is within some prescribed gap, or band, from set point (say $\pm 1\%$, $\pm 3\%$, or the like), no action is taken. The rationale is that these small deviations are due only to noise and are not really process deviations, so there is no need to take corrective action. Outside the prescribed band, the controller works as usual.

Let us now look at another option in controllers. Consider the heat exchanger of Fig. 5-3.2. The temperature controller is in automatic and is controlling at set point-say, at 140°C. That is, both the set point and the controlled variable are at 140°C. Now for some reason, the operator or engineer sets the controller in manual and increases the controller output. This opens the valve, permitting more steam into the heat exchanger. As a result, the temperature increases to a new value-say, 150°C. If the controller is then transferred to automatic, it will see an error, because the set point is still at 140°C and the controlled variable is now at 150°C. The controller, of course, suddenly closes the valve to correct for the deviation. This sudden change in signal to the valve represents a "bump" to the process and in some cases may be detrimental to the operation. If a *bumpless transfer* is desired when transferring from manual to automatic, the error must be zero; that is, the set point and the controlled variable must be equal. The error can be made zero either by manually reducing the controller output to bring the temperature back to 140°C or by increasing the set point to 150°C to match the temperature. Once either of these actions is taken, a **bumpless** transfer results when the controller is transferred to automatic. Computer-based controllers offer a standard option called *tracking* or, specifically in this case, *process variable tracking (PV-tracking),* which allows a **bumpless** transfer automatically. If this option is selected, whenever the controller is in manual, the set point is forced to be equal to the controlled variable; that is, the set point tracks the controlled variable. This action results in a zero error while the controller is in manual and, therefore, at the moment of its transfer back to automatic. Once the controller is in automatic, the set point remains at the new value, 150°C in our example, not at the original value, 140°C. Note that PV-tracking is an option and does not have to be selected when configuring the controller. The tracking options are quite useful in control strategies for safety and improved performance.

5-3.4 Reset Windup and Its Prevention

The problem of reset windup is an important and realistic one in process control. It may occur whenever a controller contains the integral mode of control. Let's use the heat exchanger control loop shown in Fig. 5-3.2 to explain this problem.

Suppose that the process inlet temperature drops by an unusually large amount. This disturbance will reduce the outlet temperature. The controller (PI or PID) will in turn ask the steam valve to open. Because this is a fail-closed valve, the signal from the controller will increase until, because of the reset action, the outlet temperature equals the desired set point. But suppose that in restoring the controlled variable to set point,
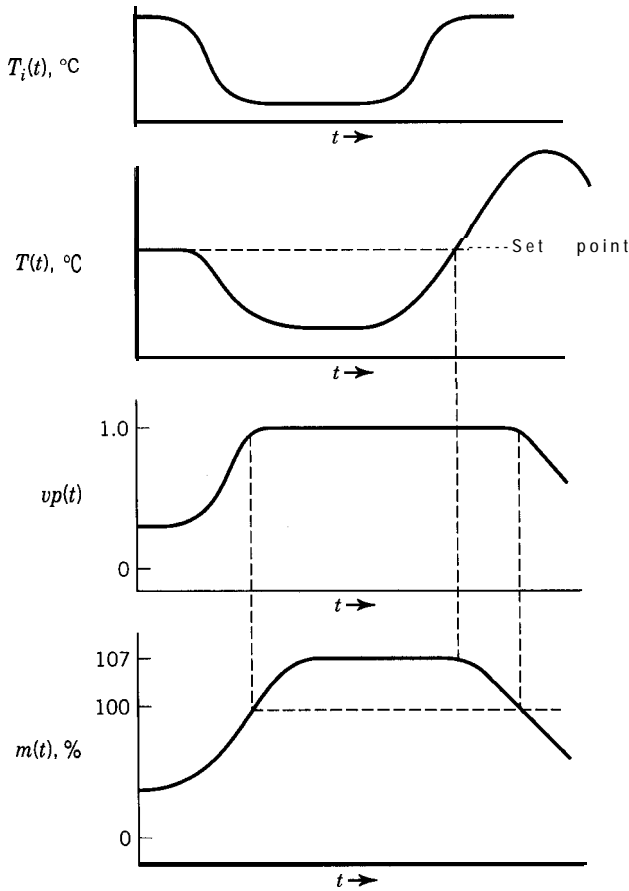
Figure 53.15 Heat exchanger control-reset windup.

the controller integrates up to 100% because the drop in inlet temperature is too large. At this point the steam valve is wide open, so the control loop cannot do any more. Essentially, the process is out of control. This is demonstrated graphically in Fig. 5-3.15, which shows that when the valve is fully open, the outlet temperature is not at set point. Because there is still an error, the controller will try to correct for it by further increasing (integrating the error) its output, even though the valve will not open more after 100%. The output of the controller can in fact integrate above 100%. Some controllers can integrate between − 15% and 115%, others between − 7% and 107%, and still others between − 5% and 105%. Analog controllers can also integrate outside their limits of 3 to 15 psig, or 4 to 20 mA. Let us suppose the controller being used can integrate up to 107%. At that point, the controller cannot increase its output anymore; its output has saturated. This state is also shown in Fig. 5-3.15. This saturation is due to the reset action of the controller and is referred to as *reset windup*.

Suppose now that the inlet temperature goes back up; the outlet process temperature will in turn start to increase, as also shown in Fig. 5-3.15. This figure shows that the outlet temperature reaches and passes the set point and that the valve remains wide open when, in fact, it should be closing. The valve is not closing because the controller
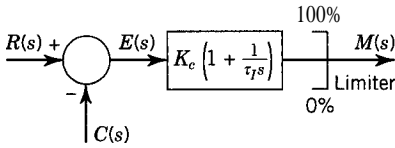
Figure 53.16 Limiting controller output.

must integrate down to 100% before it starts to close. By the time this happens, the outlet temperature has overshot the set point by a significant amount.

As we have said, this problem of reset windup may occur whenever integration is present in the controller. It can be avoided if the controller is set in manual as soon as its output reaches 100% (or 0%); this action will stop the integration. The controller can be set back to automatic when the temperature starts to decrease (or increase) again. The disadvantage of this operation is that it requires the operator's attention. Note that the prevention of reset windup requires stopping the integration, not limiting the controller output, when the controller reaches the 0% or 100% limit. Fig. 5-3.16 shows a limiter on the output of the controller that does not prevent windup. Although the output does not go beyond the limits, the controller may still be internally wound up, because it is the integral mode that winds up.

There is a very ingenious way to limit the integration when the controller output reaches its limits. Consider the PI controller transfer function

$$M(s) = K_c \left[1 + \frac{1}{\tau_I s} \right] E(s)$$

or

$$M(s) = K_c E(s) + M_I(s) \tag{53.31}$$

where

$$M_I(s) = \frac{K_c}{\tau_I s} E(s)$$

or

$$\tau_I s M_I(s) = K_c E(s) \tag{53.32}$$

From Eq. 5-3.31, we get

$$K_c E(s) = M(s) - M_I(s) \tag{5-3.33}$$

Equating Eqs. 5-3.32 and 5-3.33 and rearranging yield

$$M_I(s) = \frac{1}{\tau_I s + 1} M(s) \tag{5-3.34}$$

The implementation of Eqs. 5-3.31 and 5-3.34 is shown in Fig. 5-3.17. When the limiter is placed as shown in the figure, $M_I(s)$ will be automatically limited. $M_I(s)$ is always lagging M(s) with a gain of 1 and an adjustable parameter τ_I, so it can never get outside
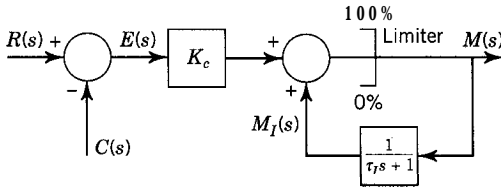
Figure 53.17 Reset-feedback implementation block diagram.

the range within which M(s) is limited. In other words, if $M(s)$ reaches one of its limits, $M_I(s)$ will approach that limit-say, 100%. Then, at the moment the error turns negative, the controller output becomes

$$m(t) = 100 + K_c e(t) < 100\% \qquad \text{as } e(t) < 0$$

That is, the controller output will come off the limit, closing the valve, the instant the controlled variable crosses the set point!

Note that at steady state, the error is zero:

$$M(s) = M_I(s) = M_I(s) + K_c E(s)$$

and for this to be true, $E(s) = 0$. Thus there is no offset. This way to implement this reset windup protection is commonly referred to as *reset feedback* (RF).

Reset windup protection is an option that must be bought in analog controllers. It is a standard feature in many computer-based controller.

Reset windup occurs any time a controller is not in charge, such as when a manual bypass valve is open or when there is insufficient manipulated variable power. It also typically occurs in batch processes, in cascade control, and when a final control element is driven by more than one controller, as in override control schemes. Cascade control is presented in Chapter 10, override control in Chapter 11.

5-3.5 Feedback Controller Summary

This section has presented the subject of process controllers. The purpose of the controllers is to adjust the manipulated variable to maintain the controlled variable at set point. We considered the significance of the controller and saw how to choose its action, reverse or direct. The different types of controllers were also presented, including the significance of the tuning parameters gain (K_c) or proportional band (PB), reset time (τ_I) or reset rate (τ_I^R), and rate time (τ_D). Finally, the subject of reset windup was presented and its significance discussed.

We have not discussed yet the important subject of obtaining the optimum setting of the tuning parameters. "Tuning the controller" is presented in Chapter 7.

5-4 SUMMARY

In this chapter, we looked at some of the hardware necessary to build a control system. The chapter began with a brief look at some terms related to sensors and transmitters

and a discussion of the parameters that describe these devices. It continued with some important considerations related to control valves, such as fail action, sizing, and characteristics. The reader is referred to Appendix C for more information on sensors, transmitters, and valves.

A discussion of feedback process controllers followed. The four most common types of controllers were presented, along with some modifications. The physical significance of their parameters was explained. The tuning of these parameters is presented in Chapter 7.

We are now ready to apply what we have learned in the first five chapters of this book to design process control systems.

REFERENCES

1. "Control Valve Handbook." 1977. Marshalltown, Iowa: Fisher Controls Co.
2. "Masoneilan Handbook for Control Valve Sizing." Norwood, Mass.: Masoneilan International, Inc.
3. "Fisher Catalog 10." 1976. Marshalltown, Iowa: Fisher Controls Co.

PROBLEMS

5-1. For each of the following cases, calculate the gain in percent transmitter output (%TO) per variable unit (specify units), write the transfer function, and draw the block diagram.

(a) A temperature transmitter with a range of 100 to 150°C and a time constant of 1.2 min.

(b) A temperature transmitter with a range of 100 to 350°F and a time constant of 0.5 min.

(c) A pressure transmitter with a range of 0 to 50 psig and a time constant of 0.05 min.

(d) A level transmitter with a range of 0 to 8 ft and a negligible time constant.

(e) A flow transmitter consisting of a differential pressure transmitter measuring the pressure drop across an orifice, sized for a maximum flow of 750 gpm, when the flow is 500 gpm. The time constant is negligible.

5-2. Liquid levels in storage tanks are frequently determined by measuring the pressure at the bottom of the tank. In one such tank, the material stored in the tank was changed, and an overflow resulted. Why? (Copyright 1992 by the American Institute of Chemical Engineers; reproduced by permission of the Center for Chemical Process Safety of AIChE.)

5-3. An operator was told to control the temperature of a reactor at 60°C. The operator set the set point of the temperature controller at 60. The scale actually indicated 0 to 100% of a temperature range of 0 to 200°C. This caused a run-away reaction that over-pressurized the vessel. Liquid was discharged and injured the operator. What was the set point temperature the operator actually set? (Copyright 1992 by the American Institute of Chemical Engineers; reproduced by permission of the Center for Chemical Process Safety of AIChE.)

5-4. Specify the proper fail-safe action for the valves in the following services. Specify either fail-open or fail-close.

(a) A flammable solvent is heated by steam in a heat exchanger. The valve manipulates the flow of steam to the exchanger.

(b) A valve manipulates the flow rate of reactant to a reactor vessel. The reaction is exothermic.

(c) A valve manipulates the flow rate of reactant to a reactor vessel. The reaction is endothermic.

(d) A valve manipulates the flow of natural gas (combustible) to a furnace. Another valve manipulates the flow of combustion air to the same furnace.

(Copyright 1992 by the American Institute of Chemical Engineers; reproduced by permission of the Center for Chemical Process Safety of AIChE.)

5-5. Size a control valve to regulate the flow of 50 psig saturated steam to a heater. The nominal flow is 1200 lb/h, and the outlet pressure is 5 psig.

(a) Obtain the C_v coefficient for 50% overcapacity (assume $C_f = 0.8$).

(b) Obtain the valve gain in (lb/h)/%CO (assume the valve is linear with constant pressure drop).

5-6. The nominal liquid flow through a control valve is 52,500 lb/h, and the required maximum flow is 160,000 lb/h. Operating conditions call for an inlet pressure of 229 psia and an outlet pressure of 129 psia. At the flowing temperature of 104°F, the liquid has a vapor pressure of 124 psia, a specific gravity of 0.92, and a viscosity of 0.2 cp. The critical pressure of the liquid is 969 psia. (See Appendix C for sizing formulas for flashing liquids.) Obtain the C_v coefficient for the valve.

5-7. A control valve is to regulate the flow of a gas with a molecular weight of 44. Process design conditions call for a nominal flow of 45,000 scfh; an inlet pressure and a temperature of 110 psig and 100°F, respectively; and an outlet pressure of 11 psig.

(a) Obtain the C_v coefficient for 100% overcapacity (assume $C_f = 0.8$).

(b) Obtain the valve gain in scfh/%CO (assume the valve is linear with constant pressure drop).

5-8. You are asked to design a control valve to regulate the flow of benzene in the line shown in Fig. P5-1. The process design calls for a nominal flow of 140,000 kg/h and a temperature of 155°C. At the design flow, the frictional pressure drop in the line between points 1 and 2 is 100 kPa. The density of benzene at the flowing temperature is 730 kg/m^3. Assume that the pressures shown in the diagram do not change with flow.

(a) Recommend a proper location for the control valve.

(b) Size the valve for 100% overcapacity.

5-9. In the line sketched in Fig. P5-2, ethylbenzene flows at 800 gpm (nominal) and 445°F (density $= 42.0$ lb/ft^3). The frictional pressure drop between points 1 and 2 is 12.4 psi.

(a) Recommend a proper location for the control valve.

(b) Size the valve for 100% overcapacity.

5-10. The nominal flow of a liquid through a control valve is 450 gpm. At this flow, the frictional pressure drop in the line is 15 psi. The total pressure drop available across the valve and line is 20 psi, independent of flow, and the specific gravity of the liquid is 0.85.

(a) Size the valve for 100% overcapacity.
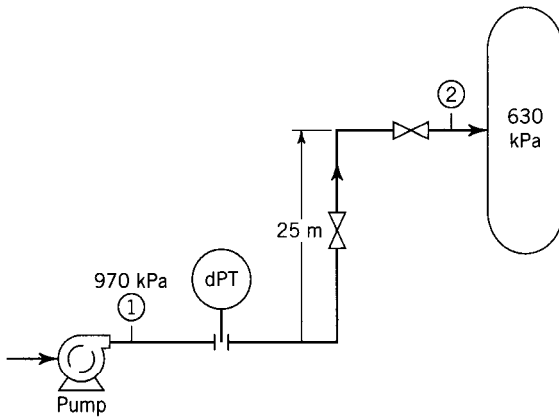
Figure P5-1 Benzene process for Problems 5-8 and 5-12.

    (b) Find the flow through the valve when it is fully opened. (Hint: It is not 900 gpm.)

    (c) Calculate the gain through the valve at design flow, assuming it has linear inherent characteristics.

    (d) Obtain the rangeability of the valve.

    State your assumptions in solving this problem.

5-11. Repeat Problem 5-10 if the total available pressure drop is increased to 35 psi to have more pressure drop across the valve. Estimate also the incremental annual cost of running the pump to provide the additional 15 psi of pressure drop. Use the economic parameters of Example 5-2.3 and a pump efficiency of 70%.

5-12. The valve of Problem 5-8 has inherent equal percentage characteristics with a rangeability parameter of 50.

    (a) Find the flow through the valve when it is fully opened. (Hint: It is not 280,000 kg/h.)

    (b) Obtain the rangeability of the control valve.
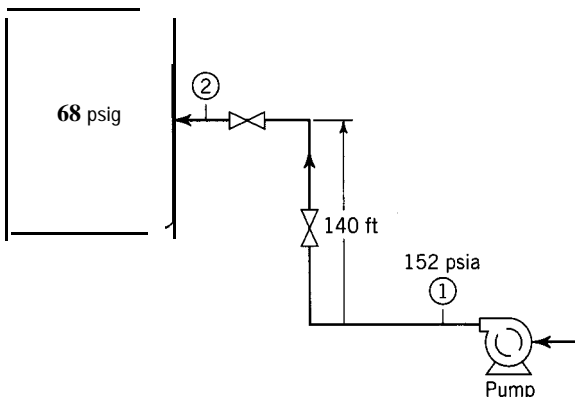
Figure P5-2 Ethylbenzene process for Problems 5-9 and 5-13.

(c) Estimate the gain of the valve at the design flow, in $(kg/h)/\%CO$.

(d) Plot the normalized installed characteristics.

State your assumptions in solving this problem.

5-13. The valve of Problem 5-9 has linear inherent characteristics.

(a) Obtain the flow through the valve when it is fully opened. (Hint: It is not 1600 gpm.)

(b) Calculate the rangeability of the control valve.

(c) Find the gain of the valve at the design flow, in $gpm/\%CO$.

(d) Plot the normalized installed characteristics.

State your assumptions in solving this problem.

5-14. Derive Eqs. 5-2.25 and 5-2.27 for a gas if inlet and outlet pressures are constant with flow. Would the equation also apply if the mass flow were replaced with the flow in scfh?

5-15. *Design of gas flow control loop.* A flow control loop, consisting of an orifice in series with the control valve, a differential pressure transmitter, and a controller, is to be designed for a nominal process flow of 150,000 scfh of air. Valve inlet conditions are 100 psig and $60°F$, and the outlet pressure is 80 psig. The valve has linear characteristics, and a square root extractor is built into the transmitter so that its output signal is linear with flow. The valve time constant is 0.06 min, and the transmitter time constant is negligible. A proportional-integral (PI) controller controls the flow.

(a) Obtain the valve capacity factor, C_v, and the gain of the valve. Size it for 100% overcapacity, and assume $C_f = 0.9$ (Masoneilan).

(b) Calculate the gain of the transmitter if it is calibrated for a range of 0 to 250,000 scfh.

(c) Draw the instrumentation diagram and the block diagram of the flow control loop, showing the specific transfer functions of the controller, the control valve, and the flow transmitter.

5-16. Consider the pressure control system shown in Fig. P5-3. The pressure transmitter, PT25, has a range of 0 to 100 psig. The controller, PC25, is a proportional-only controller, its bias value is set at mid-scale, and its set point is 10 psig. Obtain the correct action of the controller and the proportional band required so that when the pressure in the tank is 30 psig, the valve will be wide open.
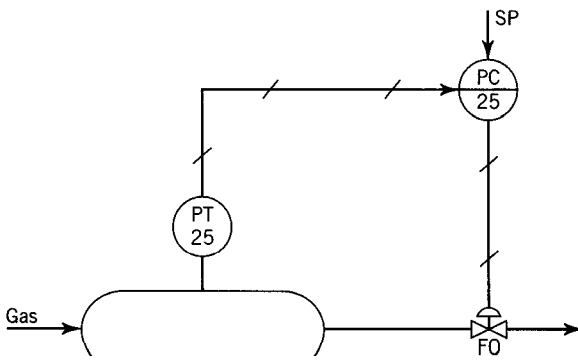
Figure P5-3 Pressure control system for Problem 5-16.
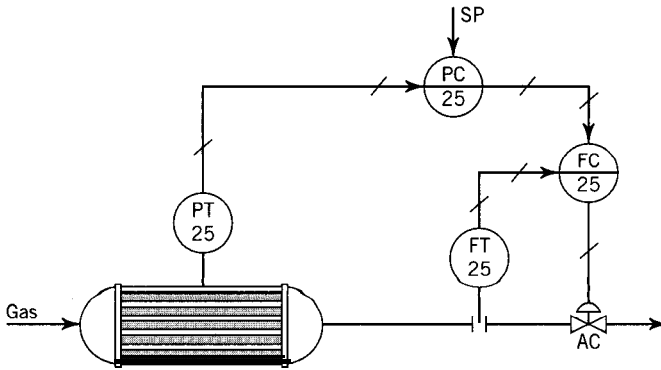
Figure P5-4 Pressure control system for Problem 5-17.

5-17. Let us change the pressure control system of Problem 5-16. The new control scheme is shown in Fig. P5-4. This control scheme is called cascade control; its benefits and principles are explained in Chapter 10. In this scheme, the pressure controller sets the set point of the flow controller. The pressure transmitter has a range of 0 to 100 psig, and the flow transmitter range is 0 to 3000 scfh. Both controllers are proportional-only. The normal flow rate through the valve is 1000 scfh, and to give this flow, the valve must be 33% opened. The control valve has linear characteristics and is fail-open (air-to-close).

(a) Obtain the action of both controllers.

(b) Choose the bias values $(\bar{m})$ for both controllers so that no offset occurs in either controller.

(c) Obtain the proportional band setting of the pressure controller so that when the tank pressure reaches 40 psig, the set point to the flow controller is 1700 scfh. The set point of the pressure controller is 10 psig.

(d) Obtain the action of both controllers if the valve were to be fail-closed (air-to-open).

5-18. Consider the level loop shown in Fig. 5-3.3. The steady-state operating conditions are $\bar{f}_i = \bar{f}_o = 150$ gpm and $\bar{h} = 6$ ft. For this steady state, the FC valve requires a 50% signal. The level transmitter has a range of 0 to 20 ft. A proportional-only controller, with $K_c = 1$, is used in this process. Calculate the offset if the inlet flow increases to 170 gpm and the valve requires 57% to pass this flow. Report the offset in % of scale and in feet.

5-19. A controller receives a signal from a temperature transmitter with a range of 100 to 150°C. Assume the controller is proportional-integral (PI) with a gain of 3 %CO/%TO and an integral (or reset) time of 5 min.

(a) Write the transfer function of the controller relating the output $M(s)$ to the error signal E(s); assume both signals are in percent of range. Show the numerical values of the controller parameters.

(b) Calculate the gain of the transmitter and write its transfer function, assuming it can be represented by a first-order lag with a time constant of 0.1 min.

(c) Draw a block diagram of the transmitter and controller, showing all transfer functions. The input signals to the diagram are the process temperature $T(s)$ and its set point $T^{set}(s)$, both in °C.

(d) Assume that a sustained step change in set point of 1°C is applied to the controller, and that because of a loss of the signal to the control valve, the process temperature remains constant and equal to the original set point. Calculate the sustained error in %TO and the controller output in %CO at the following times: right after the change in set point, 5 min later, and 10 min later. Sketch a plot of the error and the controller output versus time.

5-20. Consider the concentration control loop for the two stirred reactors shown in Fig. P5-5. The rate of consumption of reactant A in each reactor is given by the formula

$$r_A(t) = kc_A(t)$$

where $r_A(t)$ is the reaction rate, lbmoles/gal-min, and $c_A(t)$ is the concentration of reactant A in the reactor, lbmoles/gal. Assume the reactor volumes, V, and V_2, gal, the rate coefficients, k_1 and k_2, min^{-1}, and the density of the fluid ρ, lb/gal, are constant.

(a) Obtain the transfer functions for the concentration from the reactors, $C_{A_1}(s)$ and $C_{A_2}(s)$, to the input variables; the flow F(s); and the inlet concentration, C„(s). Draw the block diagram for the reactors.

(b) Size the control valve for 100% overcapacity and a nominal flow of 100 gpm. The pressure drop across the valve is constant at 9 psi, and the specific gravity of the reactant stream is 0.8 1. Assume that the time constant of the valve actuator is negligible and that the valve is air-to-open with linear characteristics. Calculate the gain of the valve and draw the block diagram for the valve.

(c) The concentration transmitter has a calibrated range of 0 to 0.2 lbmoles/gal and a time constant of 0.5 min. Calculate the transmitter gain and draw the block diagram for the transmitter.
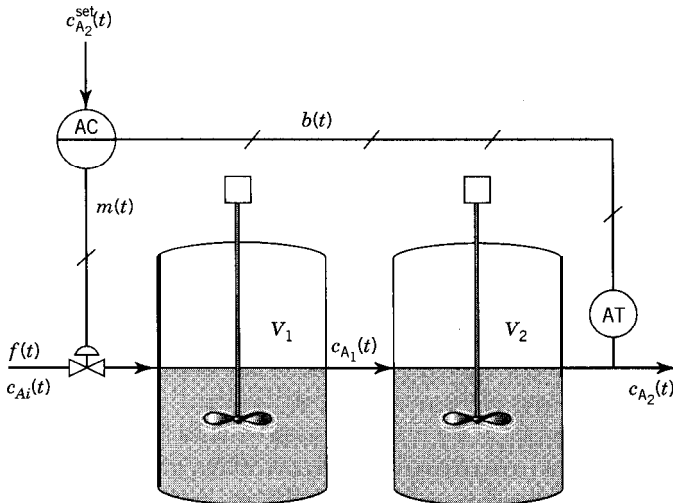
Figure P5-5 Stirred reactors in series for Problem 5-20.

(d) Draw a block diagram of the concentration control loop, showing all transfer functions. Use a proportional-integral-derivative (PID) controller. Should the controller action be direct or reverse?

5-21. Show how the relationships given in Eq. 5-3.20 were obtained from Eqs. 5-3.16 and **5-3.19**. You may assume $\alpha = 0$.

Chapter 6

Design of Single-Loop Feedback Control Systems

In previous chapters we have become familiar with the dynamic characteristics of processes, sensor/transmitters, control valves, and controllers. We have also learned how to write linearized transfer functions for each of these components and to recognize the parameters that are significant to the design of automatic control systems: the steady-state gain, the time constants, and the dead time (transportation lag or time delay). In this chapter we will see how these concepts are put together to design and tune single-loop feedback control systems. We will first analyze a simple feedback control loop and learn how to draw a block diagram for it and determine its characteristic equation. Then we will examine the significance of the characteristic equation in terms of how it can be used to determine the stability of the loop. We will use two methods to determine the stability of the loop: the direct substitution method and Routh's test.

The methods that we will study in this chapter are most applicable to the design of feedback control loops for industrial processes. Two other design techniques, root-locus and frequency response analysis, which have been traditionally applied to inherently linear systems, will be presented in Chapters 8 and 9, respectively.

6-1 THE FEEDBACK CONTROL LOOP

The concept of feedback control, though it is more than 2000 years old, did not find practical application in industry until James Watt applied it to control the speed of his steam engine about 200 years ago. Since then, industrial applications have proliferated to the point where, today, almost all automatic control systems include feedback control. None of the advanced control techniques that have been developed in the last 50 years to enhance the performance of feedback control loops have been able to replace it. We will study these advanced techniques in later chapters.

To review the concept of feedback control, let us again look at the heat exchanger example of Chapter 1. Figure 6-1.1 presents a sketch of the exchanger. Our objective is to maintain the outlet temperature of the process fluid, $T_o(t)$, at its desired value or set point, T_o^{set}, in the presence of variations of the process fluid flow, IV(t), and inlet temperature, $T_i(t)$. We select the steam flow, $W_s(t)$, as the variable that can be adjusted
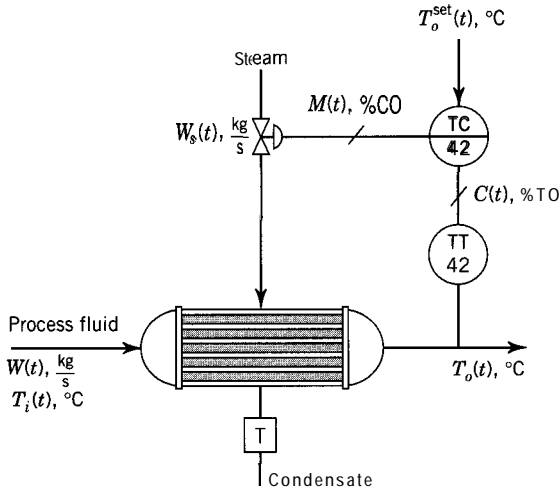
Figure **6-1.1** Feedback control loop for temperature control of a heat exchanger.

to control the outlet temperature; the amount of energy supplied to the process fluid is proportional to the steam flow.

Feedback control works as follows: A sensor/transmitter (TT42) measures the outlet temperature or *controlled variable, $T_o(t)$*; generates a signal C(t) proportional to it; and sends it to the controller (TC42), where it is compared to the set point, T_o^{set}. The controller then calculates an output signal or *manipulated variable,* M(t), on the basis of the error-that is, the difference between the measurement and the set point. This controller output signal is sent to the actuator of the steam control valve. The valve actuator positions the valve in proportion to the controller output signal. Finally, the steam flow, a function of the valve position, determines the energy rate to the exchanger and therefore the controlled outlet temperature.

The term *feedback* derives from the fact that the controlled variable is measured and this measurement is "fed back" to reposition the steam valve. This causes the signal variations to move around the loop as follows:

> *Variations in outlet temperature are sensed by the sensorltransmitter and sent to the controller, causing the controller output signal to vary. This in turn causes the control valve position and consequently the steam flow to vary. The variations in steam flow cause the outlet temperature to vary, thus completing the loop.*

This loop structure is what makes feedback control simultaneously simple and effective. When properly tuned, the feedback controller can maintain the controlled variable at or near the set point in the presence of any disturbance (such as process flow and inlet temperature) without knowledge of what the disturbance is or of its magnitude.

As we saw in Section 5-3, the most important requirement of the controller is the direction of its action (or simply *action),* direct or reverse. In the case of the temperature controller, the correct action is reverse, because an increase in temperature requires a decrease in the controller output signal to close the valve and reduce the steam flow.
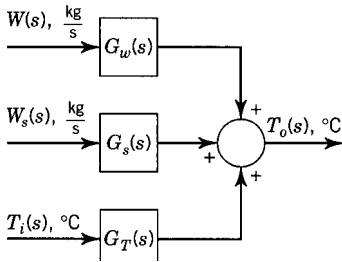
Figure 6-1.2 Block diagram of the heat exchanger of Figure 6- 1.1.

This assumes that the control valve is air-to-open so that the steam flow will be cut off in case of loss of electric power or instrument air pressure (fail-closed).

The performance of the control loop can best be analyzed by drawing the block diagram for the entire loop. To do this, we first draw the block for each component and then connect the output signal from each block to the next block. Let us start with the heat exchanger. In Chapters 3 and 4 we learned that the linear approximation to the response of the output of any process can be represented by the sum of a series of blocks, one for each input variable. As Fig. 6-1.2 shows, the block diagram for the heat exchanger consists of three blocks, one for each of its three inputs: the process flow, $W(s)$, inlet temperature, $T_i(s)$, and steam flow, $W_s(s)$. The corresponding transfer functions are G,(s), $G_T(s)$, and G,(s).

Figure 6-1.3 shows the complete block diagram for the feedback control loop. To simplify the discussion that follows, we have purposely omitted the inlet temperature, $T_i(s)$, as an input signal. This effectively assumes that the inlet temperature is constant and selects the process flow, $W(s)$, as representative of either disturbance. The symbols in Figure 6-1.3 are as follows:

$$E(s) = \text{the error, \% transmitter output (\%TO)}$$

$$G_c(s) = \text{the controller transfer function, \%CO/\%TO}$$
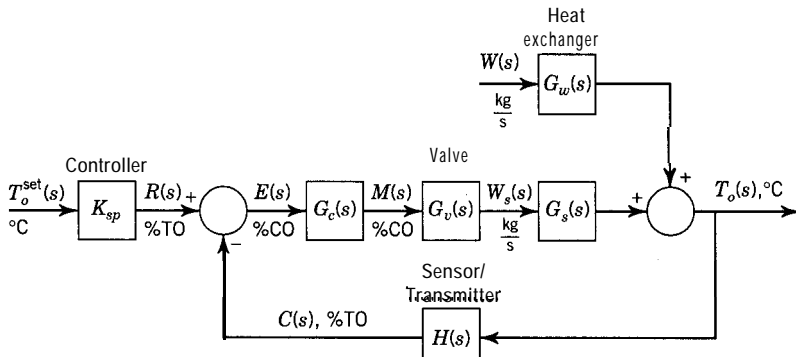
Figure 6-1.3 Block diagram of heat exchanger temperature control loop.

$$G_v(s) = \text{the control valve transfer function, (kg/s)/\%CO}$$

$$H(s) = \text{the sensor/transmitter transfer function, \%TO/°C}$$

$$K_{sp} = \text{the scale factor for the temperature set point, \%TO/°C}$$

It is important at this point to note the correspondence between the blocks (or groups of blocks) in the block diagram, Fig. 6-1.3, and the components of the control loop, Fig. 6-1.1. This comparison is facilitated by matching the symbols used to identify the various signals. It is also important to recall from Chapter 3 that the blocks on the diagram represent linear relationships between the input and output signals and that the signals are deviations from initial steady-state values and are not absolute variable values.

The term K_{sp} is a scale factor that converts the set point, usually calibrated in the same units as the controlled variable, to the same basis as the transmitter signal-that is, °C to %TO. It can be shown that for the measurement and the set point to be on the same scale, K_{sp} must be numerically equal to the transmitter gain.

The sign convention used in the block diagram of Fig. 6-1.3 agrees with the convention used in Section 5-3 for calculation of the error (set point — measurement). This convention will be used throughout this book. Note that this makes the sign around the loop negative if the gains of all the blocks and summers in the loop are positive, as they are in this case. *A negative feedback gain is a requirement for stability.* Following this convention, a reverse-acting controller must have a positive gain, and a direct-acting controller must have a negative gain, as you can verify by analyzing the controller section of the block diagram. The convention is not selected this way to confuse you, but to emphasize graphically the negative feedback gain on the block diagram (otherwise, the minus sign would be hidden in the controller gain).

6-1.1   Closed-Loop Transfer Function

We can see by inspection of the closed-loop block diagram of Fig. 6-1.3 that the loop has one output signal, the controlled variable $T_o(s)$, and two input signals, the set point $T_o^{set}(s)$, and the disturbance W(s). Because the steam flow is connected to the outlet temperature through the control loop, we might expect that the "closed-loop response" of the system to the various inputs would be different from the response when the loop is "open." Most control loops can be opened by flipping a switch on the controller from the automatic to the manual position (see Section 5-3). When the controller is in the manual position, its output does not respond to the error signal, so it is independent of the set point and measurement signals. In "automatic," on the other hand, the controller output varies when the measurement signal varies.

We can determine the closed-loop transfer function of the loop output with regard to any of its inputs by applying the rules of block diagram algebra (see Chapter 3) to the diagram of the loop. To review, suppose we want to derive the response of the outlet temperature $T_o(s)$ to the process flow W(s). We first write the equations for each block in the diagram, as follows:

$$E(s) = K_{sp}T_o^{set}(s) - C(s) \qquad \text{(6-1.1)}$$

$$M(s) = G_c(s)E(s) \tag{6-1.2}$$

$$W_s(s) = G_v(s)M(s) \tag{6-1.3}$$

$$T_o(s) = G_s(S)W_s(s) + G_w(s)W(s) \tag{6-1.4}$$

$$C(s) = H(s)T_o(s) \tag{6-1.5}$$

Next we assume that the set point does not vary-that is, its deviation variable is zero—

$$T_o^{set}(s) = 0$$

and eliminate all the intermediate variables by combining Eqs. 6-1.1 through 6-1.5. The result is

$$T_o(s) = G_s(s)G_v(s)G_c(s)[- H(s)T_o(s)] + G_w(s)W(s) \tag{6-1.6}$$

Solving for $T_o(s)$ and dividing by W(s), we get

$$\boxed{\frac{T_o(s)}{W(s)} = \frac{G_w(s)}{1 + H(s)G_s(s)G_v(s)G_c(s)}} \tag{6-1.7}$$

This is the closed-loop transfer function between the process flow and the outlet temperature. Similarly, if we let $W(s) = 0$ and combine Eqs. 6-1.1 through 6-1.5, the closed-loop transfer function between the set point and the outlet temperature results.

$$\boxed{\frac{T_o(s)}{T_o^{set}(s)} = \frac{K_{sp}G_s(s)G_v(s)G_c(s)}{1 + H(s)G_s(s)G_v(s)G_c(s)}} \tag{6-1.8}$$

As we saw in Chapter 3, the denominator is the same for both inputs, whereas the numerator is different for each input. We recall further that the denominator is 1 plus the product of the transfer functions of all the blocks that are in the loop itself and that the numerator of each transfer function is the product of the blocks that are in the direct path between the specific input and the output of the loop. These results apply to any block diagram that contains a single loop.

It is enlightening to check the units of the product of the blocks in the loop, as follows:

$$H(s) \cdot G_s(s) \cdot G_v(s) \cdot G_s(s) = \text{dimensionless}$$

$$\left(\frac{\%TO}{°C}\right) \cdot \left(\frac{°C}{kg/s}\right) \cdot \left(\frac{kg/s}{\%CO}\right) \cdot \left(\frac{\%CO}{\%TO}\right)$$

This shows that the product of the transfer functions of the blocks in the loop is dimensionless. as it should be. We can also verify that the units of the numerator of each
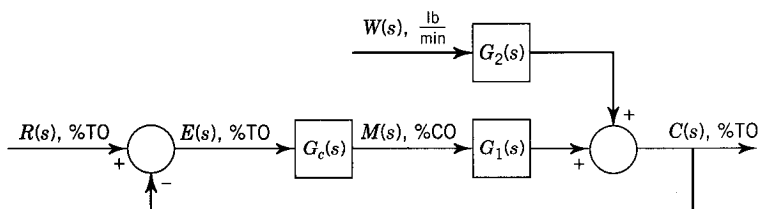
Figure 6-1.4 Simplified block diagram of a feedback control loop.

of the closed-loop transfer functions are the units of the output variable divided by the units of the corresponding input variable.

Simplified Block Diagram

It is convenient to simplify the block diagram of Fig. 6-1.3 by combining blocks. Following the rules of block diagram algebra from Chapter 3 yields Fig. 6-1.4. The transfer functions of the simplified diagram are

$$G_1(s) = G_v(s)G_s(s)H(s) \tag{6-1.9}$$

$$G_2(s) = G_w(s)H(s) \tag{6-1.10}$$

In the simplified diagram, the loop signals are in percent of range and the feedback gain is unity, which is why the loop in the diagram is sometimes called a *unity feedback loop.* The closed-loop transfer function of the output signal, which is now the transmitter output, is

$$C(s) = \frac{G_c(s)G_1(s)}{1 + G_c(s)G_1(s)} R(s) + \frac{G_2(s)}{1 + G_c(s)G_1(s)} W(s) \tag{6-1.11}$$

where $R(s)$ is the reference signal (set point) in %TO. Except for the name of the flow disturbance, the block diagram of Fig. 6-1.4 can represent any feedback control loop.

The following example demonstrates how to develop the closed-loop transfer function from the principles we learned in Chapters 3, 4, and 5.

EXAMPLE 6-1.1 *TEMPERATURE CONTROL OF A CONTINUOUS STIRRED TANK HEATER*

The stirred tank sketched in Fig. 6-1.5 is used to heat a process stream so that its premixed components achieve a uniform composition. Temperature control is important because a high temperature tends to decompose the product, whereas a low temperature results in incomplete mixing. The tank is heated by steam condensing inside a coil. A proportional-integral-derivative (PID) controller is used to control the temperature in
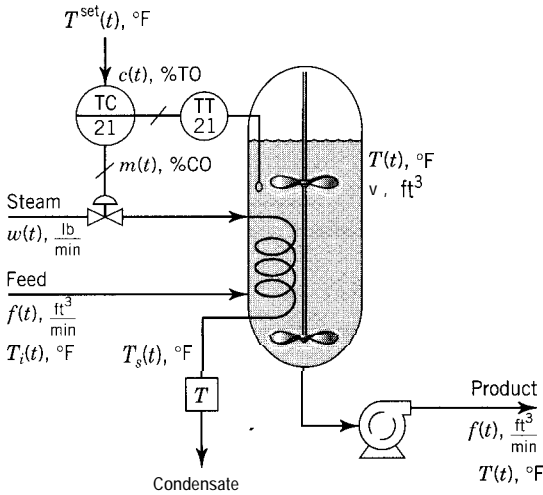
Figure 6-1.5 Temperature control of the stirred tank heater of Example 6- 1.1.

the tank by manipulating the steam valve position. Derive the complete block diagram and the closed-loop transfer function from the following design data.

Process.   The feed has a density ρ of 68.0 lb/ft^3 and a heat capacity c_p of 0.80 Btu/lb-°F. The volume V of liquid in the reactor is maintained constant at 120 ft^3. The coil consists of 205 ft of 4-in. schedule 40 steel pipe that weighs 10.8 lb/ft and has a heat capacity of 0.12 Btu/lb-°F and an outside diameter of 4.500 in. The overall heat transfer coefficient U, based on the outside area of the coil, has been estimated as 2.1 $Btu/min\text{-}ft^2\text{-}°F$. The steam available is saturated at a pressure of 30 psia; it can be assumed that its latent heat of condensation λ is constant at 966 Btu/lb. It can also be assumed that the inlet temperature T_i is constant.

Design Conditions. The feed flow f at design conditions is 15 ft^3/min, and its temperature T_i is 100°F. The contents of the tank must be maintained at a temperature T of 150°F. Possible disturbances are changes in feed rate and temperature.

Temperature Sensor and Transmitter. The temperature sensor has a calibrated range of 100 to 200°F and a time constant τ_T of 0.75 min.

Control Valve. The control valve is to be designed for 100% overcapacity, and pressure drop variations can be neglected. The valve is an equal percentage valve with a rangeability parameter α of 50. The actuator has a time constant τ_v of 0.20 min.

SOLUTION

Our approach will be to derive the equations that describe the dynamic behavior of the tank, the control valve, the sensor/transmitter, and the controller. Then we will Laplace-transform them to obtain the block diagram of the loop.

Process. An energy balance on the liquid in the tank, assuming negligible heat losses, perfect mixing, and constant volume and physical properties, results in the equation

$$V\rho c_v \frac{dT(t)}{dt} = f(t)\rho c_p T_i + UA[T_s(t) - T(t)] - f(t)\rho c_p T(t)$$

<div align="right">1 eqn., 2 unk. (T, T_s)</div>

where

A = the heat transfer area, ft^2

$T_s(s)$ = the condensing steam temperature, $°F$

and the other symbols have been defined in the statement of the problem. For the liquid contents of the tank, the c_v in the accumulation term is essentially equal to c_p .

An energy balance on the coil, assuming that the coil metal is at the same temperature as the condensing steam, results in

$$C_M \frac{dT_s(t)}{dt} = w(t)\lambda - UA[T_s(t) - T(t)] \qquad \text{2 eqn. 3 unk. } (w)$$

where

$w(t)$ = the steam rate, lb/min

C_M = heat capacitance of the coil metal, $\text{Btu/}°F$

Because the steam rate is the output of the control valve and an input to the process, our process model is complete.

Linearization and Laplace Transformation. By the methods presented in Section 2-6, we obtain the linearized tank model equations in terms of deviation variables.

$$V\rho c_p \frac{d\Gamma(t)}{dt} = \rho c_p(T_i - \bar{T})F(t) + UA\Gamma_s(t) - (UA + \bar{f}\rho c_p)\Gamma(t)$$

$$C_M \frac{d\Gamma_s(t)}{dt} = \lambda W(t) - UA\Gamma_s(t) + UA\Gamma(t)$$

where $\Gamma(t)$, $F(t)$, and $W(t)$ are the deviation variables.

Taking the Laplace transform of these equations and rearranging, as we learned in Chapters 2, 3, and 4, we get

$$\Gamma(s) = \frac{K_F}{\tau s + 1} F(s) + \frac{K_s}{\tau s + 1} \Gamma_s(s)$$

$$\Gamma_s(s) = \frac{1}{\tau_c s + 1} \Gamma(s) + \frac{K_w}{\tau_c s + 1} W(s)$$

where

$$\tau = \frac{V\rho c_p}{UA + \bar{f}\rho c_p} \qquad \tau_c = \frac{C_M}{UA}$$

$$K_F = \frac{\rho c_p(T_i - \bar{T})}{UA + \bar{f}\rho c_p}$$

$$K_s = \frac{UA}{UA + \bar{f}\rho c_p} \qquad K_w = \frac{\lambda}{UA}$$

Control Valve.   The transfer function for an equal percentage valve with constant pressure drop is, from Section 5-2,

$$G_r(s) = \frac{W(s)}{M(s)} \frac{K_v}{\tau_v s + 1} \qquad \textbf{3} \text{ eqn. 4 unk. } (M)$$

where *M(s)* is the controller output signal in percent controller output (%CO), and the valve gain is, from Section 5-2:

$$K_v = \frac{\bar{w}(\ln \alpha)}{100}$$

Sensor/Transmitter *(TT21).* The sensor/transmitter can be represented by a first-order lag:

$$H(s) = \frac{C(s)}{T(s)} = \frac{K_T}{\tau_T s + 1} \qquad \textbf{4} \text{ eqn. 5 unk. (C)}$$

where $C(s)$ is the Laplace transform of the transmitter output signal, %TO, and the transmitter gain is, from Section 5-1,

$$K_T = \frac{100 - 0}{\textbf{200} \quad 100} = 1.0 \frac{\% \text{ TO}}{°F}$$

The transfer function of the PID controller is, from Section 5-3,

$$G_c(s) = \frac{M(s)}{R(s) - C(s)} = K_c \left(1 + \frac{1}{\tau_I s} + \tau_D s \right) \qquad \textbf{5} \text{ eqn. 5 unk.}$$

where K_c is the controller gain, τ_I is the integral time, and τ_D is the derivative time. This completes the derivation of the equations for the temperature control loop.

Block Diagram of *the* Loop. Figure 6-1.6 shows the complete block diagram for the loop. All of the transfer functions in the diagram have been derived here. Using the
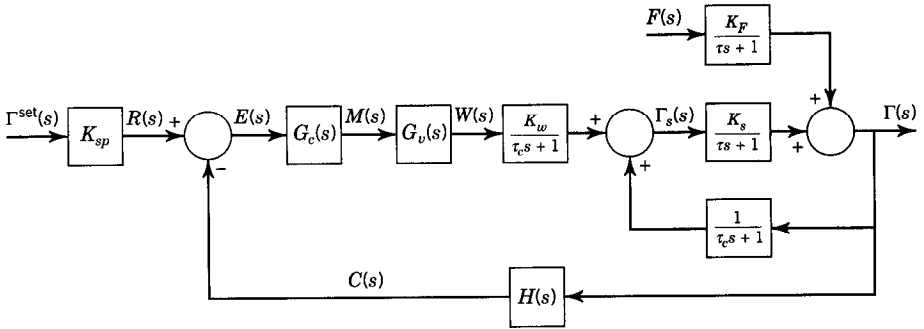
Figure 6-1.6 Block diagram of temperature control loop of stirred tank heater.

rules for block diagram manipulation we learned in Chapter 3, we obtain the simpler diagram of Fig. 6-1.7. The transfer functions in the diagram are

$$G_F(s) = \frac{K_F(\tau_c s + 1)}{(\tau s + 1)(\tau_c s + 1) - K_s}$$

$$G_s(s) = \frac{K_{..}K_c}{(\tau s + 1)(\tau_c s + 1) - K_s}$$

The closed-loop transfer functions to the inputs are

$$\frac{\Gamma(s)}{\Gamma^{set}(s)} = \frac{K_{sp}G_c(s)G_v(s)G_s(s)}{1 + H(s)G_c(s)G_v(s)G_s(s)}$$

$$\frac{\Gamma(s)}{F(s)} = \frac{G_F(s)}{1 + H(s)G_c(s)G_v(s)G_s(s)}$$

Table 6-1.1 gives the numerical values of the parameters in the transfer functions, calculated from the data given in the problem statement. The base values for the linearization are the design conditions, assumed to be the initial conditions and at steady
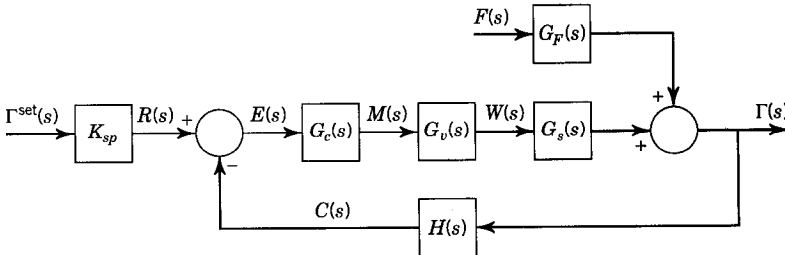
Figure 6-1.7 Simplified block diagram of temperature control loop.

Table 6-1.1 Parameters for Example 6-1.1

$A = 241.5 \text{ ft}^2$	$\tau = 4.93$ min
$C_M = 265.7$ Btu/°F	$\tau_c = 0.524$ min
$K_F = -2.06°\text{F/(ft}^3\text{/min)}$	$K_w = 1.905°\text{F/(lb/min)}$
$K_s = 0.383°\text{F/°F}$	$K_{sp} = K_T = 1.0 \%\text{TO/°F}$
$K_v = 1.652 \text{ (lb/min)/%CO}$	$\tau_v = 0.20$ min
$\tau_T = 0.75$ min	

state. From the model equations for the tank and the coil, we compute the initial steam temperature and steam flow. At steady state,

$$\bar{f}\rho c_p T_i + UA(\bar{T}_s - \bar{T}) - \bar{f}\rho C_p \bar{T} = 0$$

$$\bar{w}\lambda - UA(\bar{T}_s - \bar{T}) = 0$$

$$\bar{T}_s = \frac{(15)(68)(0.80)(150 - 100)}{(2.1)(241.5)} + 150 = 230°\text{F}$$

$$\bar{w} = \frac{(2.1)(241.5)(230 - 150)}{966} = 42.2 \text{ lb/min}$$

We can see, from Fig. 6-1.6, that the coil and the tank form a set of two interacting lags. This means that we must calculate effective time constants from the parameters in Table 6-1.1. They are 8.34 and 0.502 min, which results in the following transfer function:

$$G_1(s) = G_v(s)G_s(s)H(s)$$

$$= \frac{1.652}{0.2s + 1} \cdot \frac{1.183}{(8.34s + 1)(0.502s + 1)} \cdot \frac{1.0}{0.75s + 1}$$

where the gain is $K_w K_s /(1 - K_s) = 1.183°\text{F/(lb/min)}$. Similarly,

$$G_2(s) = G_F(s)H(s) = \frac{-3.34(0.524s + 1)}{(8.34s + 1)(0.502s + 1)} \cdot \frac{1.0}{0.75s + 1}$$

where the gain is $K_F /(1 - K_s) = -3.34°\text{F/(ft}^3\text{/min)}$. The closed-loop transform of the temperature transmitter output is then

$$C(s) = \frac{G_c(s)G_1(s)}{1 + G_c(s)G_1(s)} R(s) + \frac{G_2(s)}{1 + G_c(s)G_1(s)} F(s)$$

These transfer functions match the unity feedback loop of Fig. 6-1.4.

This example illustrates how the basic principles of process engineering can be put to work in analyzing simple feedback control loops. From the closed-loop transfer functions, we can calculate the response of the closed loop to various input forcing functions for different values of the controller tuning parameters, K_c, τ_I, and τ_D.

6-1.2 Characteristic Equation of the Loop

As we saw in the preceding discussion, the denominator of the closed-loop transfer function of a feedback control loop is independent of the location of the input to the loop and thus is characteristic of the loop. Recall from Chapter 2 that the unforced response of the loop and its stability depend on the roots of the equation that is obtained when the denominator of the transfer function of the loop is set equal to zero.

$$1 + H(s)G_s(s)G_v(s)G_c(s) = 0 \qquad (6\text{-}1.12)$$

This is the characteristic equation of the loop. Note that the controller transfer function is very much a part of the characteristic equation of the loop. **This is why the response of the loop can be shaped by tuning the controller.** The other elements that form part of the characteristic equation are the sensor/transmitter, the control valve, and that part of the process that affects the response of the controlled variable to the manipulated variable-that is, G,(s). On the other hand, the process transfer function related to the disturbance, G,(s), is not part of the characteristic equation.

To show that the characteristic equation determines the unforced response of the loop, let us derive the response of the closed loop to a change in process flow by inverting the **Laplace** transform of the output signal, as we learned to do in Chapter 2. Assume that the characteristic equation can be expressed as an nth-degree polynomial in the **Laplace** transform variable s.

$$1 + H(s)G_s(s)G_v(s)G_c(s) = a_n s^n + a_{n-1} s^{n-1} + \ldots + a, = 0 \qquad (6\text{-}1.13)$$

where a,, a_{n-1}, . . . ,a, are the polynomial coefficients. With an appropriate computer program, we can find the n roots of this polynomial and factor it as follows:

$$1 + H(s)G_s(s)G_v(s)G_c(s) = a_n(s - r_1)(s - r_2) \ldots (s - r,) \qquad (6\text{-}1.14)$$

where r_1, r_2 . . . , r_n are the roots of the characteristic equation. These roots can be real numbers or pairs of complex conjugate numbers, and some of them may be repeated, as we saw in Chapter 2.

From Eq. 6-1.7 we obtain

$$T_o(s) = \frac{G_w(s)}{1 + H(s)G_s(s)G_v(s)G_c(s)} W(s) \qquad (6\text{-}1.15)$$

Next let us substitute Eq. 6-1.14 for the denominator and assume that other terms will appear because of the input forcing function, W(s).

$$T_o(s) = \frac{\text{numerator terms}}{a_n(s - r_1) \ldots (s \quad r_n)(\text{input terms})} \qquad \textbf{(6-1.16)}$$

We then expand this expression into partial fractions.

$$T_o(s) = \frac{b_1}{s - r_1} + \frac{b_2}{s - r_2} + \cdots + \frac{b_n}{s - r_n} + (\text{input terms}) \qquad \textbf{(6-1.17)}$$

where $b_1, b_2, \ldots, b_n$ are the constant coefficients that are determined by the method of partial fractions expansion (see Chapter 2). Inverting this expression with the help of a Laplace transform table (such as Table 2-1.1), we obtain

$$T_o(t) = b_1 e^{r_1 t} + b_2 e^{r_2 t} + \ldots + b_n e^{r_n t} + (\text{input terms}) \qquad \textbf{(6-1.18)}$$
$$\underbrace{\qquad\qquad\qquad\qquad\qquad\qquad}_{\text{UNFORCED RESPONSE}} \qquad \underbrace{\qquad\qquad}_{\text{FORCED RESPONSE}}$$

We have thus shown that each of the terms of the unforced response contains a root of the characteristic equation. We recall that the coefficients $b, b_2, \ldots, b_n$ depend on the actual input forcing function and so does the exact response of the loop. However, the speed with which the unforced response terms die out $(r_i < 0)$, diverge $(r_i > 0)$, or oscillate $(r_i$ complex) is determined entirely by the roots of the characteristic equation. We will use this concept in the next section to determine the stability of the loop.

The following two examples illustrate the effect of a pure proportional and a pure integral controller on the closed-loop response of a first-order process. We will see that the pure proportional controller speeds up the first-order response and results in an **offset** or steady-state error, as discussed in Section 5-3. On the other hand, the integral controller produces a second-order response that, as the controller gain increases, changes from overdamped to underdamped. We noted in Chapter 2, the underdamped response is oscillatory.

EXAMPLE 6-1.2   **PROPORTIONAL CONTROL OF A FIRST-ORDER PROCESS**

In the simplified block diagram of Fig. 6-1.4, the process can be represented by a first-order lag.

$$G_1(s) = \frac{K}{\tau s + 1}$$

Determine the closed-loop transfer function and the response to a unit step change in set point for a proportional controller:

$$G_c(s) = K_c$$

SOLUTION

By block diagram algebra, we obtain the closed-loop transfer function

$$\frac{C(s)}{R(s)} = \frac{G_1(s)G_c(s)}{1 + G_1(s)G_c(s)}$$

Then we substitute the process and controller transfer functions and simplify.

$$\frac{C(s)}{R(s)} = \frac{KK_c}{1 + KK_c + \tau s} = \frac{\dfrac{KK_c}{1 + KK_c}}{\dfrac{\tau}{1 + KK_c}s + 1} = \frac{K'}{\tau's + 1}$$

We can easily see that the closed-loop response is first-order with steady-state gain of

$$K' = \frac{KK_c}{1 + KK_c}$$

and time constant of

$$\tau' = \frac{\tau}{1 + KK_c}$$

Note that the closed-loop gain is always less than unity and that the closed-loop time constant is always less than the open-loop time constant τ. In other words, the closed-loop system responds faster than the open-loop system but does not quite match the set point at steady state; that is, there will be *offset*.

Figure 6-1.8 shows the closed-loop unit step responses for several positive values of the loop gain, KK_c. These responses are typical first-order (see Chapter 2). The response
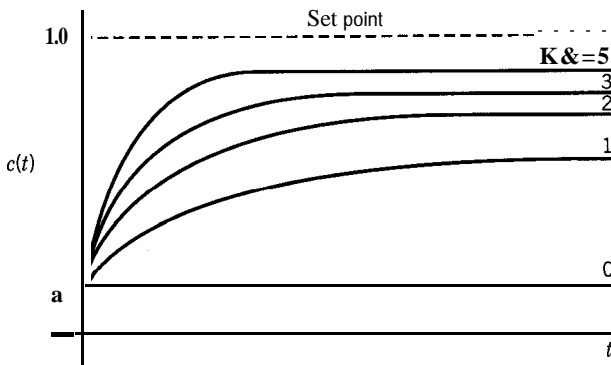
Figure 6-1.8 Unit step response to set point for closed loop of first-order system with proportional controller (Example 6-1.2).

approaches the set point as the loop gain increases. This verifies the statement in Section 5-3 that the offset decreases when the controller gain is increased.

What would be the response if the loop gain, KK_c, were negative? You can easily verify that for loop gains between 0 and -1, the response is stable, but the offset is greater than if no control action is taken at all ($K_c = 0$). You can also verify that for loop gains less than -1, the response is unstable. In contrast, positive loop gains result in a stable response with decreasing *offset*.

EXAMPLE 6-1.3   *PURE INTEGRAL CONTROL OF A FIRST-ORDER PROCESS*

Determine the closed-loop transfer function and the response of a unit step change in set point for the process of Example 6-1.2, and a pure integral controller:

$$G_c(s) = \frac{K_c}{\tau_I} \frac{1}{s} = \frac{K_I}{s}$$

where $K_I = K_c/\tau_I$ is the controller integral gain in min^{-1}.

SOLUTION

Substitute the integral controller transfer function into the closed-loop transfer function of Example 6-1.2.

$$\frac{C(s)}{R(s)} = \frac{\dfrac{KK_I}{s}}{1 + \tau s + \dfrac{KK_I}{s}} = \frac{KK_I}{\tau s^2 + s + KK_I}$$

By the extension of the final value theorem to transfer functions (see Section 3-3), we substitute $s = 0$ to obtain the steady-state gain.

$$\lim_{s \to 0} \frac{C(s)}{R(s)} = \frac{KK_I}{KK_I} = 1.0$$

This means that for the integral controller, the controlled variable will always match the set point at steady state; that is, there will not be an offset.

The characteristic equation of the loop is

$$\tau s^2 + s + KK_I = 0$$

The roots of this quadratic equation are

$$r_{1,2} = \frac{-1 \pm \sqrt{1 - 4KK_I\tau}}{2\tau}$$
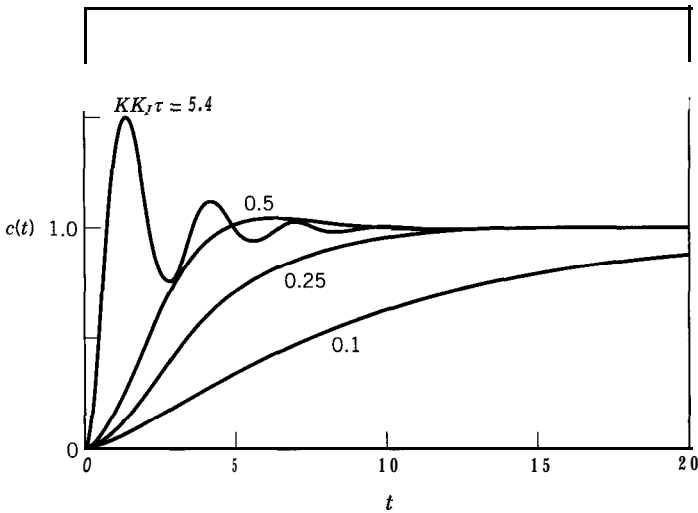
Figure 6-1.9 Unit step response to set point for closed loop of first-order system with integral controller (Example 6-1.3).

These roots are real for $0 \leq KK_I\tau \leq \frac{1}{4}$ and complex conjugates for $KK_I\tau > \frac{1}{4}$. As we saw in Chapter 2, when the roots are real, the response is overdamped, and when the roots are complex conjugates, the response is underdamped (oscillatory). This means that for the loop considered here, the response becomes oscillatory when the loop gain increases. This property is common to most feedback loops.

Figure 6-1.9 shows the closed-loop unit step responses for several positive values of the loop gain. By comparing the characteristic equation of the closed loop with that for the standard underdamped second-order system (see Chapter 2), we can calculate the damping ratio and the frequency of oscillation as a function of the loop parameters. They are

$$\zeta = \sqrt{\frac{1}{4KK_I\tau}} \qquad \omega = \frac{\sqrt{4KK_I\tau - 1}}{2\tau}$$

Table 6-1.2 gives the values of the damping ratio and the frequency of oscillation for several positive values of the loop gain that correspond to the values for the responses of Fig. 6-1.9. This demonstrates how the adjustable controller gain shapes the response of the closed loop.

Table 6-1.2 Underdamped Response for Example 6-1.3

$KK_I\tau$	ζ	$\omega\tau$	Comments
0.25	1.0	0	Critically damped
0.50	0.707	0.5	5% overshoot
5.40	0.215	2.3	Quarter decay ratio

It can be readily shown that for this loop, any negative value of the loop gain KK_s results in two real roots, one of which is positive. This means that the response will exponentially run away with time. In the other hand, for positive values of the loop gain, the roots are either negative real numbers or complex numbers with negative real roots. This, coupled with the unity gain, means that the response always converges to the set point when the loop gain is positive.

The preceding example illustrates the point (discussed in Chapter 4) that although most processes are inherently overdamped, their response can be underdamped when forming part of a closed feedback control loop.

EXAMPLE 6-1.4   *FLOW CONTROL LOOP*

As we shall see in Chapters 10 and 12, flow control loops are commonly used as the innermost loop in cascade, ratio, and feedforward control systems. Develop the closed-loop transfer function for a flow control loop with a proportional-integral (PI) controller.

SOLUTION

Figure 6-1.10 shows a schematic diagram of a flow control loop and its corresponding block diagram. To concentrate on the response of the flow F(s) to its set point $F^{set}(s)$, we will assume constant pressure drop across the control valve. However, one of the purposes of the flow controller is to compensate for changes in the pressure drop across the valve (disturbance). Note that the flow control loop does not have a process! This is because the controlled variable, the flow, is the output of the control valve.
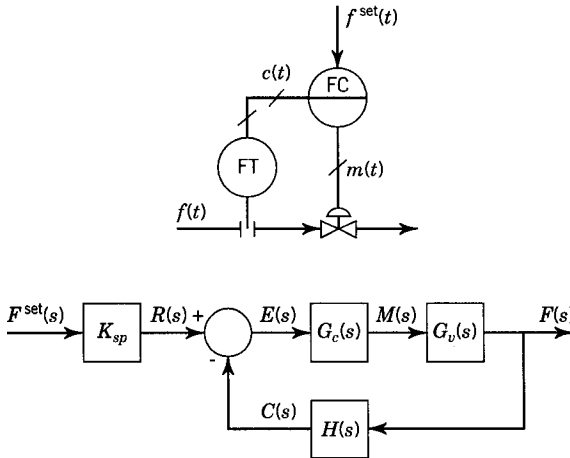
Figure **6-1.10** Schematic and block diagram of a flow control loop.

As we saw in Section 5-2, the control valve can be represented by a first-order lag.

$$G_v(s) = \frac{F(s)}{M(s)} = \frac{K_v}{\tau_v s + 1} \frac{\text{gpm}}{\%\text{CO}}$$

Flow transmitters are usually fast and can thus be represented by just a gain. Assuming a linear transmitter, the gain is, from Section 5-1,

$$H(s) = K_T = \frac{100}{f_{\max}} \frac{\%\text{TO}}{\text{gpm}}$$

We apply block diagram algebra (Chapter 3) to the diagram of Fig. 6-1.10 to obtain the transfer function of the closed loop.

$$\frac{F(s)}{F^{\text{set}}(s)} \quad \frac{K_{sp} G_v(s) G_c(s)}{1 + K_T G_v(s) G_c(s)}$$

where $K_{sp} = K_T$. From Section 5-3, the transfer function of the PI controller is

$$G_c(s) = K_c\left(1 + \frac{1}{\tau_I s}\right) = \frac{K_c(\tau_I s + 1)}{\tau_I s} \frac{\%\text{CO}}{\%\text{TO}}$$

We substitute into the closed-loop transfer function and simplify to obtain

$$\frac{F(s)}{F^{\text{set}}(s)} \quad \frac{K_T K_v K_c(\tau_I s + 1)}{\tau_I s(\tau_v s + 1) + K_T K_v K_c(\tau_I s + 1)}$$

The response is second-order. It can be underdamped (oscillatory) or overdamped, depending on the controller parameters. A fast first-order response can be obtained by setting the integral time equal to the valve time constant, $\tau_I = \tau_v$.

$$\frac{F(s)}{F^{\text{set}}(s)} = \frac{K_T K_v K_c}{\tau_v s + K_T K_v K_c} = \frac{1}{\tau_{FC} s + 1}$$

where the closed-loop time constant is

$$\tau_{FC} = \frac{\tau_v}{K_T K_v K_c}$$

Note that the closed-loop response is faster (shorter time constant) as the controller gain increases and that the steady-state gain is unity; that is, there is no offset.

When the flow control loop is part of a cascade control system, its set point is sometimes in percent of range instead of in engineering units (gpm). In such cases, the

input to the loop is R(s) instead of $F^{set}(s)$ (see Fig. 6-1.10). The transfer function is then

$$\frac{F(s)}{R(s)} = \frac{K_v K_c}{\tau_v\, s + K_T K_v K_c} = \frac{1/K_T}{\tau_{FC} s + 1}\, \frac{\text{gpm}}{\%}$$

and the gain of the flow control loop is

$$K_{FC} = \frac{1}{K_T} = \frac{f_{max}}{100}\, \frac{\text{gpm}}{\%}$$

Note that this is very similar to the gain of a linear valve with constant pressure drop, except that the maximum flow here is the upper limit of the flow transmitter range (see Section 5-2).

The formulas derived in this example apply to liquid, gas, and steam valves, with the units appropriately adjusted (e.g., gpm, scfm, lb/h). They are independent of the flow characteristics of the valve and of whether the pressure drop is constant or variable.

6-1.3 Steady-State Closed-Loop Gains

We have seen in the preceding examples that the final or steady-state value is an important aspect of the closed-loop response. This is because in industrial process control practice, the presence of steady-state error, or *offset,* is usually unacceptable. We shall learn in this section how to calculate the offset when it is present. To do this, we return to the exchanger of Fig. 6-1.1 and the corresponding block diagram of Fig. 6-1.3. As we learned earlier, this is a linearized representation of the heat exchanger. Our approach is to obtain the steady-state closed-loop relationships between the output variable and each of the inputs to the loop by applying the final value theorem to the closed-loop transfer function. From Eq. 6-1.7, the closed-loop transfer function between the outlet temperature and the process fluid flow is

$$\frac{T_o(s)}{W(s)} = \frac{G_w(s)}{1 + H(s)G_s(s)G_v(s)G_c(s)} \tag{6-1.7}$$

We recall that this expression assumes that the deviation variables for the inlet temperature T_i and the set point T_o^{set} are zero when these other inputs remain constant. We also recall, from Section 3-3, that the steady-state relationship between the output and the input to a transfer function is obtained by setting $s = 0$ in the transfer function. This follows from the final value theorem of Laplace transforms. Applying this method to Eq. 6-1.7, we obtain

$$\frac{\Delta T_o}{\Delta W} = \frac{G_w(0)}{1 + H(0)G_s(0)G_v(0)G_c(0)} \tag{6-1.19}$$

where

ΔT_o = the steady-state change in outlet temperature, °C
ΔW = the steady-state change in process fluid flow, kg/s

If we assume, as is usually the case, that the process is stable, then

$G_s(O) = K_w$, the process open-loop gain to a change in process fluid flow, °C/(kg/s)

$G_s(O) = K_s$, the process open-loop gain to a change in steam flow, °C/(kg/s)

Similarly, for the valve and the sensor/transmitter,

$G_v(O) = K_v$, the valve gain, (kg/s)/%CO

$H(O) = K_T$, the sensor/transmitter gain, %TO/%CO

Finally, if the controller does not have integral mode, then

$G_c(O) = K_c$, the proportional gain, %CO/%TO

Substituting these terms into Eq. 6-1.19 yields

$$\frac{\Delta T_o}{\Delta W} = \frac{K_w}{1 + KK_c} \frac{°C}{kg/s} \qquad (6\text{-}1.20)$$

where $K = K_T K_s K_v$ is the combined gain of the elements of the loop other than the controller, %TO/%CO. Because the change in set point is zero, the steady-state error, or offset. is

$$e = \Delta T_o^{set} - \Delta T_o = -\Delta T_o \; °C$$

and combining this relationship with Eq. 6- 1.20 gives

$$\boxed{\frac{e}{\Delta W} = \frac{-K_w}{1 + KK_c} \frac{°C}{kg/s}} \qquad (6\text{-}1.21)$$

Note that the offset decreases as the controller gain, K_c, is increased.

Following an identical procedure for Eq. 6- 1.8, we obtain the steady-state relationship to a change in set point at constant process fluid flow.

$$\frac{\Delta T_o}{\Delta T_o^{set}} = \frac{K_s K_v K_c K_{sp}}{1 + K_T K_s K_v K_c} = \frac{KK_c}{1 + KK_c} \; °C/°C \qquad (6\text{-}1.22)$$

where ΔT_o^{set} is the steady-state change in set point, "C, and we have used $K_{sp} = K_T$. The offset in this case is

$$e = \Delta T_o^{set} - \Delta T_o \; °C$$

Combining this relationship with Eq. 6-1.22 gives

$$\boxed{\frac{e}{\Delta T_o^{set}} = \frac{1}{1 + K_T K_s K_v K_c} = \frac{1}{1 + KK_c} \; °C/°C} \qquad (6\text{-}1.23)$$

Again the offset is smaller the higher the controller gain.

Effect of Integral Mode

For a proportional-integral-derivative (PID) controller,

$$G_c(0) = \lim_{s \to 0} K_c \left(1 + \frac{1}{\tau_I s} + \tau_D s \right) = \infty$$

In this case, by substitution into Eq. 6-1.21 or 6-1.23, in place of K_c, we can see that the offset is zero. The same is true for a PI controller $(\tau_D = 0)$.

EXAMPLE 6-1.5

For the heat exchanger of Fig. 6-1.1, calculate the linearized ratios for the steady-state error in outlet temperature to

(a) A change in process flow.
(b) A change in set point.

The operating conditions and instrument specifications are

Process fluid flow	$w = 12$ kg/s
Inlet temperature	$T_i = 50°C$
Set point	$T_o^{set} = 90°C$
Heat capacity of fluid	$c_p = 3.15$ kJ/kg-°C
Latent heat of steam	$\lambda = 2250$ kJ/kg
Capacity of steam valve	$w_{s,max} = 1.6$ kg/s
Transmitter range	50 to 150°C

SOLUTION

If we assume that heat losses are negligible, then we can write the following steady-state energy balance:

$$\overline{w} c_p (\overline{T}_o - T_i) = \overline{w}_s \lambda$$

and solving for $\overline{w}_s$, we find that the steam flow required to maintain T_o at 90°C is

$$\overline{w}_s = \frac{\overline{w} c_p (\overline{T} - T_i)}{A} = \frac{(12)(3.75)(90 - 50)}{2250} = 0.80 \text{ kg/s}$$

The next step is to calculate the steady-state open-loop gains of each of the elements in the loop.

Exchanger.   From the steady-state energy balance, solving for T_o yields

$$\overline{T}_o = T_i + \frac{w_s \lambda}{w c_p}$$

By linearization (see Section 2-6), we obtain

$$K_w = \frac{\partial T_o}{\partial w} = -\frac{\overline{w}_s \lambda}{\overline{w}^2 c_p} = -\frac{(0.80)(2250)}{(12)^2(3.75)} = -3.33 \frac{°C}{kg/s}$$

$$K_s = \frac{\partial T_o}{\partial w_s} = \frac{\lambda}{\overline{w} c_p} = \frac{2250}{(12)(3.75)} = 50 \frac{°C}{kg/s}$$

Control Valve. Assuming a linear valve with constant pressure drop, the gain of the valve is, from Section 5-2,

$$K_v = \frac{w_{s,max}}{100} = \frac{1.6}{100} = 0.016 \frac{kg/s}{\%CO}$$

Sensor/Transmitter. From Section 5-1, the gain of the transmitter is

$$K_T = \frac{100 \quad 0}{150 - 50} = 1.0 \frac{\%TO}{°C}$$

Then

$$K = K_T K_s K_v = (1.0)(50)(0.016) = 0.80 \frac{\%TO}{\%CO}$$

(a) Substitute into Eq. 6-1.21 to get

$$\frac{e}{Aw} = \frac{-K_w}{1 + KK_c} = \frac{3.33}{1 + 0.80K_c} \frac{°C}{kg/s}$$

(b) Substitute into Eq. 6-1.23 to get

$$\frac{e}{\Delta T_o^{set}} = \frac{1}{1 + KK_c} = \frac{1}{1 + 0.80K_c} °C/°C$$

The results for different values of K_c are given in Table 6-1.3. We see that the offset in outlet temperature approaches zero as the gain is increased. These results illustrate the point made in Section 5-3 that the offset decreases when the gain of the proportional controller is increased. As we noted there, the gain of the controller is limited by the stability of the loop. We shall see this in the next section.

Table 6-1.3 Offset for Heat Exchanger
Control Loop

$K_c,$ $\dfrac{\%CO}{\%TO}$	$\dfrac{e}{\Delta W'},$ $\dfrac{°C}{kg/s}$	$\dfrac{e}{\Delta T_o^{set}},$ $°C/°C$
0	3.33	1.00
0.5	2.38	0.714
1.0	1.85	0.556
5.0	0.67	0.200
10.0	0.37	0.111
20.0	0.20	0.059
100.0	0.04	0.012

6-2 STABILITY OF THE CONTROL LOOP

As defined in Chapter 2, a system is *stable* if its output remains bound for a bound input. Most industrial processes are open-loop stable; that is, they are stable when not a part of a feedback control loop. This is equivalent to saying that most processes are self-regulating, that is, the output moves from one steady state to another when driven by changes in its input signals. As we learned in Chapter 4, a typical example of an open-loop unstable process is an exothermic stirred tank reactor.

Even for open-loop stable processes, stability becomes a consideration when the process is a part of a feedback control loop. This is because the signal variations may reinforce each other as they travel around the loop, causing the output-and all the other signals in the loop-to become unbounded. As we noted in Chapter 1, the behavior of a feedback control loop is essentially oscillatory-"trial and error." Under some circumstances the oscillations may increase in magnitude, resulting in an unstable process. A good illustration of an unstable feedback loop is the controller whose direction of action is the opposite of what it should be. For example, in the heat exchanger sketched in the preceding section, if the controller output were to increase with increasing temperature (direct-acting controller), then the loop would be unstable because the opening of the steam valve would cause a further increase in temperature. What is needed in this case is a reverse-acting controller that decreases its output when the temperature increases, so as to close the steam valve and bring the temperature back down. However, even for a controller with the proper action, the system may become unstable because of the lags in the loop. This usually happens as the loop gain is increased. The controller gain at which the loop reaches the threshold of instability is therefore of utmost importance in the design of a feedback control loop. This maximum gain is known as the *ultimate gain.*

In this section, we will determine a criterion for the stability of dynamic systems and study two methods used to calculate the ultimate gain: direct substitution and Routh's test. Then we will study the effect of various loop parameters on its stability.

6-2.1 Criterion of Stability

We saw earlier that the response of a control loop to a given input can be represented (Eq. 6-1.18) by

$$C(t) = b_1 e^{r_1 t} + b_2 e^{r_2 t} + \ldots + b_n e^{r_n t} + (\text{input terms}) \qquad (6\text{-}2.1)$$

where $C(t)$ is the controlled variable and r_1, r_2, ..., r_n **are** the roots of the characteristic equation of the loop.

Assuming that the input terms remain bounded as time increases, the stability of the loop requires that the unforced response terms also remain bounded as time increases. This depends only on the roots of the characteristic equation and can be expressed as follows:

> For real roots: If $r < 0$, then $e^{rt} \to 0$ as $t \to \infty$.
> For complex roots: $\tau = \sigma + i\omega$ $e^{rt} = e^{\sigma t} \sin(\omega t + \theta)$
>   If $\sigma < 0$, then $e^{\sigma t} \sin(\omega t + \theta) \to 0$ as $t \to \infty$.

In other words, the real part of the complex roots and the real roots must be negative in order for the corresponding terms in the response to decay to zero. This result is not affected by repeated roots, because this only introduces a polynomial of time into the solution, which cannot overcome the effect of the decaying exponential term (see Chapter 2). Note that if any root of the characteristic equation is a positive real number, or a complex number with a positive real part, then that term on the response (Eq. 6-2.1) will be unbounded, and the entire response will be unbounded even though all the other terms may decay to zero. This brings us to the following statement of the criterion for the stability of a control loop:

> *For a feedback control loop to be stable, all of the roots of its characteristic equation must be either negative real numbers or complex numbers with negative real parts.*

If we now define the complex plane or s plane as a two-dimensional graph with the horizontal axis for the real parts of the roots and the vertical axis for the imaginary parts, we can make the following graphical statement of the criterion of stability (see Fig 6-2.1):

> *For a feedback control loop to be stable, all the roots of its characteristic equation must fall on the left-hand half of the s plane, also known as the left-hand plane.*

We must point out that both of these statements of the stability criterion in the Laplace domain apply in general to any physical system, not just to feedback control loops. In each case, we obtain the characteristic equation by setting the denominator of the linearized transfer function of the system equal to zero.

Having articulated the criterion of stability, let us turn our attention to determining the stability of a control loop.

6-2.2 Direct Substitution Method

Direct substitution is a convenient method for determining the range of controller parameters for which the closed-loop response is stable. The method is based on the fact that the roots of the characteristic equation vary continuously with the loop parameters. Consequently, at the point of instability, at least one and usually two of the roots must
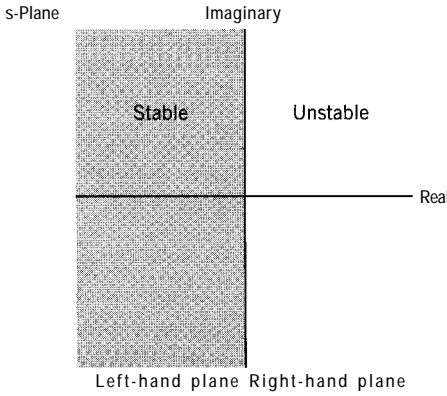
Figure 6-2.1 The s-plane, showing the regions of stability and instability for the location of the roots of the characteristic equation.

lie on the imaginary axis of the complex plane as they cross from the left half-plane to the right. This means that the roots are pure imaginary numbers-zero real parts-at the verge of instability. At this point the loop is said to be marginally stable, and the corresponding term on the loop output is, in the **Laplace** domain,

$$C(s) = \frac{b_1 s + b_2}{s^2 + \omega_u^2} + \text{(other terms)} \tag{6-2.2}$$

or, upon inverting, this term, from Table 2-1.1, is a sine wave in the time domain:

$$C(t) = b_1' \sin(\omega_u t + \theta) + \text{(other terms)} \tag{6-2.3}$$

where ω_n is the frequency of sine wave, θ is its phase angle, and b_1' is its amplitude (constant). This means that at the point of marginal stability, the characteristic equation must have a pair of pure imaginary roots at

$$r_{1,2} = \pm i\omega_u$$

The frequency ω_u with which the loop oscillates is *the ultimate frequency*. The controller gain at which this point of marginal instability is reached is called the *ultimate gain*. At a gain just below the ultimate, the loop oscillates with a decaying amplitude, whereas at a gain just above the ultimate gain, the amplitude of the oscillations increases with time. At the point of marginal stability, the amplitude of the oscillation remains constant with time. Figure 6-2.2 shows these responses, along with the graphical representation of the *ultimate period, T_u*. This is the period of the oscillations at the ultimate gain, and it is related to the ultimate frequency, ω_u, rad/s, by

$$T_u = \frac{2\pi}{\omega_u} \tag{6-2.4}$$
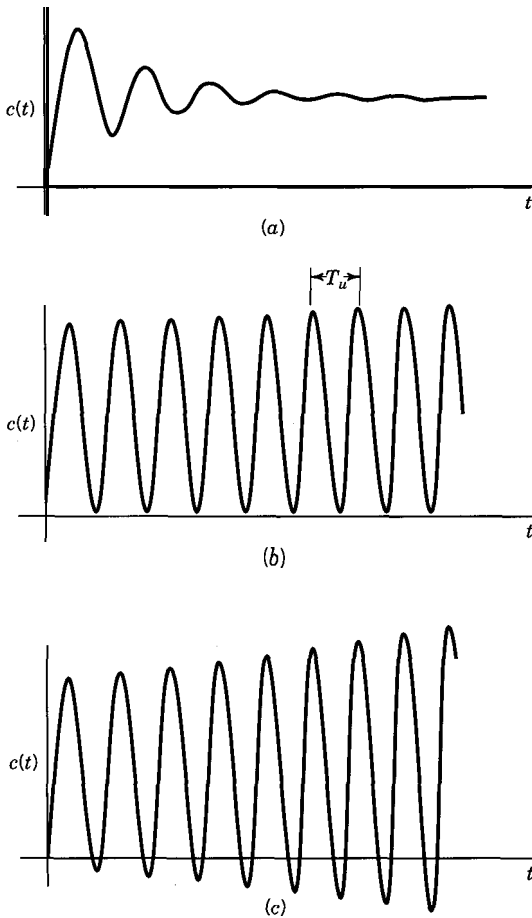
Figure 6-2.2 Response of closed loop with the controller gain less than (a), equal to (b), and greater than (c) the ultimate gain.

The method of direct substitution consists of substituting $s = i\omega_u$ in the characteristic equation. This results in a complex equation that can be converted into two simultaneous equations:

$$\text{Real part} = 0$$

$$\text{Imaginary part} = 0$$

From these we can solve for two unknowns. One is the ultimate frequency ω_u, and the other is any of the parameters of the loop, usually the controller gain at the point of marginal instability or ultimate gain. Generally, the closed-loop response is unstable when the controller gain is greater than the ultimate gain. The following example shows how the direct substitution method is used to compute the ultimate gain and period of the heat exchanger control loop.

EXAMPLE 6-2.1   *ULTIMATE GAIN AND PERIOD OF TEMPERATURE CONTROLLER BY DIRECT SUBSTITUTION*

Let us assume that the transfer functions for the various elements of the temperature control loop of Fig. 6-1.3 are as follows:

Exchanger. The exchanger response to the steam flow has a gain of $50°C/(kg/s)$ and a time constant of 30 s.

$$G_s(s) = \frac{50}{30s + 1} \frac{°C}{kg/s}$$

Sensor/Transmitter. The sensor/transmitter has a calibrated range of 50 to 150°C and a time constant of 10s.

$$H(s) = \frac{1.0}{10s + 1} \frac{\% \, TO}{°C}$$

Control Valve. The control valve has a maximum capacity of 1.6 kg/s of steam, linear characteristics, a constant pressure drop, and a time constant of 3 s.

$$G_v(s) = \frac{0.016}{3s + 1} \frac{kg/s}{\%CO}$$

Controller. The controller is proportional only.

$$G_c(s) = K_c \frac{\%CO}{\%TO}$$

The problem is then to determine the ultimate controller gain (that is, the value of K_c at which the loop becomes marginally stable) and the ultimate period.

SOLUTION

The characteristic equation is given by Eq. 6-1.12:

$$1 + H(s)G_s(s)G_v(s)G_c(s) = 0$$

or, when we substitute the transfer function for each element,

$$1 + \frac{1}{10s + 1} \cdot \frac{50}{30s + 1} \cdot \frac{0.016}{3s + 1} \cdot K_c = 0$$

We must now rearrange this equation into polynomial form.

$$(10s + 1)(30s + 1)(3s + 1) + 0.80K_c = 0$$

$$900s^3 + 420s^2 + 43s + 1 + 0.80K_c = 0$$

Next we substitute $s = i\omega_u$ at $K_c = K_{cu}$.

$$900i^3\omega_u^3 + 420i^2\omega_u^2 + 43i\omega_u + 0.80K_{cu} = 0$$

Then we substitute $i^2 = -1$ and separate the real and imaginary parts.

$$(-420\omega_u^2 + 1 + 0.80K_{cu}) + i(-900\omega_u^3 + 43\omega_u) = 0 + i0$$

From this complex equation, we obtain the following two equations, because both the real and the imaginary parts must be zero.

$$-420\omega_u^2 + 1 + 0.80K_{cu} = 0$$

$$-900\omega_u^3 + 43\omega_u = 0$$

The solution of this set has the following possibilities:

For $\omega_u = 0$, $\qquad\qquad K_{cu} = -1.25$ %CO/%TO

For $\omega_u = 0.2186$ rad/s, $\qquad K_{cu} = 23.8$ %CO/%TO

The first solution corresponds to the monotonic instability caused by having the wrong action on the controller. In this case, the system does not oscillate but moves monotonically in one direction or the other. The crossing of the imaginary axis occurs at the origin $(s = 0)$. This solution is irrelevant.

The ultimate gain for the second solution is the one that is relevant. At this gain, the loop response oscillates with a frequency of 0.2186 rad/s (0.0348 hertz) or a period of

$$T_u = \frac{2\pi}{0.2186} = 28.7 \text{ s}$$

We saw in the preceding section that the offset, or steady-state error, inherent in proportional controllers can be reduced by increasing the controller gain. We see here that stability imposes a limit on how high that gain can be. It is of interest to study how the other parameters of the loop affect the ultimate gain and period.

EXAMPLE 6-2.2 ULTIMATE GAIN AND PERIOD OF CONTINUOUS STIRRED TANK HEATER

Obtain the ultimate gain and period for the temperature control loop of the continuous stirred tank heater of Example 6-1.1, assuming a proportional controller.

SOLUTION

In Example 6-1.1, we obtained the following transfer function for the closed-loop.

$$C(s) = \frac{G_c(s)G_1(s)}{1 + G_c(s)G_1(s)}R(s) + \frac{G_2(s)}{1 + G_c(s)G_1(s)}F(s)$$

where

$$G_1(s) = \frac{1.954}{(0.2s + 1)(8.34s + 1)(0.502s + 1)(0.75s + 1)} \frac{\% \, TO}{\% CO}$$

where $1.954 = 1.652 \cdot 1.183 \cdot 1.0$. The characteristic equation of the loop is

$$1 + G_c(s)G_1(s) = 0$$

Substitute G,(s) and G,(s) $= K_c$, and clear fractions.

$$(0.2s + 1)(8.34s + 1)(0.502s + 1)(0.75s + 1) + 1.954K_c = 0$$

In polynomial form,

$$0.628s^4 + 5.303s^3 + 12.73s^2 + 9.790s + 1 + 1.954K_c = 0$$

Substitute $s = i\omega_u$ at $K_c = K_{cu}$.

$$0.628\omega_u^4 - i5.303\omega_u^3 - 12.734 + i9.790\omega_u + 1 + 1.954K_{cu} = 0$$

Solve for ω_u by setting the imaginary part to zero yields

$$\omega_u = \sqrt{\frac{9.790}{5.303}} = 1.359 \text{ rad/min}$$

Solve for K_{cu} by setting the real part to zero yields

$$K_{cu} = \frac{-0.628(1.359)^4 + 12.73(1.359)^2 - 1}{1.954} = \frac{20.37}{1.954} = 10.4 \frac{\% CO}{\% TO}$$

The ultimate periods is

$$T_u = \frac{2\pi}{\omega_u} = \frac{2\pi}{1.359} = 4.6 \text{ min}$$

In both of the examples given in this section, the controller action has been reverse. The following example shows one way to handle the case of a direct-acting controller.
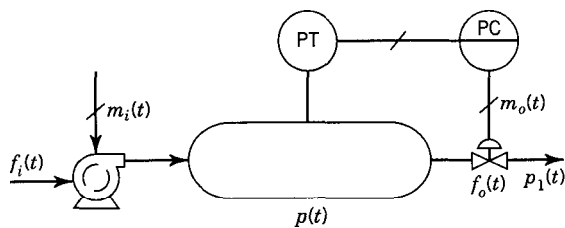
Figure 6-2.3 Pressure control of gas process.

EXAMPLE 6-2.3   *FEEDBACK CONTROL OF PRESSURE IN A GAS VESSEL*

Obtain the ultimate gain and period for the control of the pressure in the gas process of Section 3-5. Figure 6-2.3 is a revised version of Fig. 3-5.1 showing the pressure control loop. The pressure transmitter (PT) has a range of 0 to 40 psig and can be represented by a first-order lag with a time constant of 1.0 s. A first-order lag with a time constant of 3.0 s represents the control valve actuator.

SOLUTION

The block diagram for the gas process, originally given in Fig. 3-5.2, has been expanded in Fig. 6-2.4 to include the pressure controller, the lag in the control valve actuator, and the transmitter. Note that the process action is reverse; an increase in the controller output causes the outlet valve from the tank to open, resulting in a decrease in the pressure in the tank. This is shown in the block diagram by the minus sign on the summer. Because, as we saw in Section 5-3, the controller action must be such as to counteract any change in the pressure, the pressure controller must be *direct-acting;* that is, an increase in pressure must cause the controller output to increase. Because of our sign convention on the block diagram, this means that *the controller gain must be negative.*
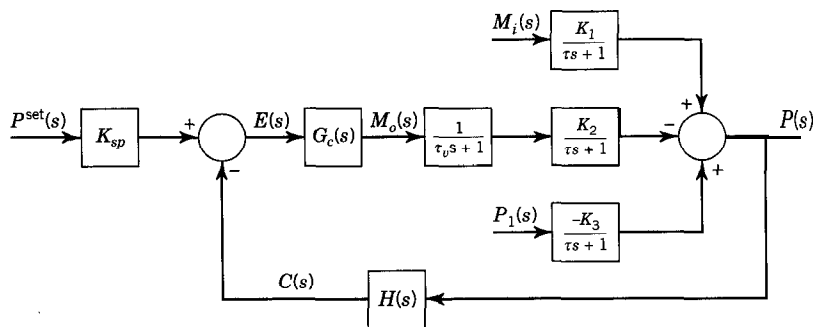
Figure 6-2.4 Block diagram of gas pressure control loop.

The numerical values of the transfer functions are given in Eqs. 3-5.17 through 3-5.19 and are $K_1 = 0.615$ psi/%, $K_2 = 0.619$ psi/%VP, $K_3 = -0.611$ psi/psi, and $\tau = 5.242$ s. The valve actuator time constant is $\tau_v = 3.0$ s, its gain is 1.0 %VP/%CO, and the transmitter transfer function is

$$H(s) = \frac{K_T}{s + 1}$$

where the time constant is 1.0 s, and the gain, from Section 5-2, is $K_T = K_{sp} = (100 \quad 0)/(40 - 0) = 2.5$ %TO/psi. The characteristic equation of the loop is

$$1 + G_c(s) \frac{1}{\tau_v s + 1} \frac{-K_2}{\tau s + 1} \frac{K_T}{\tau_T s + 1} = 0$$

A good way to handle the negative controller gain is to set, for the proportional controller, $G_c(s)$ equal to $-K_c$. This avoids having to remember that the relevant values of the controller gain in this problem are negative. Substituting all the numerical values into the characteristic equation, we obtain

$$1 + \frac{(- \; 1.548)(-K_c)}{(3.0s + 1)(5.242s + 1)(s + 1)} = 0$$

Note that the sign of the sum is positive. ***This is always the case when the correct controller action is used.*** The equation, in polynomial form, is

$$15.726s^3 + 23.968s^2 + 9.242s + 1 + 1.548K_c = 0$$

By direct substitution of $s = i\omega_u$ at $K_c = K_{cu}$, we obtain the following answers:

$$\omega_u = \sqrt{\frac{9.242}{15.726}} = 0.767 \;\; \text{rad/s}$$

$$T_u \quad \frac{2\pi}{= 0.767} = 8.20 \;\; \text{s} \;\; (0.14 \;\; \text{min})$$

$$K_{cu} = \frac{23.968(0.767)^2 - 1}{1.548} = 8.5 \;\; \frac{\%CO}{\%TO}$$

The actual controller gain, however, must be between 0 and $-$ 8.5 %CO/%TO for the loop to be stable.

6-2.3 Effect of Loop Parameters on the Ultimate Gain and Period

Let us assume that the calibrated range of the temperature sensor/transmitter in Example 6-2.1 is reduced to 75 to 125°C. The new transmitter gain is

$$K_T = \frac{100 - 0}{125 - 75} = 2.0 \ \frac{\% \ TO}{°C}$$

The characteristic equation of the loop becomes

$$900s^3 + 420s^2 + 43s + 1 + 1.60K_c = 0$$

and the ultimate gain and period are

$$K_{cu} = 11.9 \ \frac{\%CO}{\% \ TO} \qquad T_u = 28.7s$$

This is exactly half the ultimate gain for the base case, which shows that the ultimate loop gain remains the same. The loop gain is defined as the product of the gains of all the blocks in the loop.

$$\boxed{K_L = K_v K_s K_T K_c = KK_c} \tag{6-2.5}$$

where K_L is the (dimensionless) loop gain. For the two cases considered so far, the ultimate loop gains are

$$K_{Lu} = (1.0)(50)(0.016)(23.8) = 19.04$$

$$K_{Lu} = (2.0)(50)(0.016)(11.9) = 19.04$$

Similarly, if we were to double the capacity of the control valve, and thus its gain, then the ultimate controller gain would be reduced to half its value for the base case.

Next let us assume that a faster sensor/transmitter with a time constant of 5 s is installed in this service, replacing the 10-s instrument. The new transfer function is

$$H(s) = \frac{1.0}{5s + 1} \ \frac{\% \ TO}{°C}$$

The characteristic equation is now

$$1 + \frac{1.0}{5s + 1} \cdot \frac{30}{30s + 1} \cdot \frac{0.016}{3s + 1} \cdot K_c = 0$$

$$450s^3 + 255s^2 + 38s + 1 + 0.80K_c = 0$$

and the ultimate gain, frequency, and period are

$$\omega_u = \sqrt{\frac{38}{450}} = 0.2906 \ \text{rad/s}$$

$$T_u = 21.6 \ \text{s}$$

$$K_{cu} = 25.7 \ \frac{\%CO}{\%TO} \qquad (PB, = 4.0)$$

The reduction of the time constant of the sensor has resulted in a slight increase in the ultimate gain and a decrease in the period of oscillation of the loop. This is because we have reduced the measurement lag on the control loop. A similar result would be obtained if the time constant of the control valve were reduced. However, the increase in the ultimate gain would be even less, because the valve is not as slow as the sensor/transmitter. You are invited to verify this.

Finally, let us consider a case in which a change in exchanger design results in a shorter time constant for the process, namely from 30 to 20 s. The new transfer function is

$$G_,(s) = \frac{50}{20s + 1} \ \frac{°C}{kg/s}$$

The characteristic equation is then

$$1 + \frac{1.0}{10s + 1} \cdot \frac{50}{20s + 1} \cdot \frac{0.016}{3s + 1} \cdot K_c = 0$$

$$600s^3 + 290s^2 + 33s + 1 + 0.80K_c = 0$$

and the ultimate frequency, gain and, period are

$$\omega_u = \sqrt{\frac{33}{600}} = 0.2345 \ \text{rad/s}$$

$$T_u = 26.8 \ \text{s}$$

$$K_{cu} = 18.7 \ \frac{\%CO}{\% TO} \qquad (PB, = 5.4)$$

Surprisingly, the ultimate gain is reduced by a reduction in the process time constant. This is opposite to the effect of reducing the time constant of the sensor/transmitter. The reason is that when the longest or dominant time constant is reduced, the relative effect of the other lags in the loop becomes more pronounced. In other words, in terms of the ultimate gain, reducing the longest time constant is equivalent to proportionately increasing the other time constants in the loop. However, the loop with the shorter process time constant responds faster than the original one, as shown by the shorter

Table 6-2.1 Direct Substitution Results for Heat Exchanger
Control Loop

	K_{cu}	ω_u, rad/s	T_u, s
1. Base case	23.8	0.2186	28.7
2. $H(s) = \dfrac{2.0}{10s + 1}$	11.9	0.2186	28.7
3. $H(s) = \dfrac{1.0}{5s + 1}$	25.7	0.2906	21.6
4. $G_s(s) = \dfrac{50}{20s + 1}$	18.7	0.2345	26.8

ultimate period. The results of the direct substitution method for the other cases con-
sidered here are summarized in Table 6-2.1. We note that the loop can oscillate signif-
icantly faster when the time constant of the sensor/transmitter is reduced from 10 to
5 s. Also, the loop oscillates slightly faster when the exchanger time constant is reduced
from 30 to 20 s, in spite of the significant reduction in ultimate gain. Changing the
gains of the blocks on the loop has no effect on the frequency of oscillation or on the
ultimate loop gain.

6-2.4 Effect of Dead Time

We have seen how the direct substitution method allows us to study the effect of various
loop parameters on the stability of the feedback control loop. Unfortunately, the method
fails when any of the blocks on the loop contains a dead-time (transportation lag or
time delay) term. This is because the dead time introduces an exponential function of
the Laplace transform variable into the characteristic equation. This means that this
equation is no longer a polynomial, and the methods we have learned in this section
no longer apply. An increase in dead time tends to reduce the ultimate loop gain very
rapidly. This effect is similar to the effect of increasing the nondominant time constants
of the loop in that it is relative to the magnitude of the dominant time constant. We
will study the exact effect of dead time on loop stability when we consider the method
of frequency response in Chapter 9.

We must point out that the exchanger we have used in this chapter is a distributed-
parameter system; that is, the temperature of the process fluid is distributed throughout
the exchanger. The transfer functions for such systems usually contain a least one dead-
time term, which, for simplicity, we have ignored.

An estimate of the ultimate gain and frequency of a loop with dead time may some-
times be obtained by using an approximation to the dead-time transfer function. A
popular approximation is the first-order Padé approximation, which is given by

$$e^{-t_0 s} = \frac{1 - \dfrac{t_0}{2}s}{1 + \dfrac{t_0}{2}s} \tag{6-2.6}$$

where t_0 is the dead time. More accurate higher-order approximations are also available, but they are too complex to be practical. The following example illustrates the use of the Padé approximation with the direct substitution method.

EXAMPLE 6-2.4   ULTIMATE GAIN AND FREQUENCY OF FIRST-ORDER PLUS DEAD-TIME PROCESS

Let the process transfer function of the loop of Fig. 6-1.4 be

$$G_1(s) = \frac{Ke^{-t_0 s}}{\tau s + 1}$$

where K is the gain, t_0 is the dead time, and τ is the time constant. Determine the ultimate gain and frequency of the loop as a function of the process parameters if the controller is a proportional controller:

$$G_c(s) = K_c$$

SOLUTION

From Example 6-1.2, the characteristic equation of the loop is

$$1 + G_c(s)G_1(s) = 0$$

or, for the transfer functions considered here,

$$1 + \frac{KK_c e^{-t_0 s}}{\tau s + 1} = 0$$

Substitute the first-order Padt approximation, Eq. 6-2.6, to get

$$1 + \frac{KK_c\left(1 - \dfrac{t_0}{2}s\right)}{(\tau s + 1)\left(1 + \dfrac{t_0}{2}s\right)} = 0$$

Clear the fraction.

$$\frac{t_0}{2}\tau s^2 + \left(\tau + \frac{t_0}{2} - KK_c\frac{t_0}{2}\right)s + 1 + KK_c = 0$$

The direct substitution method, $s = i\omega_u$ at $K_c = K_{cu}$, yields

$$\frac{t_0}{2} \tau i^2 \omega_u^2 + \left(\tau + \frac{t_0}{2} - KK_{cu} \frac{t_0}{2} \right) i\omega_u + 1 + KK_{cu} = 0$$

$$\left(-\frac{t_0}{2} \tau \omega_u^2 + 1 + KK_{cu} \right) + i \left(\tau + \frac{t_0}{2} - KK_{cu} \frac{t_0}{2} \right) \omega_u = 0$$

After setting the real and imaginary parts equal to zero and solving the two equations simultaneously, we find that the solution is

$$KK_{cu} = 1 + 2 \frac{\tau}{0 t_0}$$

$$\omega_u = \frac{2}{t0} \sqrt{\frac{t_0}{\tau} + 1}$$

These formulas show that the ultimate loop gain goes to infinity-with no stability limit-as the dead time approaches zero, which agrees with the results of Example 6-1.2. However, any finite amount of dead time imposes a stability limit on the loop gain. The ultimate frequency increases with decreasing dead time and becomes very small as the dead time increases. This means that dead time slows the response of the loop.

Summary

The following general effects of various loop parameters emerge from the results of direct substitution analysis in the preceding examples.

- Stability imposes a limit on the overall loop gain, so that an increase in the gain of the control valve, the transmitter, or the process results in a decrease in the ultimate controller gain.
- An increase in dead time or in any of the nondominant (smaller) time constants of the loop results in a reduction of the ultimate gain.
- A decrease in the dominant (longest) time constant of the loop results in a decrease in the ultimate loop gain and an increase in the ultimate frequency of the loop.

6-2.5 Routh's Test

Routh's test is a procedure that enables us to determine how many of the roots of a polynomial have positive real parts without actually finding the roots by iterative techniques. Because the stability of a system requires that none of the roots of its characteristic equation have positive real parts, Routh's test is useful to determine stability.

With today's availability of computer and calculator programs to solve for the roots of polynomials, Routh's test would not be useful if the problem were merely to find out whether a feedback loop is stable once all the parameters of the loop have been

specified. However, the more relevant problem is to find the limits on a given loop parameter-usually the controller gain-for which the loop is stable. Routh's test is useful for solving this problem.

The mechanics of Routh's test can be presented as follows: Given the nth-degree polynomial

$$a_n s^n + a_{n-1} s^{n-1} + \ldots + a_1 s + a, = 0 \tag{6-2.7}$$

where $a,, a_{n-1}, \ldots, a_1, a,$ are the coefficients of the polynomial, determine how many roots have positive real parts.

To perform the test, we must first prepare the following array:

Row 1	a_n	a_{n-2}	a_{n-4}	. .·	a_1	0
Row 2	a_{n-1}	a_{n-3}	a_{n-5}	·· .	a_0	0
Row 3	b_1	b_2	b_3	· · ·	0	0
Row 4	c_1	c_2	c_3	· · ·	**0**	**0**
				· ·	·	
				· · ·	·	
				· · ·	·	
Row n	d_1	d_2	0	· · ·	0	0
Row $n+1$	e,	0	0	· · ·	0	0

where rows 3 through $n + 1$ are calculated by

$$b, = \frac{a_{n-1} a_{n-2} - a_n a_{n-3}}{a_{n-1}} \qquad b_2 = \frac{a_{n-1} a_{n-4} - a_n a_{n-5}}{a_{n-1}} \qquad \ldots$$

$$c_1 = \frac{b_1 a_{n-3} - a_{n-1} b_2}{b_1} \qquad c_2 = \frac{b_1 a_{n-5} - a_{n-1} b_3}{b_1} \qquad \ldots$$

and so on. The process is continued until all new terms are zero. Once the array is completed, the number of roots of the polynomial that have positive real parts can be determined by counting the number of changes of sign in the extreme left-hand column of the array. In other words, for the polynomial to have all its roots in the left half of the s-plane, all the terms in the left-hand column of the array must be of the same sign.

To illustrate the use of Routh's test, let us apply it to determination of the ultimate gain of the temperature controller for the exchanger discussed previously.

EXAMPLE 6-2.5

Determine the ultimate gain of the temperature controller of Example 6-2.1 by Routh's method.

SOLUTION

From Example 6-2.1, the characteristic equation for the base case is, in polynomial form:

$$900s^3 + 420s^2 + 43s + 1 + 0.80K_c = 0$$

The next step is to prepare Routh's array.

Row 1	900	43	0
Row 2	420	$1 + 0.80K_c$	0
Row 3	b_1	0	0
Row 4	$1 + 0.80K_c$	0	0

where

$$b_1 = \frac{(420)(43) - 900(1 + 0.80K_c)}{420} = \frac{17160 - 720K_c}{420}$$

For the control loop to be stable, all the terms in the left-hand column must be of the same sign, in this case positive. This requires that

$$b_1 \geq 0 \quad \text{or} \quad 17160 - 720K_c \geq 0$$

$$K_c \leq 23.8$$

$$1 + 0.80K_c \ IO \quad \text{or} \quad 0.80K_c \geq - 1$$

$$K_c \geq -1.25$$

In this case the lower limit on K_c is negative. This is meaningless, because a negative gain means that the controller has the wrong action (opens the steam valve in response to increasing temperature). The upper limit on the controller gain is the ultimate gain that we seek.

$$K_{cu} = 23.8 \ \frac{\%CO}{\%TO}$$

This tells us that in tuning the proportional controller for this loop, we must not exceed the gain of 23.8 or reduce the proportional band below $100/23.8 = 4.2\%$. This is the same result we got with the direct substitution method.

Although it is of historical importance, Routh's test is not as useful as the direct substitution method. This is because it does not give the ultimate period for the loop. The direct substitution method is also easier to apply.

6-3   SUMMARY

This chapter presented the analysis of feedback control loops. We learned how to develop the closed-loop transfer function and the characteristic equation of the loop and how to estimate the closed-loop steady-state gain, the ultimate gain, and the ultimate period of the loop. We also saw how the various loop parameters affect the ultimate gain and period. The next chapter looks at various important methods for tuning feedback controllers.

PROBLEMS

6-1. A feedback control loop is represented by the block diagram of Fig. 6-1.4. The process can be represented by two lags in series:

$$G_1(s) = \frac{K}{(\tau_1 s + 1)(\tau_2 s + 1)}$$

where the process gain is $K = 0.10$ %TO/%CO and the time constants are

$$\tau_1 = 1 \text{ min}$$

$$\tau_2 = 0.8 \text{ min}$$

The controller is a proportional controller: $G_c(s) = K_c$.
(a) Write the closed-loop transfer function and the characteristic equation of the loop.
(b) For what values of the controller gain is the loop response to a step change in set point overdamped, critically damped, and underdamped? Can the loop be unstable?
(c) Find the effective time constants, or the (second-order) characteristic time and the damping ratio, of the closed loop for $K_c = 0.1$, 0.125, and 0.20 %CO/%TO.
(d) Determine the steady-state offset for each of the gains in part (c) and a unit step change in set point.

6-2. Do Problem 6-1 for the process transfer function

$$G(s) = \frac{6(1 - s)}{(s + 1)(0.5s + 1)} \text{ %/%}$$

Transfer functions such as this are typical of processes that consist of two lags in parallel with opposite action (see Section 4-4.3). The controller is a proportional controller as in Problem 6-1.

6-3. A feedback control loop is represented by the block diagram of Fig. 6-1.4. The

process can be represented by a first-order lag, and the controller is **proportional-integral** (PI):

$$G_1(s) = \frac{K}{\tau s + 1}$$

$$G_c(s) = K_c \left(1 + \frac{1}{\tau_I s} \right)$$

Without loss of generality, you can set the process time constant τ equal to 1 and the process gain K equal to 1.

(a) Write the closed-loop transfer function and the characteristic equation of the loop. Is there offset?

(b) Is there an ultimate gain for this loop?

(c) Determine the response of the closed loop to a step change in set point for $\tau_I = \tau$ as the controller gain varies from zero to infinity.

6-4. Consider the feedback control loop of Problem 6-1 and a pure integral controller:

$$G_c(s) = \frac{K_I}{s}$$

(a) Determine the ultimate controller gain and the ultimate period.

(b) Recalculate the ultimate controller gain for $\tau_2 = 0.10$ and for $\tau_2 = 2$. Are your results what you expected?

(c) Use Routh's test to check the ultimate gains you calculated in parts (a) and (b).

6-5. Consider the feedback control loop of Problem 6-1 and a proportional-integral controller:

$$G_c(s) = K_c \left(1 + \frac{1}{\tau_I s} \right)$$

(a) Determine the ultimate loop gain KK_{cu} and the ultimate period of oscillation as functions of the integral time τ_I.

(b) Determine the damping ratio and the decay ratio with the controller gain set equal to one-half the ultimate gain and with the integral time set equal to 1.

6-6. **Design of Gas Flow Control Loop.** A flow control loop, consisting of an orifice in series with the control valve, a differential pressure transmitter, and a controller, is to be designed for a nominal process flow of 150 kscf/h (kscf = 1000 cubic feet of gas at standard conditions of 60°F and 1 atm). The upstream conditions are constant at 100 psig and 60°F, the downstream pressure is constant at 80 psig, and the fluid is air (mol. wt. = 29). The valve has equal percentage characteristics with $\alpha = 50$, and a square root extractor is built into the transmitter so that its output signal is linear with flow. The valve time constant is 0.06 min, and the transmitter time constant is negligible. A proportional-integral (PI) controller controls the flow. Draw the block diagram of the flow control

loop, showing the specific transfer functions of the controller, the control valve, and the flow transmitter. Write the closed-loop transfer function for the loop, and find the time constant of the loop for $K_c = 0.9$ %CO/%TO and $\tau_I = \tau_v$.

6-7. Steam Flow Control Loop. A process heater requires 3500 lb/h of steam to heat a process fluid. A control valve and linear flow transmitter are installed to control the flow of the steam. The conditions are as follows: Control valve upstream conditions are 45 psig, superheated 50°F; downstream pressure is 20 psig; critical flow factor is 0.8. A linear valve sized for 100% overcapacity is proposed. The flow transmitter is sized to measure a maximum flow of 5000 lb/h, and its output is linear with flow; that is, it has a built-in square root extractor. Draw the block diagram for the flow control loop, showing all transfer functions, and write the closed-loop transfer function. Use a proportional-integral (PI) feedback controller with the integral time set equal to the time constant of the control valve. Find the time constant of the loop for $K_c = 0.5$ %CO/%TO.

6-8. For a feedback control loop represented by the block diagram of Fig. 6-1.4, determine the ultimate gain and period for a proportional controller and each of the following process transfer functions.

(a) $G_1(s) = \dfrac{1}{(s + 1)^4}$

(b) $G_1(s) = \dfrac{1}{(s + 1)^2}$

(c) $G_1(s) = \dfrac{1}{(4s + 1)(2s + 1)(s + 1)}$ /

(d) $G_1(s) = \dfrac{(0.5s + 1)}{(4s + 1)(2s + 1)(s + 1)}$

(e) $G_1(s) = \dfrac{1}{(4s + 1)(0.2s + 1)(0.1s + 1)}$

(f) $G_1(s) = \dfrac{e^{-0.6s}}{6s + 1}$

6-9. Check the ultimate gains of Problem 6-8 using Routh's test.

6-10. An open-loop unstable process can be represented by the block diagram of Fig. 6- 1.4 and the following transfer function:

$$G_1(s) = \frac{K}{(5s - 1)(\tau_v s + 1)(\tau_T s + 1)}$$

where τ_v and τ_T are, respectively, the time constants of the control valve and the transmitter. Assuming a proportional controller, find the range of the loop gain KK_c for which the loop is stable if

(a) The valve and transmitter time constants are negligible.
(b) The valve time constant is negligible and $\tau_T = 1.0$ min.
(c) $\tau_v = 0.1$ min and $\tau_T = 1.0$ min.

Hint: Parts (a) and (b) can be solved by writing the roots of the characteristic equation as functions of the loop gain. Part (c) requires the direct substitution method. Note that there is a lower stability limit on the loop gain for each case.

6-11. Calculate the ultimate gain and period of oscillation for a proportional analyzer controller installed on the blending tank of Problem 3-18. The control valve is to be installed on the dilute stream and sized for 100% overcapacity. The valve has linear characteristics, has a constant pressure drop of 5 psi, and can be represented by a first-order lag with a time constant of 0.1 min. The analyzer transmitter has a range of 20 to 70 kg/m^3 and can be represented by a first-order lag with a time constant of 3 min. Also draw the block diagram of the loop, showing all transfer functions, and calculate the offset caused by a change of 0.1 m^3/min in the flow of the concentrated solution when the controller gain is one-half the ultimate gain.

6-12. In Section 4-2.3, a nonisothermal chemical reactor is modeled in detail. Calculate the ultimate gain and period of a proportional temperature controller for the reactor, assuming that the control valve is installed on the cooling water line to the jacket to manipulate the cooling water flow $f_c(t)$. The valve is equal percentage with constant pressure drop and $\alpha = 50$, and the temperature transmitter has a range of 640 to 700°R. The time constants of the valve and the transmitter can be neglected. Also draw the block diagram of the temperature control loop, showing all transfer functions. *Hint:* For simplicity, you may consider the cooling water flow as the only input variable; that is, assume all other input variables are constant.

6-13. Solve Problem 6-12 assuming that the control valve is installed on the reactants line to manipulate the flow of reactants $f(t)$. Assume that all other input variables, including the coolant flow, are constant. Use a linear control valve that has constant pressure drop and is sized for 100% overcapacity.

6-14. In Problem 4-4 you are asked to model three mixing tanks in series. Find the ultimate gain and period of a proportional controller that is to control the outlet composition from the third tank, $x_6(t)$, by manipulating the flow of water into the first tank, $f_1(t)$. Assume an equal percentage control valve with constant pressure drop and $\alpha = 50$. The analyzer transmitter has a range of 0.30 to 0.70 mass fraction units, and the time constants of the valve and transmitter can be neglected. Also draw the block diagram of the loop and calculate the offset caused by a change of 10 gpm in flow f_2 when the controller gain is set equal to one-half the ultimate gain.

6-15. *Feedback Control of Reactors in Series.* Consider the control of the concentration out of the second of the two reactors in series of Problem 4-5 by manipulating the reactants flow. Each of the reactors has a volume of 125 ft^3, the inlet concentration is initially $c_{Ai}(0) = 7.0$ lbmole/ft³, and the rate coefficient in each reactor is $k = 0.2$ min⁻¹. The initial inlet flow is $f(0) = 10$ ft³/min., and the recycle flow is zero. The control valve is linear, is sized for 100% overcapacity, and has a negligible time lag. The analyzer transmitter has a range of 0 to 5 lbmoles/ft³, and it can be represented by a first-order lag with a time constant of 0.5 min. The set point of the analyzer controller is initially equal to the initial steady-state value of the concentration from the second reactor. Draw the block diagram of the control loop, showing all transfer functions, and calculate the ultimate gain and period of the loop for a proportional controller. Also calculate the offset of the controller for a change in inlet concentration from 7.0 to 8.0 lbmoles/ft³, assuming the controller gain is set equal to one-half the ultimate gain. What would the offset be if the loop were opened (controller on "manual,"

or $K_c = O.$)? What would it be if the controller were a proportional-integral (PI) controller?

6-16. In Problem 4-14 you were asked to model a tank in which steam is mixed with a liquid stream. Find the ultimate gain and period for a proportional controller that is to control the temperature of the stream leaving the tank by manipulating the steam valve position. Also draw the block diagram for the temperature control loop, showing all transfer functions. Calculate the offset caused by a 2-gpm change in the inlet liquid flow when the controller gain is set equal to one-half the ultimate gain.

6-17. *Composition Control of Three Isothermal Reactors in Series.* Consider the concentration control loop for the three stirred reactors shown in Fig. P6-1. Each reactor has a volume of 1000 gal, and the design flow of reactants is 100 gpm. The initial inlet concentration of reactant A into the first reactor is 4 lb/gal, and the reaction rate is proportional to the concentration of A in each reactor with a constant coefficient of 0.1 min- ^{1}. The outlet concentration transmitter has a range of 0 to 1.0 lb/gal and a negligible time constant. The reactant control valve is linear with constant pressure drop of 5 psi and is sized for 100% overcapacity. The time constant of the valve is negligible.

(a) Draw a block diagram of the loop, showing all transfer functions.

(b) If a proportional controller with a gain of 1 .O %CO/%TO is installed on this system, what will be the offset caused by a 1 .0 lb/gal change in inlet reactant concentration? What would be the offset if the controller gain were zero? What if the controller were proportional-integral (PI)?

(c) Calculate the ultimate gain and period for the loop, assuming a proportional controller.

6-18. *Compressor Suction Pressure Control.* Figure P6-2 shows the schematic of a compressor suction pressure control loop. A mass balance on the suction volume results in the following approximate linear model for the suction pressure.

$$P_s(s) = \frac{0.5}{7.5s + 1} [F_i(s) - F_c(s)] \text{ psi}$$

where $F_i(s)$ and $F_c(s)$ are, respectively, the inlet and compressor flows, kscf/min (1 kscf = 1000 ft^3 at standard conditions of 1 atm and 60°F), and the time constant is in seconds. The response of the compressor flow to the controller output signal, M(s), %CO, is

$$F_c(s) = \frac{0.36}{2.5s + 1} M(s)$$

The pressure transmitter has a range of 0 to 20 psig and can be represented by a first-order lag with a time constant of 1.2 s.

(a) Draw the block diagram for the loop and write the closed-loop transfer function and the characteristic equation. Must the controller be direct-acting or reverse-acting?

(b) Calculate the ultimate gain and period, assuming a proportional controller.
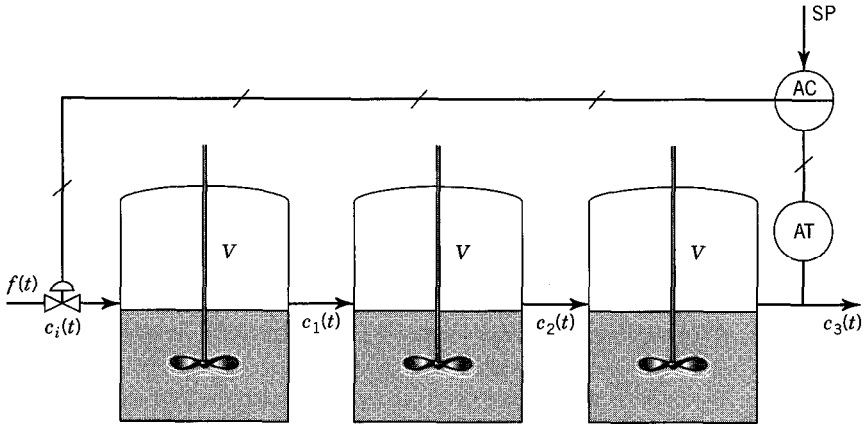
Figure P6-1 Three stirred reactors in series for Problem 6-17.

(c) Calculate the offset caused by a change of 1.0 kscf/min in the inlet flow when the controller gain is one-half the ultimate gain.

6-19. The parameters for the stirred tank cooler of Problem 4-10, which is sketched in Fig. P4-9, are $V = 5.0$ m^3, $U = 200$ kJ/min-m^2-°C, $A = 4.0$ m^2, $V_c = 1.1$ m^3, $\rho = 800$ kg/m^3, $c_p = 3.8$ kJ/kg-°C, $\rho_c = 1000$ kg/m^3, and $c_{pc} = 4.2$ kJ/kg-°C. The design conditions are as follows: process flow = 0.10 m^3/min, inlet process
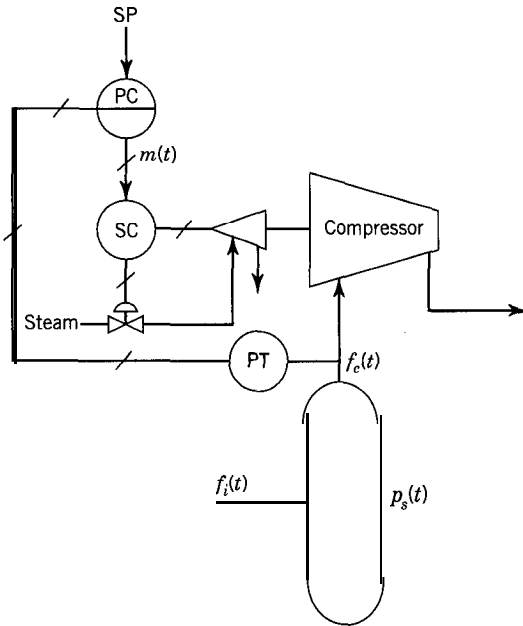
SC = Turbine speed controller

Figure P6-2 Compressor pressure control for Problem 6-18.

temperature = 70°C, outlet process temperature = 45°C, and coolant inlet temperature = 25°C. The temperature transmitter (TT) has a range of 20 to 70°C and can be represented by a first-order lag with a time constant of 0.6 min. The coolant flow transmitter (FT) has a range of 0 to 0.8 m³/min, and the time constant of the flow control loop (FC) is negligible.

(a) Draw the complete block diagram for the temperature control loop, showing all transfer functions. Should the coolant valve fail open or closed? Must the controller be direct-acting or reverse-acting?

(b) Calculate the ultimate gain and period for the temperature control loop, assuming a proportional controller (TC).

(c) Calculate the offset caused by a 5°C increase in inlet process temperature when the temperature controller gain is one-half the ultimate gain. What is the offset when the controller gain is zero (manual state)? What is the offset if the controller is proportional-integral (PI)?

6-20. The gas storage tank shown in Figure P6-3 supplies a gas with a molecular weight of 50 to two processes. The first process receives a normal flow of 500 scf/min (scf = ft³ at 1 atm and 60°F) and operates at a pressure of 30 psig, and the second process operates at a pressure of 15 psig. A process operating at 90 psig supplies gas to the storage tank at a rate of 1500 scf/min. The tank has a capacity of 550,000 ft³ and operates at 45 psig and 350°F. You may assume that the pressure transmitter responds instantaneously with a calibrated range of 0 to 100 psig.

(a) Size all three valves for 100% overcapacity. For all valves, you can use the factor $C_f = 0.9$ (Masoneilan).

(b) Draw the complete block diagram for the system. You can consider as disturbances $P_1(t)$, $P_3(t)$, $P_4(t)$, $vp_3(t)$, $vp_4(t)$ and the set point to the controller.

(c) Can the feedback loop go unstable? If so, what is its ultimate gain?

(d) When a proportional-only controller with a gain of 50 %CO/%TO is used, what is the offset observed for a set point change of + 5 psi?

6-21. Consider the electric heater shown in Fig. P6-4. Two liquid streams with variable mass rates $w_A(t)$ and $w_B(t)$ come together in a tee and pass through the heater, where they are thoroughly mixed and heated to temperature T(t). The outlet temperature is controlled by manipulating the current through an electric coil. The outlet mass fraction of component B is also controlled by manipulating the inlet flow of stream B. The following information is known:
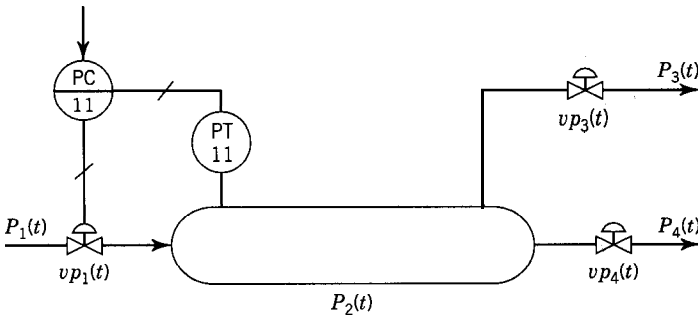
Figure P6-3 Gas storage tank for Problem 6-20.
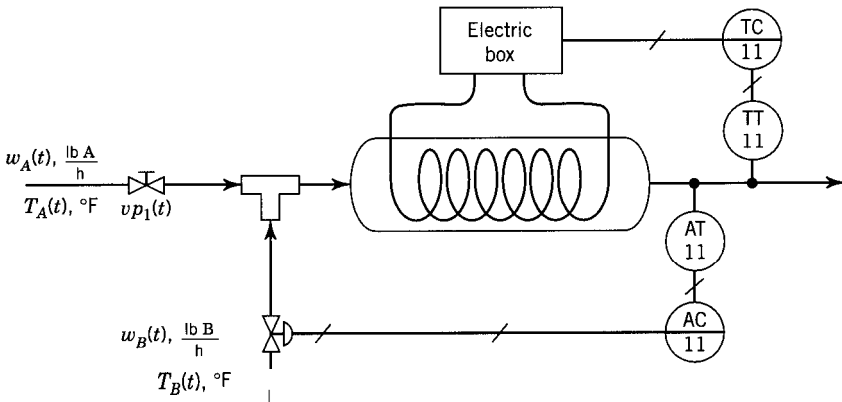
Figure P6-4 Electric heater for Problem 6-21.

- The pressure drop across the valves can be assumed constant, so the flow through the valves is given by

$$w_A(t) = K_{v_1}vp_1(t) \qquad\qquad w_B(t) = K_{v_2}vp_2(t)$$

These two streams are pure in components A and B, respectively.
- The mass fraction of B in the outlet stream is related to the electrical conductivity of the stream. The conductivity of this stream is inversely proportional to the mass fraction, x_B; that is,

$$\text{Conductivity} = \frac{\beta}{x_B}$$

where β is a constant, mho-mass fraction/m. The conductivity transmitter has a range of C_L to C_H mho/m.
- You may assume that the heat transfer rate, q, is linear with the output of the controller in the range 0 to q_{max}.

The disturbances to this system are $vp_1(t)$, $T_A(t)$, and $T_B(t)$.
(a) Derive, from basic principles, the set of equations that describes the composition (conductivity) control loop. State all assumptions.
(b) Linearize the equations from part (a) and draw the complete block diagram for the conductivity loop. Show the transfer function of each block. Specify the required action of the composition controller, assuming that the control valve is air-to-open.
(c) Derive, from basic principles, the set of equations that describes the temperature control loop. State all assumptions.
(d) Linearize the equations from part (c) and draw the complete block diagram for the temperature loop. Show the transfer function of each block. Specify the action of the temperature controller.
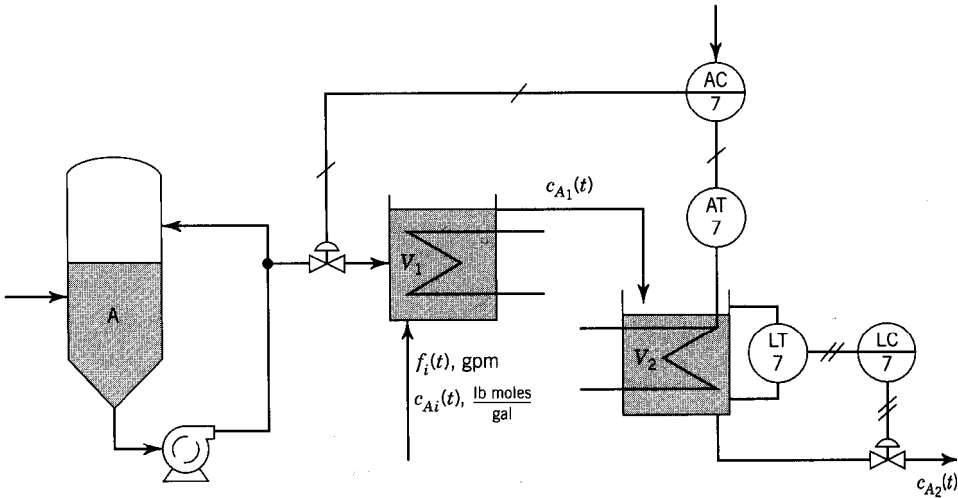
Figure P6-5 Reactors in series for Problem 6-22.

(e) Write the characteristic equation for each of the control loops. Can either loop be made unstable by increasing the controller gain? Discuss briefly.

6-22. Consider the system shown in Fig. P6-5. In each of the two tanks, the reaction A → E takes place. The rate of reaction is

$$r(t) = kC_A(t), \quad \frac{\text{lbmoles}}{\text{gal-min}}$$

where k is the reaction rate coefficient, min^{-1}, and C_A is the concentration, lbmoles/gal. The disturbances to this process are $f_i(t)$ and $c_{Ai}(t)$. The concentration out of the second reactor is controlled by manipulating a stream of pure A to the first reactor. The density of this stream is ρ_A in lbmoles/gal. The temperature in each reactor can be assumed constant. The following design data are known.

Reactor volumes:	$V_1 = 500$ gal	$V_2 = 500$ gal
Reaction rate coefficients:	$k_1 = 0.25$ min^{-1}	$k_2 = 0.50$ min^{-1}
Properties of stream A:	$\rho_A = 2.0$ lbmoles/gal	$MW_A = 25$
Design conditions:	$\bar{c}_{Ai} = 0.8$ lbmole/gal	$\bar{f}_i = 50$ gal/min
	$\bar{f}_A = 50$ gal/min	
Control valve:	$\Delta p_v = 10$ psi, linear characteristics	

Concentration transmitter range: 0.05 to 0.5 lbmole/gal. The dynamics of this transmitter can be represented by a first-order lag with a time constant of 0.5 min.

(a) Size the control valve for 100% overcapacity. Report the C_v and the gain of the valve.

(b) Derive, from basic principles, the set of equations that describes the composition control loop. State all assumptions.

(c) Linearize the equations from part (b) and draw the complete block diagram of the composition control loop. Show all transfer functions with the numerical values and units of all gains and time constants, except for the controller.

(d) Obtain the closed-loop transfer functions

$$\frac{C_{A_2}(s)}{C_{A_2}^{set}(s)} \qquad \frac{C_{A_2}(s)}{F_i(s)} \qquad \frac{C_{A_2}(s)}{C_{A_i}(s)}$$

(e) Calculate the ultimate gain and period of the loop.

6-23. Consider the process shown in Fig. P6-6. In the first tank, two streams of rates $f_1(t)$ and $f_2(t)$ are being mixed and heated. The heating medium flows at such a high rate that its temperature change from inlet to exit is not significant. Thus the heat transfer rate can be described by $UA[T_{c1}(t) - T_3(t)]$. It can also be assumed that the densities and heat capacities of all streams are not strong functions of temperature or composition. The outlet flow from the first tank flows into the second tank, where it is again heated, this time by condensing steam. The rate of heat transfer can be described by $w_s(t)\lambda$, where $w_s(t)$ is the mass flow rate of steam and λ is the latent heat of vaporization of the steam. Assume that the pressure drop across the steam valve is constant and that its time constant is τ_v. The temperature transmitter has a range of T_L to T_H and a time constant τ_T. Assuming that the heat losses from both tanks are negligible and that the important disturbances are $T_1(t)$, $T_2(t)$, and $T_{c1}(t)$, obtain the complete block diagram
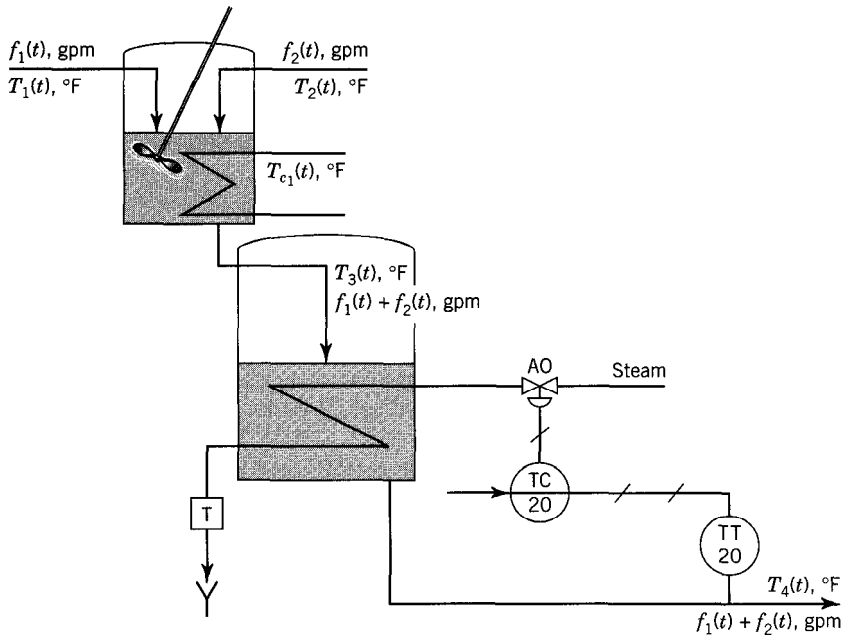

Figure P6-6 Heaters for Problem 6-23.
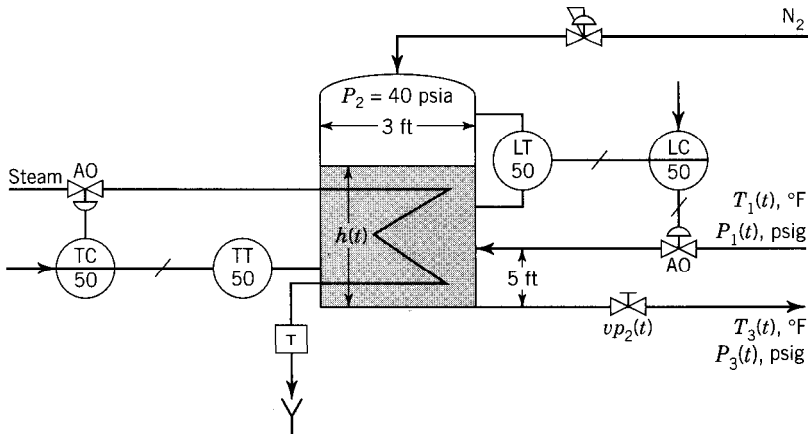
Figure **P6-7** Oil heater for Problem 6-24.

of the temperature control loop and its characteristic equation. Derive the transfer function of each block.

6-24. Consider the process shown in Fig. **P6-7**. The process fluid entering the tank is an oil with a density of 53 lb/ft³, a heat capacity of 0.45 Btu/lb-°F, and an inlet temperature of 70°F. This oil is to be heated up to 200°F by saturated steam at 115 psig. The pressure in the tank, above the oil level, is maintained at 40 psia by a blanket of inert gas, N_2. Assume that the tank is well insulated, that the physical properties of the oil are not strong functions of temperature, that the liquid is well mixed, and that the level covers the heating coil. The following data are also known:

$\bar{p}_1 = 4.5$ psig $\bar{p}_3 = 15$ psig

Heat transfer coefficient = 136

Heating surface area = 127.5 ft²

Heating coil: $\frac{1}{2}$ in. O.D., 20 BWG tubes, 974 linear ft., mass of tube metal = 0.178 lb/ft, c_p of tube metal = 0.12 Btu/lb-°F

Tank diameter: 3 ft

Level transmitter: 7 to 10-ft range, 0.01 min time constant

Temperature transmitter: 100 to 300°F range, 0.5 min time constant

(a) Size the control valves for 50% overcapacity. The nominal oil flow rate is 100 gpm. The pressure drop across the steam valve can be assumed to be constant.

(b) Obtain the complete block diagram for the level control loop. Use a proportional-only controller.

(c) Obtain the complete block diagram and the characteristic equation of the temperature control loop. Use a P controller. Show the numerical values of all the gains and time constants in the transfer functions.

(d) Calculate the ultimate gain and period of oscillation of the temperature control loop.

6-25. Consider the process shown in Fig. **P6-8**. The two outlet valves remain at constant opening and their downstream pressure is also constant; the pump flow is linear with the controller output in the range 0 to f_{max} ; the variable speed pump has a
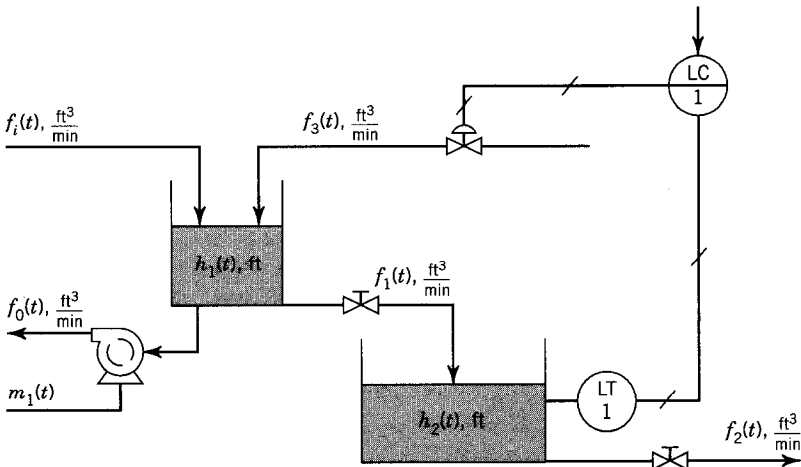
Figure P6-8 Level controller for Problem 6-25.

time constant, relating the flow to the input signal, m,(t), of τ_p s; the control valve is linear and has a time constant, relating the flow to the pneumatic signal, of τ_v s; the pressure drop across the control valve is constant; the level transmitter has a range of 0 to h_{max} and a negligible time constant. The diameters of the tanks are D, and D,. The valve coefficients are C_{v_1}, C_{v_2}, and C_{v_3}.

(a) Draw the block diagram and derive the transfer functions for this control system. The disturbances **are** $f_i(t)$ and m,(t).

(b) Write the characteristic equation of the level control loop and determine its ultimate gain and period as functions of the system parameters.

6-26. Consider the process shown in Fig. **P6-9.** In this process, a waste gas is enriched with natural gas to be used as fuel in a small furnace. The enriched waste gas must have a certain heating value to be used as fuel. The control strategy calls for measuring the heating value of the gas leaving the process and manipulating the natural gas flow (using a variable-speed fan) to maintain the heating value set point. The waste gas is composed of methane (CH,) and some low-heating-value combustibles. The natural gas is composed mainly of methane and some small amount of other hydrocarbons, and its composition can be considered con-
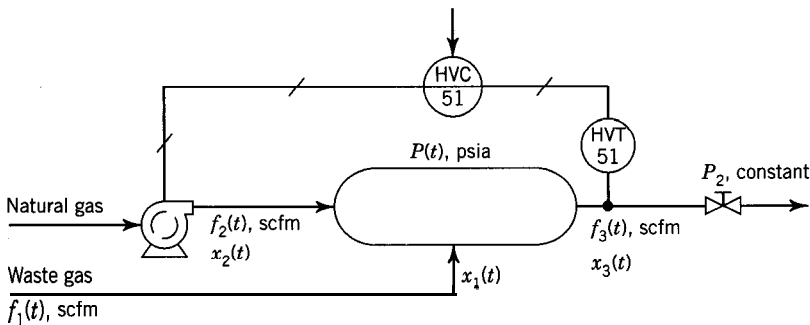
Figure P6-9 Gas enriching tank for Problem 6-26.

stant. The heating value of the enriched waste gas is related to the mole fraction of methane by the following relation:

$$hv(t) = c + gx_3(t)$$

where $hv(t)$ is the heating value, $x_3(t)$ is the mole fraction of methane, and c and g are constants. The variable-speed fan is such that at full speed its flow is $f_{2,max}$. It can be assumed that the relationship between the flow and the input signal to the fan driver is linear. This driver has a time constant τ_F. The outlet valve has a constant opening. A proportional-integral controller is used to control the heating value. The sensor/transmitter has a time constant of τ_T min. The specific gravity of the enriched gas is related to the mole fraction of methane by

$$G(t) = a + bx_3(t)$$

where a and b are constants.

(a) Draw the block diagram for this control system and derive all transfer functions. The possible disturbances are $f_1(t)$, $x_1(t)$, and $x_2(t)$.
(b) Write the characteristic equation of the feedback control loop.

Chapter 7

Tuning of Feedback Controllers

In this chapter we will study the tuning of feedback controllers-that is, the adjustment of the controller parameters to match the characteristics (or personality) of the rest of the components of the loop. We will look at two methods for characterizing the process dynamic characteristics: the on-line or closed-loop tuning method, and the step-testing or open-loop method. We will also look at three different specifications of control loop performance: quarter decay ratio response, minimum error integral, and controller synthesis. This latter method, in addition to providing some simple controller-tuning relationships, will give us some insight into the selection of the proportional, integral, and derivative modes for various process transfer functions.

Tuning is the adjusting of the feedback controller parameters to obtain a specified closed-loop response. The tuning of a feedback control loop is analogous to the tuning of an automobile engine, a television set, or a stereo system. In each of these cases, the difficulty of the problem increases with the number of parameters that must be adjusted. For example, tuning a simple proportional-only or integral-only controller is similar to adjusting the volume of a stereo sound system. Because only one parameter or "knob" needs to be adjusted, the procedure consists of moving it in one direction or the other until the desired response (or volume) is obtained. The next degree of difficulty is the tuning of a two-mode or proportional-integral (PI) controller, which is similar to adjusting the bass and treble on a stereo system. Two parameters, the gain and the reset time, must be adjusted, so the tuning procedure is significantly more complicated than when only one parameter is involved. Finally, the tuning of three-mode or proportional-integral-derivative (PID) controllers represents the next higher degree of difficulty. Here three parameters-the gain, the reset time, and the derivative time-must be adjusted.

Although we have drawn an analogy between the tuning of a stereo system and that of a feedback control loop, we do not want to give the impression that the two tasks have the same degree of difficulty. The main difference lies in the speed of response of the stereo system versus that of a process loop. With the stereo system, we get almost immediate feedback on the effect of our tuning adjustments. On the other hand, although some process loops do have relatively fast responses, for many process loops we may have to wait several minutes and maybe even hours to observe the response that results

303

from our tuning adjustments. This makes tuning feedback controllers by trial and error a tedious and time-consuming task. Yet this is the method most commonly used by control and instrument engineers in industry. A number of procedures and formulas have been introduced to help enhance tuning effectiveness and give insight into tuning itself. We will study some of these procedures in this chapter. However, keep in mind that no one procedure will give the best results for all process control situations.

The values of the tuning parameters depend on the desired closed-loop response and on the dynamic characteristics, or personality, of the other elements of the control loop, particularly the process. We saw in Chapters 3 and 4 that if the process is nonlinear, as is usually the case, then its characteristics change from one operating point to the next. This means that a particular set of tuning parameters can produce the desired response at only one operating point, given that standard feedback controllers are basically linear devices. For operation in a range of operating conditions, a compromise must be reached in arriving at an acceptable set of tuning parameters, because the response will be sluggish at one end of the range and oscillatory at the other.

One characteristic of feedback control that greatly simplifies the tuning procedure is that the performance of the loop is not a strong function of the tuning parameters. In other words, the performance does not vary much with the tuning parameters. Changes of less than 50% in the values of the tuning parameters seldom have significant effects on the response of the loop. Accordingly, we will not show the values of the tuning parameters with more than two significant digits. With this in mind, let us look at some of the procedures that have been proposed for tuning industrial controllers.

7-1 QUARTER DECAY RATIO RESPONSE BY ULTIMATE GAIN

This pioneer method, also known as the closed-loop or on-line tuning method, was proposed by Ziegler and Nichols in 1942. Like all the other tuning methods, it consists of two steps:

Step 1. Determination of the dynamic characteristics, or personality, of the control loop.
Step 2. Estimation of the controller tuning parameters that produce a desired response for the dynamic characteristics determined in the first step-in other words, matching the personality of the controller to that of the other elements in the loop.

In this method the dynamic characteristics of the process are represented by the *ultimate gain* of a proportional controller and the *ultimate period* of oscillation of the loop. These parameters, introduced in Section 6-2, can be determined by the direct substitution method if the transfer functions of all the components of the loop are known quantitatively. But because this is not usually the case, we must often experimentally determine the ultimate gain and period from the actual process by the following procedure:

1. Switch off the integral and derivative modes of the feedback controller so as to have a proportional controller. In some controllers, the integral mode cannot be switched off but can be de-tuned by setting the integral time to its maximum value or, equivalently, the integral rate to its minimum value.
2. With the controller in *automatic (i.e.,* the loop closed), increase the proportional gain (or reduce the proportional band) until the loop oscillates with constant amplitude. Record the value of the gain that produces sustained oscillations as K_{cu},
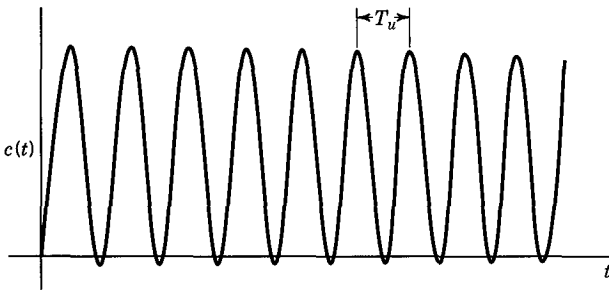
Figure 7-1.1 Response of the loop with the controller gain set equal to the ultimate gain K_{cu}. T_u is the ultimate period.

the ultimate gain. This step is carried out in discrete gain increments, bumping the system by applying a small change in set point at each gain setting. To prevent the loop from going unstable, smaller increments in gain are made as the ultimate gain is approached.

3. From a time recording of the controlled variable such as Fig. 7-1.1, the period of oscillation is measured and recorded as T_u, the ultimate period.

For the desired response of the closed loop, Ziegler and Nichols specified a decay ratio of one-fourth. The decay ratio is the ratio of the amplitudes of two successive oscillations. It should be independent of the input to the system and should depend only on the roots of the characteristic equation for the loop, Typical quarter decay ratio responses for a disturbance input and a set point change are shown in Fig. 7-1.2.

Once the ultimate gain and period are determined, they are used in the formulas of Table 7-1.1 for calculating the controller-tuning parameters that produce quarter decay ratio responses.

Table 7-1.1 shows that the introduction of integral mode forces a reduction of 10% in the gain of the PI controller as compared to the proportional controller gain. Derivative mode, on the other hand, allows an increase in both the proportional gain and the integral rate (a decrease in integral time) of the PID controller as compared to the PI controller. This is because the integral mode introduces a lag in the operation of the feedback controller, whereas the derivative mode introduces an advance or lead. This will be discussed in more detail in Chapter 9.
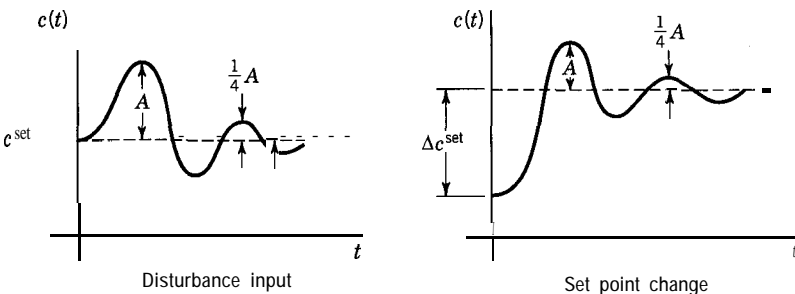
Figure   7-1.2 Quarter   decay   ratio   response to   disturbance   input   and   to change in set point.

Table 7-1.1 Quarter Decay Ratio Tuning Formulas

Controller Type	Proportional Gain, K_c'	Integral Time, τ_I'	Derivative Time, τ_D'
Proportional-only, P	$\dfrac{K_{cu}}{2}$	—	—
Proportional-integral, PI	$\dfrac{K_{cu}}{2.2}$	$\dfrac{T_u}{1.2}$	—
Proportional-integral- derivative, PID	$\dfrac{K_{cu}}{1.7}$	$\dfrac{T_u}{2}$	$\dfrac{T_u}{8}$

The PID formulas are for the actual PID controller, Eq. 5-3.19. To convert to the ideal PID controller, Eq. 5-3.17: $K_c = K_c'(1 + \tau_D'/\tau_I'); \tau_I = \tau_I' + \tau_D'; \tau_D = \tau_D'\tau_I'/(\tau_I' + \tau_D')$.

The quarter decay ratio response is very desirable for disturbance inputs, because it prevents a large initial deviation from the set point without being too oscillatory. However, it is not so desirable for step changes in set point, because it causes a 50% overshoot. This is because the maximum deviation from the new set point in each direction is one-half the preceding maximum deviation in the opposite direction (see Fig. 7-1.2). This difficulty can easily be corrected by reducing the proportional gain from the value predicted by the formulas of Table 7-1.1. In fact, the decay ratio is a direct function of the controller gain and can be adjusted at any time by simply changing the gain. In other words, if for a given process the quarter decay ratio response is too oscillatory, then a reduction of the gain will smooth out the response.

The discussion in the preceding paragraph brings out the main advantage of the closed-loop tuning formulas: they reduce the tuning procedure to the adjustment of a single parameter, the controller gain. On the assumption that a good estimate of the ultimate period can be obtained by observing the closed-loop response, the reset and derivative times can be set on the basis of this value. The response can then be molded by adjusting the proportional gain. Because of the insensitivity of the response to the precise values of the tuning parameters, it is not absolutely necessary to make the closed-loop response oscillate with sustained oscillations. Any oscillation caused by the proportional controller can be used to obtain an approximate value of the ultimate period that is usually good enough for tuning.

It has been said that one difficulty in using the quarter decay ratio response is that, except for the case of the proportional controller, the set of tuning parameters necessary to obtain it is not unique. In the case of PI controllers, we can easily verify that for each value of the integral time, we could find a value of the gain that produces a quarter decay ratio response, and vice versa. The same is true for the PID controller. The simple tuning formulas proposed by Ziegler and Nichols give ball-park figures that produce fast response for most industrial loops.

The PID tuning formulas of Table 7- 1.1 are for the "actual" PID controller transfer function given by Eq. 5-3.19. We know this because these were the only PID controllers available when Ziegler and Nichols developed their tuning formulas. Today, many computer control packages use the "ideal" PID transfer function of Eq. 5-3.17. Table 7-1.1 includes formulas for calculating the parameters of the ideal PID controller.

EXAMPLE 7-1.1

Given the characteristic equation of the continuous stirred tank heater derived in Example 6- 1.1, determine the quarter decay ratio tuning parameters for the PID controller by the ultimate gain method. Also calculate the roots of the characteristic equation for the controller tuned with these parameters, and calculate the actual decay ratio.

SOLUTION

In Example 6-2.2 the ultimate gain and period of the loop for a proportional controller were obtained by direct substitution.

$$K_{cu} = 10.4 \ \frac{\%CO}{\%TO}, \quad T_u = 4.6 \ \text{min}$$

According to Table 7- 1.1, the tuning parameters for the quarter decay ratio response of a PID controller are

$$K'_c = \frac{K_{cu}}{1.7} = 6.1 \ \frac{\%CO}{\%TO} \quad \tau'_I = \frac{4.6}{2} = 2.3 \ \text{min} \quad \tau'_D = \frac{4.6}{8} = 0.58 \ \text{min}$$

The transfer function of the PID controller is

$$G_c(s) = 6.1 \left(1 + \frac{1}{2.3s} \right) (1 + 0.58s)$$

Substitute into the characteristic equation from Example 6-2.2, and clear fractions.

$$s(8.34s + 1)(0.502s + 1)(0.2s + 1)(0.75s + 1)$$
$$+ (1.954)(6.1)(s + 0.435)(1 + 0.58s) = 0$$

where $0.435 = 1/2.3$. In polynomial form,

$$0.628s^5 + 5.303s^4 + 12.73s^3 + 16.74s^2 + 15.89s + 5.125 = 0$$

Using a computer program, we find that the roots of this characteristic equation are

$$0.327 \pm i1.232 \quad - 0.519 \quad - 1.784 \quad - 5.49$$

The response of the closed loop has the following form:

$$T(t) = b_1 e^{-0.327t} \sin(1.232t + \theta_1) + b_2 e^{-0.519t} + b_3 e^{-784t} + b_4 e^{-5.49t} + \text{(input terms)}$$

where the parameters b_1, b_2, b_3, b_4, and θ_1, must be evaluated by partial fractions expansion for the particular input (set point or inlet flow) under consideration. The technique of partial fractions expansion was discussed in Chapter 2.

The pair of complex conjugate roots dominate the response, because the corresponding term in the response decays the slowest. It has a period of oscillation of $T = 2\pi/1.232 = 5.1$ min, a 1% settling time of $-5/-0.327 = 15.3$ min, and a decay ratio of

$$e^{-0.327(5.1)} = 0.19$$

which is close to the theoretical decay ratio of 0.25. This shows that the tuning formulas are not exact in terms of the response specification.

7-2 OPEN-LOOP PROCESS CHARACTERIZATION

The Ziegler-Nichols on-line tuning method we have just introduced is the only one that characterizes the process by the ultimate gain and the ultimate period. Most of the other controller-tuning methods characterize the process by a simple first- or second-order model with dead time. In order to understand better the assumptions involved in such characterization, let us consider the block diagram of a feedback control loop given in Fig. 7-2.1. The symbols shown in the block diagram are

$R(s)$ = the Laplace transform of the set point
$M(s)$ = the Laplace transform of the controller output
$C(s)$ = the Laplace transform of the transmitter output
$E(s)$ = the Laplace transform of the error signal
$U(s)$ = the Laplace transform of the disturbance
$G_c(s)$ = the controller transfer function
$G_v(s)$ = the transfer function of the control valve (or final element)
$G_p(s)$ = the process transfer function between the controlled and
     manipulated   variables
$G_u(s)$ = the process transfer function between the controlled variable
     and the disturbance
$H(s)$ = the transfer function of the sensor/transmitter

Using the simple block diagram algebra manipulations of Chapter 3, we can draw the equivalent block diagram shown in Fig. 7-2.2. In this diagram there are only two blocks in the control loop (one for the controller and the other for the rest of the components of the loop) plus one block for each disturbance. The advantage of this simplified representation is that it highlights the two signals in the loop that can be
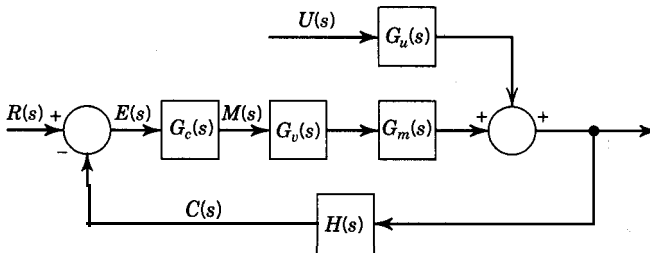

Figure 7-2.1 Block diagram of typical feedback control loop.
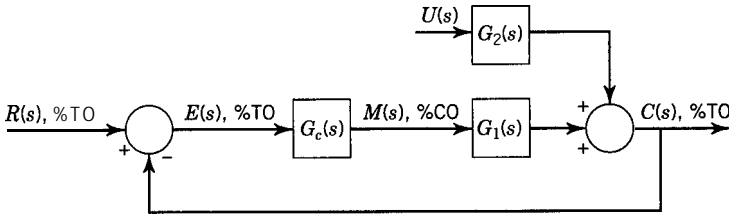
Figure 7-2.2 Equivalent simplified block diagram in which all the field instruments and the process have been lumped into single blocks.

usually observed and recorded: the controller output $M(s)$ and the transmitter signal $C(s)$. For simple loops, no signal or variable can be observed except these two. Therefore, lumping the transfer functions of the control valve, the process, and the sensor/ transmitter into a single block is not just a convenience but a practical necessity. Let us call these combinations of transfer functions G,(s) and $G_2(s)$:

$$G_1(s) = G_v(s)G_m(s)H(s)$$

$$G_2(s) = G_u(s)H(s)$$

(7-2.1)

The combined transfer function that is in the loop, G,(s), is precisely what is approximated by low-order models for the purpose of characterizing the dynamic response of the process. The point is that the characterized "process" includes the dynamic behavior of the control valve and of the sensor/transmitter. The two models most commonly used to characterize the process are

First-Order-Plus-Dead-Time (FOPDT) Model

$$G_1(s) = \frac{Ke^{-t_0s}}{\tau s + 1}$$

(7-2.2)

Second-Order-Plus-Dead-Time (SOPDT) Model

$$G_1(s) = \frac{Ke^{-t_0s}}{(\tau_1 s + 1)(\tau_2 s + 1)}$$

(7-2.3)

$$G_1(s) = \frac{Ke^{-t_0s}}{\tau^2 + 2\zeta\tau s + 1}$$

(7-2.4)

where

K = the process steady-state gain
t_0 = the effective process dead time
τ, τ_1, τ_2 = the effective process time constants
ζ = the effective process damping ratio

For underdamped processes, $\zeta < 1$.

Of these, the FOPDT model is the one on which most controller-tuning formulas are based. This model characterizes the process by three parameters: the gain K, the dead

time t_0, and time constant τ. The question, then, is how these parameters can be determined for a given loop. The answer is that some dynamic test must be performed on the actual system or on a computer simulation of the process. The simplest test that can be performed is a step test.

7-2.1 Process Step Testing

The step test procedure is carried out as follows:

1. With the controller on "manual" (that is, with the loop opened) apply a step change in the controller output signal m(t) to the process. The magnitude of the change should be large enough for the consequent change in the transmitter signal to be measurable, but it should not be so large that the response will be distorted by the process   nonlinearities.
2. Record the response of the transmitter output signal c(t) on a strip chart recorder or equivalent device, making sure that the resolution is adequate in both the amplitude and the time scale. The resulting plot of c(t) versus time must cover the entire test period from the introduction of the step test until the system reaches a new steady state. Typically, a step test lasts between a few minutes and several hours, depending on the speed of response of the process.

It is of course imperative that no disturbances enter the system while the step test is performed. A typical test plot, also known as a **process reaction** curve, is sketched in Fig. 7-2.3. As we saw in Chapter 2, the S-shaped response is characteristic of **second-** and higher-order processes with or without dead time. The next step is to match the process reaction curve to a simple, process model in order to determine the model parameters. Let us do this for the first-order-plus-dead-time (FOPDT) model.

In the absence of disturbances, and for the conditions of the test, the block diagram of Fig. 7-2.2 can be redrawn as in Fig. 7-2.4. The response of the transmitter output signal is given by

$$C(s) = G(s)M(s) \tag{7-2.5}$$
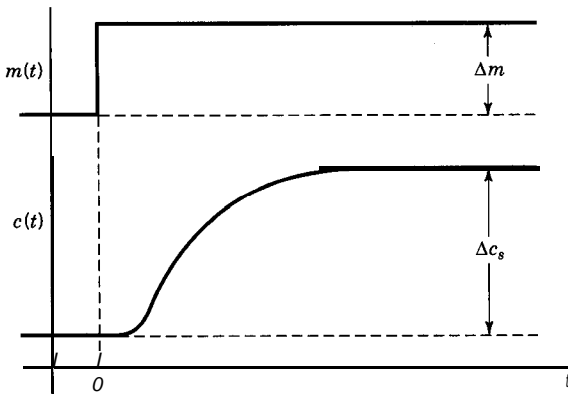
Figure 7-2.3 Process reaction curve or open-loop step response.
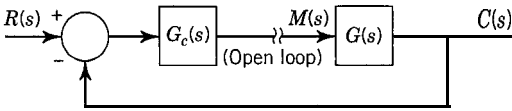
Figure 7-2.4 Block diagram for open-loop step test.

For a step change in controller output of magnitude Am and a FOPDT model, Eq. 7-2.2, we have

$$C(s) = \frac{Ke^{-t_0s}}{\tau s + 1} \cdot \frac{A\ m}{s}$$

(7-2.6)

Expanding this expression by partial fractions (see Chapter 2), we obtain

$$C(s) = K\ Am\ e^{-t_0s} \left[\frac{1}{s} - \frac{\tau}{\tau s + 1} \right]$$

(7-2.7)

Inverting with the help of a Laplace transform table (Table 2-1.1) and applying the real translation theorem of Laplace transforms (see Chapter 2), we get

$$C(t) = K\ Am\ u(t - t_0)[1 - e^{-(t-t_0)/\tau}]$$

(7-2.8)

where the unit step function $u(t - t_0)$ is included to indicate explicitly that

$$C(t) = 0 \qquad \text{for } t \leq t_0$$

Variable C is the perturbation, or change, of the transmitter output from its initial value.

$$C(t) = c(t) - c(0)$$

(7-2.9)

A graph of Eq. 7-2.8 is shown in Fig. 7-2.5. In this figure, the term ΔC, is the steady-state change in $c(t)$. From Eq. 7-2.8, we find

$$\Delta C, = \lim_{t \to \infty} C(t) = K\ Am$$

(**7-2.10**)

From this equation, and realizing that the model response must match the process reaction curve at steady state, we can calculate the steady-state gain of the process, which is one of the model parameters.

$$K = \frac{\Delta c_s}{\Delta m}$$

(**7-2.11**)

This formula, as shown in Chapters 2 and 3, is the definition of the gain.

Three methods have been proposed for estimating the dead time t_0 and time constant τ. Each of them results in different values.
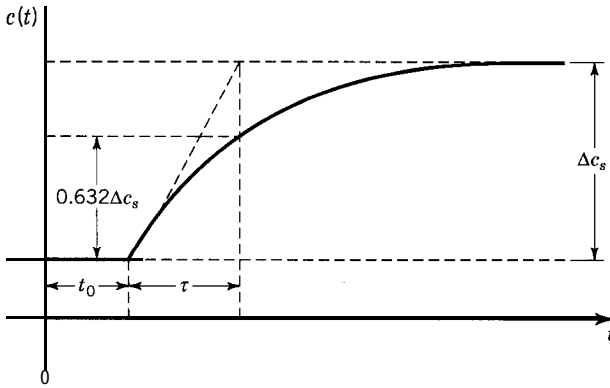
Figure 7-2.5 Step response of first-order-plus-dead-time process, showing the graphical definition of the dead time t_0 and time constant τ.

Fit 1. This method uses the line that is tangent to the process reaction curve at the point of maximum rate of change. For the FOPDT model this happens at $t = t_0$, as is evident from inspecting the model response of Fig. 7-2.5. From Eq. 7-2.8, this initial (maximum) rate of change is

$$\left.\frac{dC}{dt}\right|_{t_0} = K \, \Delta m \left[\frac{1}{\tau}\right] = \frac{\Delta c_s}{\tau} \qquad (7\text{-}2.12)$$

From Fig. 7-2.5, we see that this result tells us that the line of maximum rate of change crosses the initial value line at $t = t_0$ and the final value line at $t = t_0 + \tau$. This finding suggests the construction for determining t_0 and τ shown in Fig. 7-2.6a. The line is drawn tangent to the actual process reaction curve at the point of maximum rate of change. The model response using these values of t_0 and τ is shown by the dashed line in the figure. Evidently, the model response obtained with this fit overestimates the process time constant.
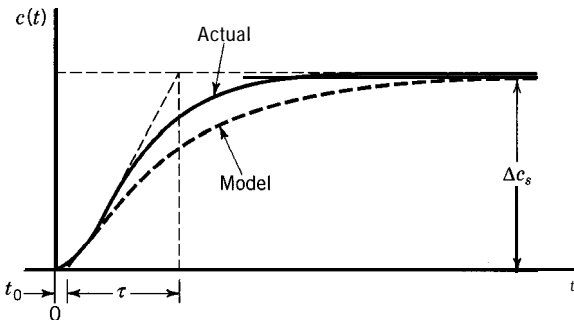
Figure 7-2.6a FOPDT model parameters by fit 1.
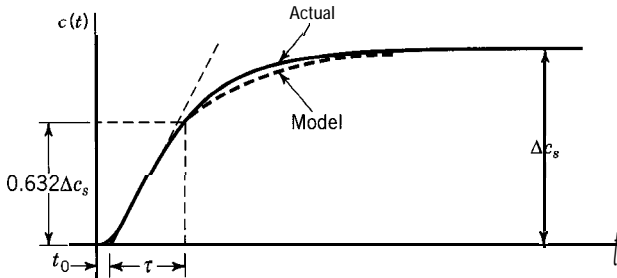
Figure 7-2.6b FOPDT model parameters by fit 2.

Fit 2. In this fit, t_0 is determined in the same manner as in fit 1, but the value of τ is the one that forces the model response to coincide with the actual response at $t = t_0 + \tau$. According to Eq. 7-2.8, this point is

$$C(t_0 + \tau) = K \text{ Am} \left[1 - e^{-1}\right] = 0.632 \text{ Ac},\tag{7-2.13}$$

Figure 7-2.6b shows that the model response for fit 2 is much closer to the actual response than is that for fit 1. The value of the time constant obtained by fit 2 is usually less than that obtained by fit 1, but the dead time is exactly the same.

Fit 3. The least precise step in the determination of t_0 and τ by the previous two methods is the drawing of the line tangent to the process reaction curve at the point of maximum rate of change. Even for fit 2, for which the value of $(t_0 + \tau)$ is independent of the tangent line, the estimated values of both t_0 and τ depend on the line. To eliminate this dependence on the tangent line, Dr. Cecil L. Smith (1972) proposes that the values of t_0 and τ be selected such that the model and actual responses coincide at two points in the region of high rate of change. The two points recommended are $(t_0 + \tau/3)$ and $(t_0 + \tau)$. The second of these points is located as in fit 2, whereas the first point is located from Eq. 7-2.8:

$$C\left(t_0 + \frac{\tau}{3}\right) = K \text{ Am} \left[1 - e^{-1/3}\right] = 0.283 \text{ AC}_s\tag{7-2.14}$$

These two points **are labeled** t_2 and t_1, respectively, in Fig. 7-2.6c. The values of t_0 and τ can then be easily obtained by the simple solution of the following set of equations:

$$t_0 + \frac{\tau}{3} = t_1$$

$$t_0 + \tau = t_2$$

<div align="right">(7-2.15)</div>
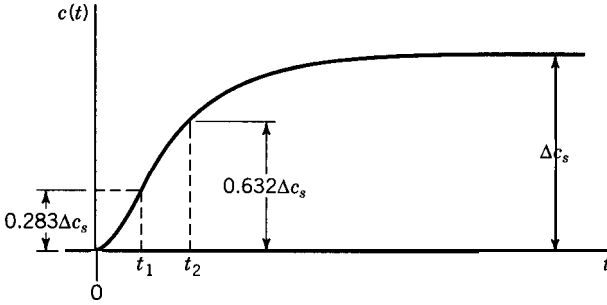
Figure 7-2.6c FOPDT model parameters by fit 3.

which reduces to

$$
\begin{aligned}
\tau &= \frac{3}{2}(t_2 - t_1) \\[2mm]
t_0' &= t_2 - \tau
\end{aligned}
$$

(7-2.16)

where

t_1 = time at which C = 0.283 Ac,

t_2 = time at which C = 0.632 Ac,

Our past experience shows that results obtained by this method are more reproducible than those obtained by the other two. We therefore recommend this method for the estimation of t_0 and τ from the process reaction curve.

Various models have been proposed in the literature for estimating the parameters of a second-order-plus-dead-time (SOPDT) model to the process reaction curve. Our experience is that the precision of these methods is very low. The reason is that the step test does not provide enough information for us to extract the additional parameters-second time constant or damping ratio-required by the SOPDT. In other words, the increased complexity of the model demands a more sophisticated dynamic test. Pulse testing is an adequate method for obtaining second and higher-order model parameters. It will be presented in Chapter 9.

Because most of the controller-tuning formulas that we are about to introduce are based on FOPDT model parameters, we may find ourselves in some situation in which we have the parameters of a high-order model and need to estimate the equivalent first-order model parameters. Although there is no general procedure for doing this, the following rule of thumb might provide a rough estimate for a first approximation:

If one of the time constants of the high-order model is much longer than the others, the effective time constant of the first-order model can be estimated to be equal to the longest time constant. The effective dead time of the first-order model

can then be approximated by the sum of all of the smaller time constant plus the dead time of the high-order model.

EXAMPLE 7-2.1

Estimate the FOPDT parameters for the temperature control loop of the exchanger of Example 6-2.1. The combined transfer function for the control valve, exchanger, and sensor/transmitter for that example is

$$G(s) = \frac{0.80}{(10s + 1)(30s + 1)(3s + 1)}$$

SOLUTION

Assuming that the 30 s time constant is much longer than the other two, we can roughly approximate

$$\tau = 30 \text{ s}$$

$$t_0 = 10 + 3 = 13 \text{ s}$$

and the gain is of course the same-that is, $K = 0.80$. The resulting FOPDT model transfer function is then

$$G(s) = \frac{0.80e^{-13s}}{30s + 1} \qquad \textbf{(Model A)}$$

We will next compare this rough approximation with the experimentally determined FOPDT parameters from the process reaction curve. Figure 7-2.7 shows the process reaction curve for the three first-order lags in series that we have assumed represent the heat exchanger, control valve, and sensor/transmitter. The response of Fig. 7-2.7 was obtained by simulating the three first-order lags on a computer, applying a 5% step change on the controller output signal, and recording the output of the sensor/transmitter versus time. From this result, we can calculate the FOPDT parameters using the three fits presented in this section.

Process Gain

$$K = \frac{\Delta c_s}{Am} = \frac{4°C}{5\%} \cdot \frac{100\%TO}{(150 - 50)°C} = 0.80 \frac{\%TO}{\%CO}$$
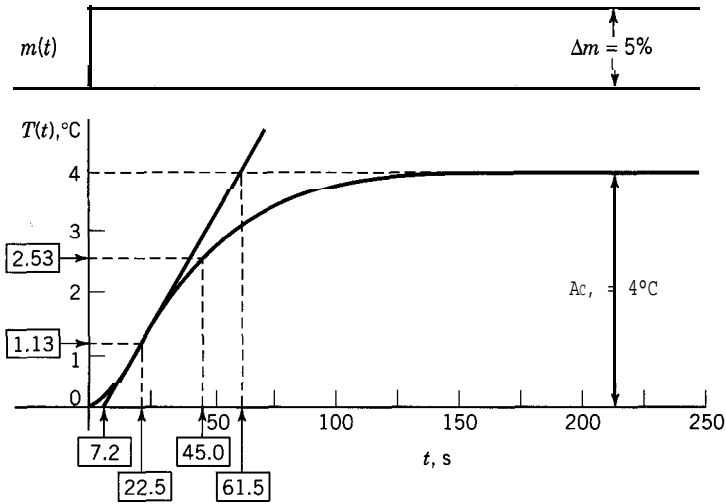
Figure 7-2.7 Step response for the heat exchanger temperature (Example 7-2.1).

Fit I

$$t_0 = 7.2 \text{ s} \qquad t_3 = 61.5 \text{ s} \qquad \text{(see Fig. 7-2.7)}$$

$$\tau = 61.5 - 7.2 = 54.3 \text{ s}$$

$$G(s) = \frac{0.80e^{-7.2s}}{54.3s + 1} \qquad \text{(Model B)}$$

Fit 2

$$t_0 = 7.2 \text{ s}$$

$$\text{At } C(t_2) = 0.632(4°\text{C}) = 2.53°\text{C} \qquad t_2 = 45.0 \text{ s}$$

$$\tau = 45.0 - 7.2 = 37.8 \text{ s}$$

$$G(s) = \frac{0.80e^{-7.2s}}{37.8s + 1} \qquad \text{(Model C)}$$

Fit 3

$$\text{At } C(t_1) = 0.283(4°\text{C}) = 1.13°\text{C} \qquad t_1 = 22.5 \text{ s}$$

$$\tau = \frac{3}{2} (t_2 - t_1) = \frac{3}{2} (45.0 - 22.5) = 33.8 \text{ s}$$

$$t_0 = 45.0 - 33.8 = 11.2 \text{ s}$$

$$G(s) = \frac{0.80e^{-11.2s}}{33.8s + 1} \qquad \text{(Model D)}$$

As we shall see in the following sections, an important parameter in terms of tuning is the ratio of the dead time to the time constant. The values for the four FOPDT model approximations are as follows:

Model	A (Rough)	B (Fit 1)	C (Fit 2)	D (Fit 3)
t_0, s	13.0	7.2	7.2	11.2
τ, s	30.0	54.3	37.8	33.8
t_0/τ	0.433	0.133	0.190	0.331

We see that the ratio t_0/τ is the most sensitive parameter, varying by a factor of slightly over 3 : 1. Recall that fits 2 and 3 provide the closest approximations to the actual step response.

EXAMPLE 7-2.2

Given a second-order process

$$G(s) = \frac{C(s)}{M(s)} = \frac{K}{(\tau_1 s + 1)(\tau_2 s + 1)} \qquad \tau_1 \geq \tau_2$$

determine the parameters of a first-order-plus-dead-time (FOPDT) model

$$G'(s) = \frac{Ke^{-t_0 s}}{\tau' s + 1}$$

using fit 3, as a function of the ratio τ_2/τ_1.

SOLUTION

To obtain the unit step response of the actual process, we solve for $C(s)$ and substitute $M(s) = 1/s$ into the transfer function.

$$C(s) = \frac{K}{(7\$ + 1)(\tau_2 s + 1)} \cdot \frac{1}{s}$$

Expanding in partial fractions, and inverting with the help of a **Laplace** transform table (Table 2-1.1), we obtain

$$C(t) = K \left[1 - \frac{\tau_1}{\tau_1 - \tau_2} e^{-t/\tau_1} + \frac{\tau_2}{\tau_1 - \tau_2} e^{-t/\tau_2} \right]$$

As $t \to \infty$, $C(t) \to K$. For fit 3, at $t_1 = t_0' + \tau'/3$,

$$C(t_1) = (1 - e^{-1/3})K = K \left[1 - \frac{\tau_1}{\tau_1 - \tau_2} e^{-t_1/\tau_1} + \frac{\tau_2}{\tau_1 - \tau_2} e^{-t_1/\tau_2} \right]$$

and at $t_2 = t_0' + \tau'$,

$$C(t_2) = (1 - e^{-1})K = K \left[1 - \frac{\tau_1}{\tau_1 - \tau_2} e^{-t_2/\tau_1} + \frac{\tau_2}{\tau_1 - \tau_2} e^{-t_2/\tau_2} \right]$$

or

$$e^{-1/3} = \frac{\tau_1}{\tau_1 - \tau_2} \left[e^{-t_1/\tau_1} - \frac{\tau_2}{\tau_1} e^{-t_1/\tau_2} \right] \qquad \textbf{(A)}$$

$$e^{-1} = \frac{\tau_1}{\tau_1 - \tau_2} \left[e^{-t_2/\tau_1} - \frac{\tau_2}{\tau_1} e^{-t_2/\tau_2} \right] \qquad \textbf{(B)}$$

Similar formulas can be obtained for the special case $\tau_1 = \tau_2$. The values of t_1 and t_2 are obtained by trial and error from these equations. Then Eq. 7-2.16 gives

$$\tau' = \frac{3}{2}(t_2 - t_1) \qquad t_0' = t_2 - \tau'$$

Martin (1975) used a computer program to solve this problem, and the results are plotted in Fig. 7-2.8. As the figure shows, the maximum effective dead time occurs when the two time constants are equal:

For $\tau_1 = \tau_2$, $t_0' = 0.5057$, $\tau' = 1.64\tau_1$

For $\tau_2 \ll \tau_1$, $t_0' \to \tau_2$ $\tau' \to \tau_1$

This is the basis for the rule of thumb presented earlier. We can use Fig. 7-2.8 to refine this rule of thumb for systems represented by three or more first-order lags in series. For instance, for the heat exchanger example, we can refine the rough approximation model as follows:
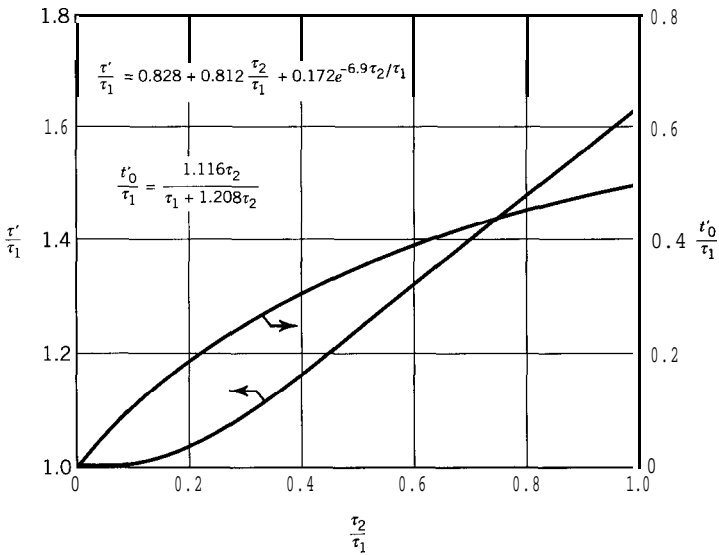
Figure 7-2.8 FOPDT model dead time and time constant for fit 3 approximation to overdamped second-order system. (Reproduced by permission of Reference 3.)

Assume $\tau_1 = 30$ s and $\tau_2 = 10 + 3 = 13s$, so that $\tau_2/\tau_1 = 13/30 = 0.433$. Then, from Fig. 7-2.8,

$$t_0' = 0.33\tau_1 = 9.9 \ s$$

$$\tau' = 1.2\tau_1 = 36 \ s$$

These values are closer to the values obtained by fit 3 in Example 7-2.1 than those obtained by the rough approximation (Model A). The reason is that in this case, τ_2 is not much smaller than τ_1.

7-2.2 Tuning for Quarter Decay Ratio Response

In addition to their on-line tuning formulas, Ziegler and Nichols (1942) proposed a set of formulas based on the parameters of a first-order model fit to the process reaction curve. These formulas are given in Table 7-2.1. Although the parameters they used were not precisely the gain, time constant, and dead time, their formulas can be modified and expressed in terms of these parameters. Ziegler and Nichols used fit 1 to determine the model parameters.

As Table 7-2.1 shows, the relative magnitudes of the gain, integral time, and derivative time among the P, PI, and PID controllers are the same as for the on-line tuning formulas that are based on the ultimate gain and period (Table 7-1.1). The formulas for the gain show that the loop gain, KK_c, is inversely proportional to the ratio of the effective dead time to the effective time constant.

Table 7-2.1 Tuning Formulas for Quarter Decay Ratio Response

Controller Type	Proportional Gain, K_c'	Integral Time, τ_I'	Derivative Time, τ_D'
Proportional-only, P	$\dfrac{1}{K}\left(\dfrac{t_0}{\tau}\right)^{-1}$	—	—
Proportional-integral, PI	$\dfrac{0.9}{K}\left(\dfrac{t_0}{\tau}\right)^{-1}$	$3.33t_0$	—
Proportional-integral-derivative, PID	$\dfrac{1.2}{K}\left(\dfrac{t_0}{\tau}\right)^{-1}$	$2.0t_0$	$\dfrac{1}{2}$ to

The PID formulas are for the actual PID controller, Eq. 5-3.19. To convert to the ideal PID controller, Eq. 5-3.17: $K_c = K_c'(1 + \tau_D'/\tau_I'); \tau_I = \tau_I' + \tau_D'; \tau_D = \tau_D'\tau_I'/(\tau_I' + \tau_D')$.

In using the formulas of Table 7-2.1, we must keep in mind that they are empirical and apply only to a limited range of ratios of dead time to time constant. In our experience, they are most applicable for a range of t_0/τ of around 0.10 to 0.5.

As we noted in the discussion of on-line tuning, the quarter decay ratio formulas can be adjusted to less oscillatory responses by simply reducing the proportional gain from the value given by the tuning formula. Unfortunately, the formulas in Table 7-2.1 give the reset and derivative times as functions of the process dead time, which cannot be estimated so readily as the ultimate period.

EXAMPLE 7-2.3

Compare the values of the tuning parameters for the temperature control of the exchanger of Figure 6- 1.1 by using the quarter decay ratio on-line tuning and the FOPDT parameters estimated in Example 7-2.1. In earlier examples we determined the following results for the exchanger temperature control loop.

By direct substitution method: (see Example 6-2.1)

$$K_{cu} = 23.8 \ \%CO/\%TO$$
$$T_u = 28.7 \ s$$

By fit 1 approximation: (see Example 7-2.1)

$$K = 0.80 \ \%CO/\%TO$$
$$t_0 = 7.2 \ s$$
$$\tau = 54.3 \ s$$

SOLUTION

The following are the tuning parameters for the quarter decay ratio.

On-Line Tuning (Table 7-1.1)   **Process Reaction Curve (Table 7-2.1)**

Proportional   Only

$$K_c = \frac{1}{2} (23.8) = 12 \frac{\%CO}{\%TO} \qquad K_c = \frac{1}{0.80} \left(\frac{7.2}{54.3} \right)^{-1} = 9.4 \frac{\%CO}{\%TO}$$

Proportional   Integral

$$K_c = \frac{23.8}{2.2} = 11 \frac{\%CO}{\%TO} \qquad K_c = \frac{0.9}{0.80} \left(\frac{7.2}{54.3} \right)^{-1} = 8.5 \frac{\%CO}{\%TO}$$

$$\tau_I = \frac{28.7}{1.2} = 24 \text{ s } (0.40 \text{ min}) \qquad \tau_I = 3.33(7.2) = 24 \text{ s } (0.40 \text{ min})$$

Proportional-Zntegral-Derivative

$$K_c' = \frac{23.8}{1.7} = 14 \frac{\%CO}{\%TO} \qquad K_c' = \frac{1.2}{0.80} \left(\frac{7.2}{54.3} \right)^{-1} = 11 \frac{\%CO}{\%TO}$$

$$\tau_I' = \frac{28.7}{2.0} = 14 \text{ s } (0.24 \text{ min}) \qquad \tau_I' = 2.0(7.2) = 14 \text{ s } (0.24 \text{ min})$$

$$\tau_D' = \frac{28.7}{8} = 3.6 \text{ s } (0.06 \text{ min}) \qquad \tau_D' = 0.5(7.2) = 3.6 \text{ s } (0.06 \text{ min})$$

The agreement is evident. Note, however, that this agreement depends on using the fit 1 model parameters, which happens to be what Ziegler and Nichols did.

7-2.3 Tuning for Minimum Error Integral Criteria

Because the quarter decay ratio tuning parameters are not unique, a substantial research project was conducted at Louisiana State University under Professors Paul W. Murrill and Cecil L. Smith to develop tuning relationships that were unique. They used the first-order-plus-dead-time (FOPDT) model parameters to characterize the process. Their specification of the closed-loop response is basically a minimum error or deviation of the controlled variable from its set point. The error is a function of time for the duration of the response, so the sum of the error at each instant of time must be minimized. This is by definition the integral of the error with time, or the shaded area in the responses illustrated in Fig. 7-2.9. Because the tuning relationships are intended to minimize the integral of the error, their use is referred to as *minimum error integral tuning.* However,

Disturbance input
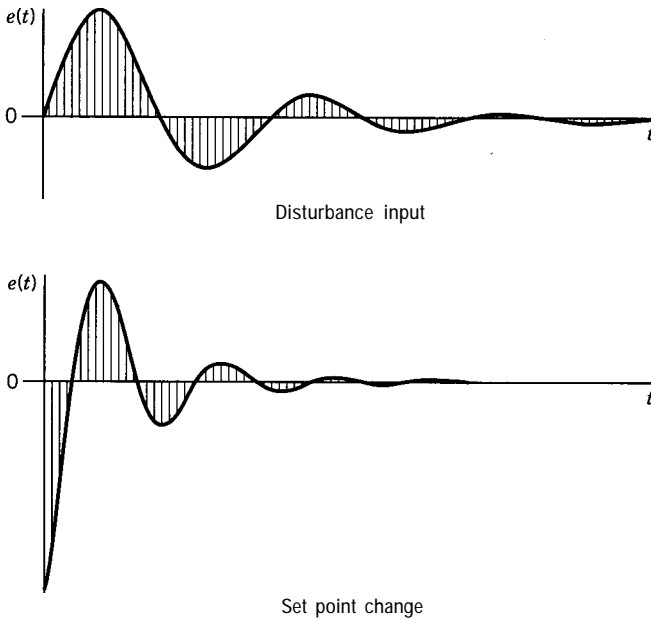
Set point change

Figure 7-2.9 Definition of error integrals for disturbance and for set
point   changes.

the integral of the error cannot be minimized directly, because a very large negative
error would be the minimum. In order to prevent negative values of the performance
function, the following formulations of the integral can be proposed.

Integral of the Absolute Value of the Error (ZAE)

$$IAE = \int_0^\infty |e(t)| \, dt \qquad (7\text{-}2.17)$$

Integral of the Square of the Error (ZSE)

$$ISE = \int_0^\infty e^*(t) \, dt \qquad (7\text{-}2.18)$$

The integrals extend from the occurrence of the disturbance or change in set point
($t = 0$) to a very long time thereafter ($t = \infty$). This is because the ending of the re-
sponses cannot be fixed beforehand. The only problem with this definition of the integral
is that it becomes undetermined when the error is not forced to zero. This happens only
when, because of offset or steady-state error, the controller does not have integral mode.
In this case, the error in the definition is replaced with the difference between the
controlled variable and its final steady-state value.

The difference between the IAE and ISE criteria is that the ISE puts more weight on
large errors, which usually occur at the beginning of the response, and less weight on
smaller errors, which happen toward the end of the response. In trying to reduce the
initial error, the minimum ISE criterion results in high controller gains and very oscil-

latory responses (i.e., a high decay ratio), with the error oscillating around zero for a relatively long time. This phenomenon suggests that the performance criteria should contain a penalty for the time elapsed from the start of the response. The following error integrals contain such a penalty by including a weight for the elapsed time.

Integral of the Time-Weighted Absolute Value of the Error (ZTAE)

$$ITAE = \int_{0}^{\infty} t I e(t) I \ dt \tag{7-2.19}$$

Zntegral of the Time-Weighted Square of the Error (ZTSE)

$$ITSE = \int_{0}^{\infty} te^2(t) \ dt \tag{7-2.20}$$

Equations 7-2.17 through 7-2.20 constitute the four basic error integrals that can be minimized for a given loop by adjusting the controller parameters. Unfortunately, the optimum set of parameter values is a function not only of which of the four integral definitions is selected but also of the type of input (that is, disturbance or set point) and of its shape (for example, step change, ramp, and so on). In terms of the shape of the input, the step change is usually selected because it is the most disruptive that can occur in practice, whereas in terms of the input type, we must select either set point or disturbance input for tuning, according to which one is expected to affect the loop more often. When set point inputs are more important, the purpose of the controller is to have the controlled variable track the set point, and the controller is referred to as a *servo regulator.* When the purpose of the controller is to maintain the controlled variable at a constant set point in the presence of disturbances, the controller is called a *regulator.* The optimum tuning parameters in terms of minimum error integral are different for each case. Most process controllers are considered regulators, except for the slave controllers in cascade control schemes, which are servo regulators. We will study cascade control in Chapter 10.

When the controller is tuned for optimum response to a disturbance input, an additional decision must be made regarding the process transfer function to the particular disturbance. This is complicated by the fact that the controller response cannot be optimum for each disturbance if there is more than one major disturbance signal entering the loop. Because the process transfer function is different for each disturbance and for the controller output signal, the optimum tuning parameters are functions of the relative speed of response of the controlled variable to the disturbance input. The slower the response to the disturbance input, the tighter the controller can be tuned-that is, the higher the controller gain can be. At the other extreme, if the controlled variable were to respond instantaneously to the disturbance, then the controller tuning would be the least tight it can be, which would be exactly equivalent to the tuning for set point changes.

Lopez *et* al. (1967) developed tuning formulas for minimum error integral criteria based on the assumption that the process transfer function to disturbance inputs is identical to the transfer function to the controller output signal. In other words, with reference to Fig. 7-2.2, they assumed that $G_2(s) = G_{,}(s)$. The tuning formulas are listed in Table 7-2.2.

Table 7-2.2 Minimum Error Integral Tuning Formulas for Disturbance Inputs

Process Model:
$$G(s) = \frac{Ke^{-t_0 s}}{\tau s + 1}$$

Proportional (P) Controller: $\quad G_e(s) = K_c$

Error Integral	ISE	IAE	ITAE
$K_c = \dfrac{a}{K}\left(\dfrac{t_0}{\tau}\right)^b$	$a = 1.411$	0.902	0.490
	$b = -0.917$	-0.985	-1.084

Proportional-Integral (PI) Controller:
$$G_c(s) = K_c\left(1 + \frac{1}{\tau_I s}\right)$$

Error Integral	ISE	IAE	ITAE
$K_c = \dfrac{a_1}{K}\left(\dfrac{t_0}{\tau}\right)^{b_1}$	$a_1 = 1.305$	0.984	0.859
	$b_1 = -0.959$	-0.986	0.977
$\tau_I = \dfrac{\tau}{a_2}\left(\dfrac{t_0}{\tau}\right)^{b_2}$	$a_2 = 0.492$	0.608	0.674
	$b_2 = 0.739$	0.707	0.680

Proportional-Integral-Derivative (PID) Controller:
$$G_c(s) = K_c\left(1 + \frac{1}{\tau_I s} + \tau_D s\right)$$

Error Integral	ISE	IAE	ITAE
$K_c = \dfrac{a_1}{K}\left(\dfrac{t_0}{\tau}\right)^{b_1}$	$a_1 = 1.495$	1.435	1.357
	$b_1 = 0.945$	-0.921	-0.947
$\tau_I = \dfrac{\tau}{a_2}\left(\dfrac{t_0}{\tau t_0}\right)^{b_2}$	$a_2 = 1.101$	0.878	0.842
	$b_2 = 0.771$	0.749	0.738
$\tau_D = a_3\tau\left(\dfrac{t_0}{\tau}\right)^{b_3}$	$a_3 = 0.560$	0.482	0.381
	$b_3 = 1.006$	1.137	0.995

These formulas indicate the same trend as the quarter decay ratio formulas, except that the integral time depends to some extent on the effective process time constant and less on the process dead time. We must again keep in mind that these formulas are empirical and should not be extrapolated beyond a range of t_0/τ of between 0.1 and 1.O. (This is the range of values used by Lopez in his correlations.) As is the case for the quarter decay ratio tuning formulas, these formulas predict that both the proportional and the integral actions go to infinity as the process approaches a first-order process without dead time. This behavior is typical of tuning formulas for disturbance inputs.

The set point tuning formulas given in Table 7-2.3 were developed by Rovira (1981), who believed that the minimum ISE criterion was unacceptable because of its highly

Table 7-2.3 Minimum Error Integral Tuning Formulas for Set Point Changes

Process Model:
$$G(s) = \frac{Ke^{-t_0 s}}{\tau s + 1}$$

Proportional-Integral (PI) Controller:
$$G_c(s) = K_c \left(1 + \frac{1}{\tau_I s} \right)$$

Error Integral	IAE	ITAE
$K_c = \dfrac{a_1}{K} \left(\dfrac{t_0}{\tau} \right)^{b_1}$	$a_1 = 0.758$ $b_1 = -0.861$	0.586 -0.916
$\tau_I = \dfrac{\tau}{a_2 + b_2(t_0/\tau)}$	$a_2 = 1.02$ $b_2 = -0.323$	1.03 -0.165

Proportional-Integral-Derivative (PID) Controller:
$$G_c(s) = K_c \left(1 + \frac{1}{\tau_I s} + \tau_D s \right)$$

Error Integral	IAE	ITAE
$K_c = \dfrac{a_1}{K} \left(\dfrac{t_0}{\tau} \right)^{b_1}$	$a_1 = 1.086$ $b_1 = -0.869$	0.965 -0.855
$\tau_I = \dfrac{\tau}{a_2 + b_2(t_0/\tau)}$	$a_2 = 0.740$ $b_2 = -0.130$	0.796 0.147
$\tau_D = a_3 \tau \left(\dfrac{t_0}{\tau} \right)^{b_3}$	$a_3 = 0.348$ $b_3 = 0.914$	0.308 0.9292

oscillatory nature. He also omitted relationships for pure proportional controllers on the assumption that the minimum error integral criteria are not appropriate for those applications for which a pure proportional controller is indicated-for example, flow averaging by proportional level control (see Section 7-3). These formulas are also empirical and should not be extrapolated beyond the range of t_0/τ between 0.1 and 1.O. They predict that for a single-capacitance (first-order) process without dead time, the integral time approaches the time constant of the process while the proportional gain goes to infinity and the derivative time to zero. These results are typical of set point tuning formulas.

EXAMPLE 7-2.4

Compare the tuning parameters that result from the various error integral criteria for disturbance inputs for the heat exchanger temperature controller using the FOPDT

model transfer function of Example 7-2.1. Consider (a) a P controller, (b) a PI controller, and (c) a PID controller.

SOLUTION

The FOPDT model parameters from Example 7-2.1 are, for fit 3,

$$K = 0.80 \ \%/\%; \qquad \tau = 33.8 \ s; \qquad t_0 = 11.2 \ s$$

The minimum error integral tuning parameters for disturbance inputs can be calculated by using the formulas from Table 7-2.2.

(a) P CONTROLLER

$$\text{ISE: } K_c = \frac{1.411}{K} \left(\frac{t_0}{\tau}\right)^{-0.917} = \frac{1.411}{0.80} \left(\frac{11.2}{33.8}\right)^{0.917} = 4.9 \ \frac{\%CO}{\%TO}$$

$$\text{IAE: } K_c = \frac{0.902}{K} \left(\frac{t_0}{\tau}\right)^{-0.985} = 3.3 \ \frac{\%CO}{\%TO}$$

$$\text{ITAE: } K_c = \frac{0.490}{K} \frac{t_0}{\tau}^{-1.084} = 2.0 \ \frac{\%CO}{\%TO}$$

(b) PI CONTROLLER

A similar application of **the** formulas from Table 7-2.2 results in the following parameters.

	Criteria		
	ISE	IAE	ITAE
K_c, %CO/%TO	4.7	3.7	3.2
τ_I, mm	0.51	0.42	0.39

(c) PID CONTROLLER

	Criteria		
	ISE	IAE	ITAE
K_c, %CO/%TO	5.3	5.0	4.8
τ_I, min	0.22	0.28	0.30
τ_D, min	0.10	0.077	0.072

The conclusion we can draw from comparing these tuning parameters is that all of these formulas result in values of the same order of magnitude, or "in the same ball

park." Note that, except for the proportional controller, the tuning parameters for the various criteria are within 50% of each other. The only noticeable difference is that the ISE formulas resulted in a lower reset-to-derivative time ratio than the other two.

EXAMPLE 7-2.5

Compare the responses to unit step changes in disturbance and in set point obtained when a PI controller is tuned for minimum IAE for disturbance inputs with those obtained for the same criteria for set point changes. The loop can be represented by the block diagram Fig. 7-2.2, and the process is modeled by the FOPDT model

$$G_1(s) = G_2(s) = \frac{1.0e^{-0.5s}}{s+1} \frac{\%TO}{\%CO}$$

where the time parameters are in minutes.

SOLUTION

The FOPDT parameters are

$$K = 1.0 \ \%TO/\%CO \qquad \tau = 1.0 \ \text{min} \qquad t_0 = 0.5 \ \text{min}$$

We can now calculate the tuning parameters for a PI controller using the formulas from Tables 7-2.2 and 7-2.3.

IAE Disturbance Criterion (Table 7-2.2)

$$K_c = \frac{0.984}{1.0}(0.5)^{-0.986} = 1.9 \ \frac{\%CO}{\%TO}$$

$$\tau_I = \frac{1.0 \ \text{min}}{0.608}(0.5)^{0.707} = 1.0 \ \text{min}$$

ZAE Set Point Criterion (Table 7-2.3)

$$K_c = \frac{0.758}{1.0}(0.50)^{-0.861} = 1.4 \ \frac{\%CO}{\%TO}$$

$$\tau_I = \frac{1.0 \ \text{min}}{1.02 - 0.323(0.50)} = 1.2 \ \text{min}$$

To calculate the responses, we must solve for the output variable in the block diagram of Fig. 7-2.2. However, the presence of the dead-time term in the FOPDT transfer function, G,(s), makes it impractical to invert the Laplace transform by partial fractions expansion. A more practical approach is to solve the loop equations on a computer. The differential equations for this loop are as follows:

FOPDT Process

$$\tau \, \frac{dc(t)}{dt} + c(t) = K[m(t - t_0) + U(t - t_0)]$$

PZD   Controller

$$m(t) = m(0) + K_c \left[e(t) + \frac{1}{\tau_1} \int e(t) \, dt + \tau_D \frac{de(t)}{dt} \right]$$

Disturbance   Input

$$U(t) = u(t) \qquad \text{(unit step)} \qquad r(t) = 0$$

Set Point Input

$$r(t) = u(t) \qquad \text{(unit step)} \qquad u(t) = 0$$

Initial   Conditions

$$c(0) = 0 \qquad\qquad m(O) = 0$$

Rovira (1981) solved these equations using a computer program that generated the response plots shown in Fig. 7-2.10. The plot in Fig. 7-2.10a is for a unit step change in disturbance; it shows that the disturbance tuning parameters result in a slightly smaller initial deviation and a faster return to the set point than the set point tuning parameters. Figure 7-2. 10b is for a unit step change in set point and shows that the set point tuning parameters result in a significantly smaller overshoot, less oscillatory behavior, and a shorter settling time than the disturbance tuning parameters. As would be expected, each set of tuning parameters performs best for the input for which it is designed. The
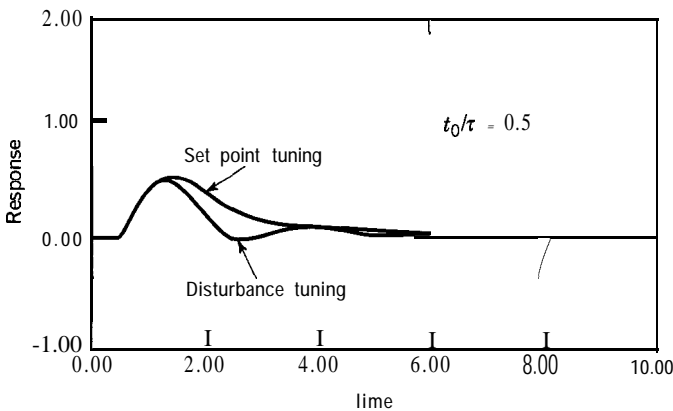
Figure 7-2.10a Response to a step change in disturbance with PI controller tuned for minimum IAE criteria. (Reproduced by permission of Reference 6.)
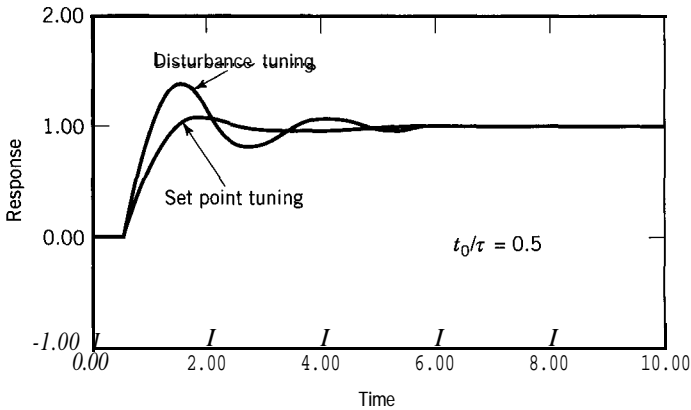
Figure **7-2.10b** Response to a step change in set point with PI controller tuned for minimum IAE criteria. (Reproduced by permission of Reference 6.)

responses obtained are a direct result of the higher gain and shorter reset time obtained with disturbance tuning.

7-2.4   Tuning   Sampled-Data   Controllers

In industry today, most control functions are implemented using microprocessors (distributed controllers) and digital computers. A common characteristic of these installations is that the control calculations are performed at regular intervals of time T, the *sample time*. This is in contrast to analog (electronic and pneumatic) instruments that perform their functions continuously with time. Sampling is also characteristic of some analyzers, such as on-line gas chromatographs.

The discrete mode of the operation characteristic of computers requires that at each sampling instant, the transmitter signals be sampled, the value of the manipulated variable be calculated, and the controller output signal be updated. The output signals are then held constant for a full sample time until the next update. This is illustrated in Fig. 7-2.11. As might be expected, this sampling and holding operation has an effect on the performance of the controller and thus on its tuning parameters.

The sampling time of computer controllers varies from about one-third of a second to several minutes, depending on the application. *A good rule of thumb is that the sample time should be about one-tenth of the effective process time constant.* When the sample time is of this order of magnitude, its effect can be taken into consideration in the tuning formulas by adding one-half the sample time to the process dead time and then using this corrected dead time in the tuning formulas for continuous controllers (Tables 7-2.1, 7-2.2, and 7-2.3). This method, proposed by Moore *et al.* (1969), holds that the dead time to use in the tuning formulas is

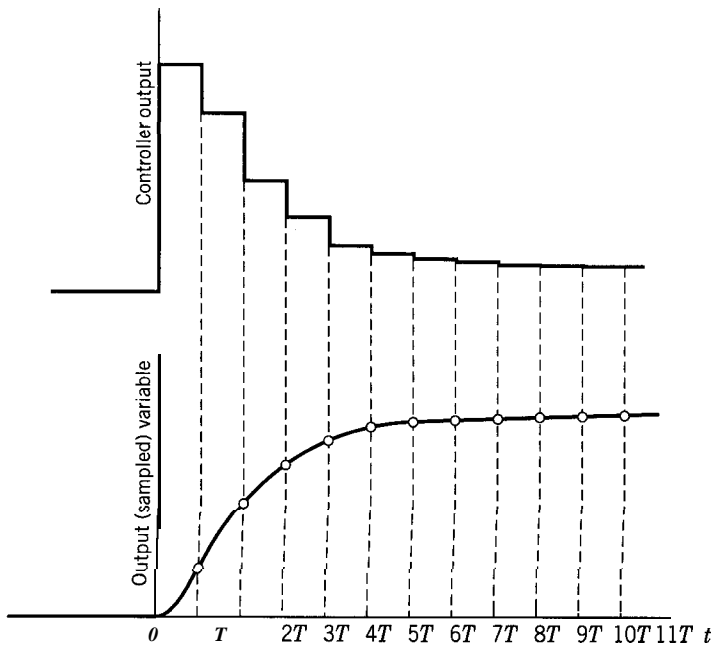$$t_{0c} = t_0 + \frac{T}{2} \qquad\qquad (7\text{-}2.21)$$

Figure 7-2.11 The output of a sampled-data controller is updated at uniform intervals of time T and held constant between updates.

where

t_{0c} = the corrected dead time
t_0 = the dead time of the process
T = the time interval between samples (sample time)

Note that the on-line tuning method inherently incorporates the effect of sampling when the ultimate gain and period are determined for the loop with the sampled-data controller in automatic.

Tuning formulas that are specific for sampled-data controllers have been developed by Chiu et al. (1973) and reproduced by Corripio (1990). They will be presented in Chapter 15.

7-2.5 Summary of Controller Tuning

We have thus far presented two methods for measuring the dynamic characteristics of the process in a feedback control loop: the ultimate gain method and the step test or process reaction curve. We have also presented one set of tuning formulas for the ultimate gain method and three sets of formulas for the first-order-plus-dead-time model parameters. For a given process, all four sets of tuning formulas result in controller parameters that are in the same ball park. These tuning parameters-are just starting values that must be adjusted in the field so that the true "personality" of the specific process can be matched by the controller. We must reiterate a point made at the beginning of this chapter. As discussed in Chapters 3 and 4, most processes are nonlinear, and their dynamic characteristics (e.g., ultimate gain and frequency, FOPDT model

parameters) vary from one operating point to another. It follows that the controller parameters arrived at by the tuning procedure are at best a compromise between slow behavior at one end of the operating range and oscillatory behavior at the other. In short, tuning is not an exact science. However, we must also keep in mind that the tuning formulas offer us insight into how the various controller parameters depend on such process parameters as the gain, the time constant, and the dead time.

7-3 TUNING CONTROLLERS FOR INTEGRATING PROCESSES

Integrating processes represent a special tuning case in that the process cannot be represented by the standard first- or second-order models presented in the preceding section. This is because the process is not self-regulating, that is, it will not reach a steady state if driven by a sustained disturbance with the loop opened. As a further consequence, feedback control is absolutely required to operate an integrating process, and open-loop step tests can be carried out only for very brief periods of time. By far the most common integrating process is the control of liquid level.

A special feature of liquid level control is that there are two opposite specifications of control loop performance: *tight level control* and *averaging level control.* Tight level control requires that the level be kept at or very near its set point, as in natural circulation evaporators and reboilers, because of the large sensitivity of the heat transfer rate on the level. Averaging level control is specified for surge tanks and accumulators, where the objective is to average out, or attenuate, variations in inlet flow so that the outlet flow does not vary suddenly. An intermediate specification is required for reactors and similar equipment where the objective of controlling the level is to keep the volume of liquid in the tank approximately constant. In such applications, it is acceptable to allow the level to vary about $\pm$ 5% around its set point, which is "looser" than for an evaporator or reboiler.

In this section, we will use level control as an example of an integrating process. We will derive a model of a simple tank level control and show that, in many cases, it is possible to compute the tuning parameters from the process design parameters.

7-3.1   Model of Liquid Level Control System

Liquid level control is one of the few continuous processes that can be treated as an integrating process. Figure 7-3.1 shows a schematic of the level control loop with the control valve on the discharge of the pump that draws the liquid from the tank. In Section 4-1.1, similar level processes were modeled as self-regulating because of the effect of the level on the pressure drop across the valve and, through it, on the outlet flow. However, this self-regulating effect is usually negligible, especially when the pressure drop across the valve is provided by a pump, as in this case. (A numerical demonstration of this fact appears as a problem at the end of this chapter.)

To model the response of the level in the tank, we write a mass balance around the tank.

$$A \frac{dh(t)}{dt} = f_i(t) - f_o(t) \tag{7-3.1}$$

where A is the cross-sectional area of the tank, $h(t)$ is the level in the tank, $f_i(t)$ is the
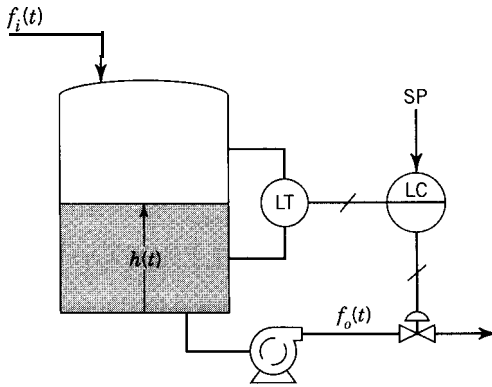
Figure 7-3.1 Schematic of a level control loop with manipulation of the outlet stream.

total flow into the tank, $f_o(t)$ is the flow out of the tank, and we are assuming consistent units. Subtracting the initial steady state, and Laplace-transforming the resulting equation in terms of the deviation variables, we obtain

$$H(s) = \frac{1}{As} [F_i(s) - F_o(s)] \qquad (7\text{-}3.2)$$

1 eqn. 2 unk. *(H, F,)*

Assuming the outlet flow is a function only of the valve position, and modeling the valve as a first-order lag (see Section 5-2), we get

$$F_o(s) = \frac{K_v}{\tau_v s + 1} M(s) \qquad (7\text{-}3.3)$$

2 eqn. 3 unk. *(M)*

where K_v is the valve gain, τ_v is the valve time constant, and M(s) is the controller output signal, %CO. Level transmitters are usually very fast and can be modeled as simple gains:

$$C(s) = K_T H(s) \qquad (7\text{-}3.4)$$

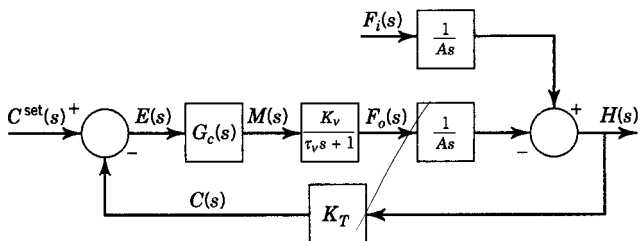*3* eqn. 4 unk. (C)

where C(s) is the transmitter output signal, %TO, and K_T is the transmitter gain, from Section 5-1,

$$K_T = \frac{100}{h_{max} - h_{min}} \qquad (7\text{-}3.5)$$

where h_{max} and h_{min} are the limits of the transmitter range.

A level controller is usually calibrated in percent of transmitter output (%TO), be-

Figure 7-3.2a Block diagram of the level control loop of Figure 7-3.1.

cause this value tells how full the tank is without our having to know the actual limits on the level transmitter. Because of this, the set point and the level are displayed in %TO. In terms of these variables, the controller output is

$$M(s) = G_c(s)[C^{set}(s) - C(s)] \qquad (7\text{-}3.6)$$

4 eqn. 4 unk.

where $C^{set}(s)$ is the set point in %TO, $G_c(s)$ is the controller transfer function, and we have followed our standard sign convention in calculating the error.

This completes the model of the level control loop. Figure 7-3.2a shows the detailed block diagram of the loop, Fig. 7-3.2b the simplified diagram. The two process gains shown in the latter diagram are

$$K = \frac{K_v K_T}{A} \cdot \frac{\%TO}{\%CO\text{-min}} \qquad (7\text{-}3.7)$$

$$K_u = \frac{K_T}{A}$$

Note that these formulas assume consistent units. For example, for K_v in **gpm/%CO**, and K_T in **%TO/ft**, A must be in **gal/ft**, and the inlet flow must also be in gpm so that K_u can be in **%TO/gal**.
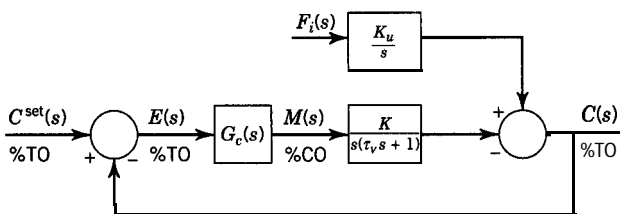
Figure 7-3.2b Simplified block diagram of the level control loop of Figure 7-3.1.

From the block diagram of Fig. 7-3.2b, after some algebraic manipulation, we can derive the following closed-loop transfer function for the level.

$$C(s) = \frac{-KG_c(s)}{s(\tau_v s + 1) \quad KG_c(s)} C^{\text{set}}(s) + \frac{K_u(\tau_v s + 1)}{s(\tau_v s + 1) \quad KG_c(s)} F_i(s) \quad (7\text{-}3.8)$$

where the minus signs indicate that we have a positive feedback loop, which means the controller must have a negative gain (direct acting). We can arrive at exactly the same conclusion by analyzing Fig. 7-3.1. An increase in level must cause the controller output to increase to open the outlet valve and increase the outlet flow. This demands a **direct-acting** controller.

7-3.2 Proportional Level Controller

Many level controllers, whether tuned for tight or averaging level control, are proportional controllers. As a matter of fact, we will see in the following section that controller synthesis for an integrating process results in a proportional controller. Let us then look at a proportional controller for the level control loop.

To look at a direct-acting proportional controller, we substitute its transfer function, $G_c(s) = -K_c$, into Eq. 7-3.8 to obtain

$$C(s) = \frac{KK_c}{s(\tau_v s + 1) + KK_c} C^{\text{set}}(s) + \frac{K_u(\tau_v s + 1)}{s(\tau_v s + 1) + KK_c} F_i(s) \quad (7\text{-}3.9)$$

Because a proportional controller usually results in offset or steady-state error, let us obtain an expression for the offset before we move into controller tuning. To obtain the offset caused by a sustained change in set point, Δc^{set}, and a sustained change in inlet flow, Δf_i, we let $s = 0$ in Eq. 7-3.9 to obtain

$$\Delta c = \Delta c^{\text{set}} + \frac{K_u}{KK_c} \Delta f_i$$

where Δc is the resulting steady-state change in level %TO. The offset is then the steady-state error.

$$e = \Delta c^{\text{set}} - \Delta c = -\frac{K_u}{KK_c} \Delta f_i = -\frac{1}{K_v K_c} \Delta f_i \quad \textbf{(7-3.10)}$$

where we have substituted Eq. 7-3.7. This result shows that for an integrating process, there is no offset for changes in set point, only for changes in flow. The offset is inversely proportional to the controller gain K_c.

The characteristic equation of the loop is obtained by setting the denominator of Eq. 7-3.9 equal to zero.

$$\tau_v s^2 + s + KK_c = 0$$

The roots of this equation can be obtained directly by the quadratic formula and rearranged into the following form:

$$r_{1,2} = \frac{-1 \pm \sqrt{1 - 4\tau_v K K_c}}{2\tau_v} \tag{7-3.11}$$

This expression tells us that the roots of the characteristic equation are real and negative as long as the gain is limited to

$$0 < K_c < \frac{1}{4\tau_v K} = \frac{A}{4\tau_v K_v K_T} \tag{7-3.12}$$

where we have substituted the value of K from Eq. 7-3.7. The expression given in Eq. 7-3.12 represents the maximum proportional controller gain that will result in a non-oscillatory response. We notice, from Eq. 7-3.11, that if the controller gain is increased beyond this maximum value, then the response becomes oscillatory, but it cannot be unstable, no matter how high the controller gain, because the real part of the complex roots is always negative.

If the controller gain is set equal to the expression of Eq. 7-3.12, then the closed-loop transfer function will have two equal roots equal to $-1/2\tau_v$. The effective time constants of the closed loop will then be equal to each other and to twice the valve time constant. Because most valves are very fast, this means that a proportional level controller can be tuned to give a very fast non-oscillatory response.

EXAMPLE 7-3.1

Calculate the maximum gain of a proportional level controller that results in a non-oscillatory response. The level transmitter has a range of 2 to 10 ft above the bottom of a distillation column 8.0 ft in diameter. The design flow of the bottoms product is 500 gpm, and the control valve on that line is linear and is sized for 100% overcapacity. The time constant of the valve is 3 s (0.05 min). Calculate also the effective time constant of the closed loop and the offset caused by a sustained change of 100 gpm in flow.

SOLUTION

Assuming constant pressure drop across the valve, the valve gain, from Section 5-2, is

$$K_v = \frac{f_{o,\text{max}}}{100} = \frac{2(500)}{100} = 100 \frac{\text{gpm}}{\%\text{CO}}$$

The transmitter gain, from Eq. 7-3.5, is

$$K_T = \frac{100}{10 - 2} = 12.5 \ \frac{\% \ TO}{ft}$$

The area of the tower is

$$A = \frac{\pi(8.0)^2}{4} \ 7.48 \ gal/ft^3 = 376 \ gal/ft$$

The maximum controller gain that will produce non-oscillatory response is, from Eq. 7-3.12,

$$K_c = \frac{A}{4\tau_v K_v K_T} = \frac{376}{4(0.05)(10.0)(12.5)} = 15.0 \ \frac{\%CO}{\% \ TO}$$

This represents a proportional band of less than 7%. The effective time constants of the closed loop, when the gain is set equal to this value, are $2(0.05) = 0.10$ min, or 6 s, each; this is very fast, considering the size of the column.

Finally, the offset caused by a 100-gpm change in flow is, from Eq. 7-3.10,

$$e = -\frac{1}{K_v K_c} \Delta f_i = -\frac{1}{(10.0)(15.0)} \ 100 = -0.66 \ \%TO$$

This imperceptible change is less than 1% of the transmitter range.

The preceding example shows that a proportional level controller can be tuned for non-oscillatory response from the design parameters of the process. Obviously this tuning is for tight level control-that is, for reducing variations in level. Let us next look at averaging level control.

7-3.3 Averaging Level Control

The purpose of an averaging level controller is to average out sudden variations in the disturbance flows so as to produce a smoothly varying manipulated flow. For example, if the tank of Fig. 7-3.1 were a surge tank on the feed to a continuous distillation column, it would be very desirable that the column not be subjected to sudden variations in flow, because this could cause flooding and upset the product compositions.

A proportional controller is ideal for averaging level control, but obviously we would like its gain to be as low as possible so that it lets the level in the tank and absorb the variations in disturbance flows. How low can the gain be? The minimum controller gain is the gain that prevents the level from exceeding the range of the level transmitter

at any time. To derive this minimum controller gain, let us recall the formula for a proportional controller, Eq. 5-3.4:

$$m(t) = \bar{m} + K_c e(t) \qquad\qquad \textbf{(5-3.4)}$$

If we set $\bar{m}$ equal to 50 %CO, then the control valve will be exactly half opened when the level is at the set point. If we further set the set point to 50 %TO, then the maximum value of the error is ± 50 %TO, because the transmitter can read only from 0 to 100 %TO. From the foregoing equation, we see that the minimum gain that prevents the level from exceeding the limits of the transmitter is 1 .O %CO/%TO (100 %PB). If the gain were lower than unity, the level would have to exceed the transmitter range for the valve either to open fully ($m = 100\%$) or to close fully ($m = 0\%$). For gains greater than unity, the valve can reach one of its limits before the level reaches the corresponding limit on the transmitter range, but then the manipulated flow will vary more than necessary. In summary,

> *The ideal averaging level controller is a proportional controller with the set point at 50 %TO, the output bias at 50 %CO, and the gain set at 1 .0 %CO/%TO.*

7-3.4 Summary

In this section we explored the interesting problem of tuning proportional controllers for integrating processes. We developed a formula, Eq. 7-3.12, that provides the maximum controller gain resulting in a non-oscillatory response. We also found that the minimum gain that prevents the controlled variable from exceeding the transmitterrange limits is unity. Gains near the maximum are to be used for tight level controller tuning, whereas the minimum gain should be used for averaging level control. A problem at the end of this chapter explores the tuning of proportional-integral (PI) level controllers.

7-4 SYNTHESIS OF FEEDBACK CONTROLLERS

In the preceding sections, we have taken the approach of tuning a feedback controller by adjusting the parameters of the proportional-integral-derivative (PID) control structure. In this section, we will take a different approach to controller design, that of controller synthesis, which is performed as follows:

> *Given the transfer functions of the components of a feedback loop, synthesize the controller required to produce a specific closed-loop response.*

Although we get no assurances that the controller resulting from our synthesis procedure can be built in practice, we stand to gain some insight into the selection of the various controller modes and their tunings.

7-4.1 Development of the Controller Synthesis Formula

Let us consider the simplified block diagram of Fig. 7-4.1, in which the transfer functions of all the loop components other than the controller have been lumped into
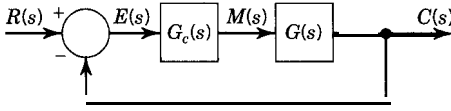
Figure 7-4.1 Simplified block diagram for controller synthesis.

a single block, G(s). From block diagram algebra, the transfer function of the closed loop is

$$\frac{C(s)}{R(s)} = \frac{G_c(s)G(s)}{1 + G_c(s)G(s)} \tag{7-4.1}$$

Next we use this expression to solve for the controller transfer function.

$$G_c(s) = \frac{1}{G(s)} \cdot \frac{C(s)/R(s)}{1 - C(s)/R(s)} \tag{7-4.2}$$

This is the *controller synthesis formula.* It gives us the controller transfer function G,(s) from the process transfer function G(s) and the specified closed-loop response C(s)/R(s). In order to illustrate how this formula is used, consider the specification of perfect control-that is, C(s) = R(s) or $C(s)/R(s) = 1$. The resulting controller is

$$\mathbf{G,(s)} = \frac{1}{G(s)} \cdot \frac{1}{1} = G(s) \cdot \frac{1}{0} \tag{7-4.3}$$

This says that in order to force the output to equal the set point at all times, the controller gain must be infinite. In other words, perfect control cannot be achieved with feedback control. This is because any feedback corrective action must be based on an error.

The controller synthesis formula, Eq. 7-4.2, results in different controllers for different combinations of closed-loop response specifications and process transfer functions. Let us look at each of these elements in turn.

7-4.2 Specification of the Closed-Loop Response

The simplest achievable closed-loop response is a first-order lag response. In the absence of process dead time, this response is the one shown in Fig. 7-4.2 and results from the closed-loop transfer function

$$\frac{C(s)}{R(s)} = \frac{1}{\tau_c s + 1} \tag{7-4.4}$$

where τ_c is the time constant of the closed-loop response and, being adjustable, becomes the single *tuning parameter* for the synthesized controller; the shorter τ_c, the tighter the controller tuning. *Note:* This response was originally proposed by Dahlin (1968), who defined the tuning parameter as the reciprocal of the closed-loop time constant, A = $1/\tau_c$. In this book we will use τ_c.
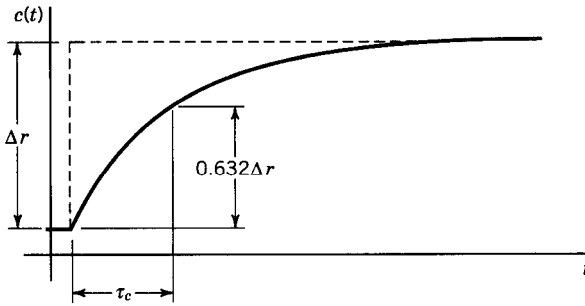
Figure 7-4.2 First-order closed-loop response specification for synthesized controller.

Substituting Eq. 7-4.4 into Eq. 7-4.2, we obtain

$$G_s(s) = \frac{1}{G(s)} \cdot \frac{\dfrac{1}{\tau_c s + 1}}{1 - \dfrac{1}{\tau_c s + 1}} = \frac{1}{G(s)} \cdot \frac{1}{\tau_c s + 1 - 1}$$

or

$$\boxed{G_c(s) = \frac{1}{G(s)} \cdot \frac{1}{\tau_c s}} \qquad (7\text{-}4.5)$$

We can see that this controller has integral mode, which results from the specification of unity gain in the closed-loop transfer function, Eq. 7-4.4. This ensures the absence of offset.

Although second- and higher-order closed-loop responses could be specified, it is seldom necessary to do so. However, when the process contains dead time, the closed-loop response must also contain a dead-time term, with the dead time equal to the process dead time. We will look at this shortly, but first let us see how controller synthesis can guide us in the selection of controller modes for various process transfer functions.

7-4.3   Controller Modes and Tuning Parameters

Controller synthesis allows us to establish a relationship between the process transfer function and the modes of a PID controller. This is so because, for simple transfer functions without dead time, the synthesized controller can be expressed in terms of the proportional, integral, and derivative modes. Controller synthesis also provides us with relationships for the controller tuning parameters in terms of the closed-loop time constant, τ_c, and the parameters of the process transfer function. In what follows, we will derive these relationships by substituting process transfer functions of increasing complexity into Eq. 7-4.5.

Instantaneous   Process   Response

$$G(s) = K$$

From Eq 7-4.5,

$$G_I(s) = \frac{1}{K\tau_c} \cdot \frac{1}{s} \tag{7-4.6}$$

where **K** is the process gain.

This is a **pure integral** controller, which is indicated for very fast processes such as flow controllers, steam turbine governors, and the control of outlet temperatures from reformer   furnaces.

First-Order Process

$$G(s) = \frac{K}{\tau s + 1}$$

From Eq. 7-4.5,

$$G_c(s) = \frac{\tau s + 1}{K} \cdot \frac{1}{\tau_c s}$$

$$= \frac{\tau}{K\tau_c}\left(1 + \frac{1}{\tau s}\right) \tag{7-4.7}$$

where τ is the process time constant.

This is a proportional-integral (PI) controller with tuning parameters

$$\boxed{K_c = \frac{\tau}{K\tau_c} \qquad \tau_I = \tau} \tag{7-4.8}$$

or, in words, the integral time is set equal to the process time constant, and the proportional gain is adjustable or tunable. Note that if the process time constant τ is known, tuning is reduced to the adjustment of a single parameter, the controller gain. This is because the tuning parameter τ_c affects only the controller gain.

Second-Order Process

$$G(s) = \frac{K}{(\tau_1 s + 1)(\tau_2 s + 1)}$$

From Eq. 7-4.5,

$$G(s) = \frac{(\tau_1 s + 1)(\tau_2 s + 1)}{K} \cdot \frac{1}{\tau_c s}$$

$$= \frac{\tau_1}{K \tau_c}\left(1 + \frac{1}{\tau_1 s}\right)(\tau_2 s + 1) \qquad \text{(7-4.9)}$$

where

τ_1 = the longer or dominant process time constant
τ_2 = the shorter process time constant

Equation 7-4.9 matches the transfer function of the actual PID controller discussed in Chapter 5, Eq. 5-3.19, ignoring the noise filter term ($\alpha \tau_D' s + 1$).

$$G_c(s) = K_c'\left(1 + \frac{1}{\tau_I' s}\right)\left(\frac{\tau_D' s + 1}{\alpha \tau_D' s + 1}\right) \qquad \text{(7-4.10)}$$

The tuning parameters are then

$$\boxed{K_c' = \frac{\tau_1}{K \tau_c} \qquad \tau_I' = \tau_1 \qquad \tau_D' = \tau_2} \qquad \text{(7-4.11)}$$

Again the tuning procedure is reduced to adjustment of the process gain with the integral time set equal to the longer time constant and the derivative time set equal to the shorter time constant. This is dictated by experience, which indicates that the derivative time should always be smaller than the integral time. In industrial practice, PID controllers are commonly used for temperature control loops so that the derivative mode can compensate for the sensor lag. Here we have arrived at this same result by controller synthesis.

We can easily see that a third-order process would demand a second derivative term in series with the first and with its time constant set to the third-longest process time constant, and so on. A reason why this idea has not caught on in practice is that the controller would be very complex and expensive. Besides, the values of the third and subsequent process time constants are very difficult to determine. The common practice has been to approximate high-order processes with low-order-plus-dead-time models. Let us next synthesize the controller for such an approximation of the process transfer function.

First-Order-Plus-Dead-Time Process

$$G(s) = \frac{K e^{-t_0 s}}{\tau s + 1}$$
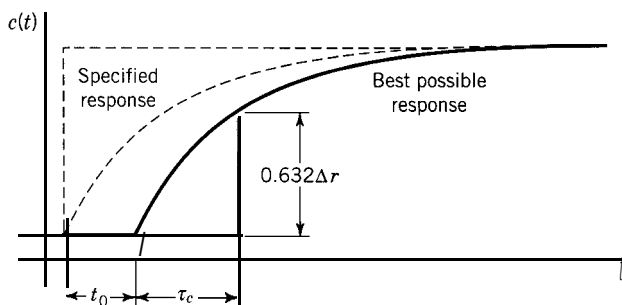
Figure 7-4.3 Specification for closed-loop response of system with dead time t_0.

From Eq. 7-4.5,

$$G_c(s) = \frac{\tau s + 1}{Ke^{-t_0 s}} \cdot \frac{1}{\tau_c s} = \frac{\tau}{K\tau_c}\left(1 + \frac{1}{\tau s}\right) e^{t_0 s} \tag{7-4.12}$$

where t_0 is the process dead time.

We note immediately that this an **unrealizable** controller because it requires knowledge of the future-that is, a negative dead time. This is even more obvious when the specified and the best possible closed-loop responses are graphically compared, as in Fig. 7-4.3. It is evident from this comparison that the specified response must be delayed by one process dead time.

$$\frac{C(s)}{R(s)} = \frac{e^{-t_0 s}}{\tau_c s + 1} \tag{7-4.13}$$

This results in the following synthesized controller transfer function:

$$G_c(s) = \frac{\tau s + 1}{Ke^{-t_0 s}} \cdot \frac{e^{-t_0 s}}{\tau_c s + 1 - e^{-t_0 s}}$$

or

$$G_c(s) = \frac{\tau s + 1}{K} \cdot \frac{1}{\tau_c s + 1 - e^{-t_0 s}} \tag{7-4.14}$$

Although this controller is now realizable in principle, its implementation is far from common practice. This is largely because the original PID controllers were implemented with analog components, and the term $e^{-t_0 s}$ cannot be implemented in practice with analog devices. Modem implementation of PID controllers on microprocessors and digital computers makes it possible to implement the dead-time term. When this is done, the term is called a predictor or dead-time compensation term. Chapter 15 presents the development of dead-time compensation controllers.

In order to convert the algorithm of Eq. 7-4.14 to the standard PI form, we expand

the exponential term by the first-order Padé approximation presented earlier, Eq. 6-2.6:

$$e^{-t_0 s} \doteq \frac{1 - \dfrac{t_0}{2} s}{1 + \dfrac{t_0}{2} s} \tag{7-4.15}$$

Substituting Eq. 7-4.15 into Eq. 7-4.14 and simplifying, we obtain the following synthesized controller:

$$G_c(s) = \frac{\tau}{K(\tau_c + t_0)} \left(1 + \frac{1}{\tau s}\right) \left(\frac{1 + \dfrac{t_0}{2} s}{1 + \tau' s}\right) \tag{7-4.16}$$

where

$$\tau' = \frac{\tau_c t_0}{2(\tau_c + t_0)}$$

This is equivalent to an actual PID controller, Eq. 7-4.10 with tuning parameters

$$\boxed{K_c' \quad \frac{\tau}{\overline{K}(\tau_c + t_0)} \qquad \tau_I' = \tau \qquad \tau_D' = \frac{t_0}{2}} \tag{7-4.17}$$

Although a lag term is present in the transfer function of the actual controller to prevent high-frequency noise amplification, the time constant τ' is usually fixed and much shorter than τ_D. In order to interpret the meaning of the term $1 + r's$, we first note that for small dead time $(t_0 \ll \tau_c)$,

$$\tau' = \frac{t_0}{2} \tag{7-4.18}$$

Substitution of this equation into Eq. 7-4.16 results in a PI controller $(\tau_D = 0)$. This suggests that a PI controller is indicated when the dead time is short. For long dead time and tight control $(\tau_c \rightarrow 0)$, the value of τ' becomes

$$\tau' = \frac{\tau_c}{2} \rightarrow 0 \tag{7-4.19}$$

Therefore, for long dead time, the tighter the control, the closer the synthesized algorithm, Eq. 7-4.16, is to the actual PID controller of Eq. 5-3.19 with the tuning parameters of Eq. 7-4.17.

It is interesting to note that the derivative time of Eq. 7-4.17 is exactly the same as the value from the Ziegler-Nichols quarter decay ratio formulas (see Table 7-2.1). However, the proportional gain for quarter decay ratio is 20% higher than the maximum synthesis gain $(\tau_c = 0)$, and the integral time of the synthesis formula is related to the

model time constant, whereas that of the quarter decay ratio formula is related to the model dead time.

Note that the tuning relationship of Eq. 7-4.17 indicates that an increase in dead time results in a reduction of the controller gain for a given closed-loop time constant specification. When we compare Eqs. 7-4.8 and 7-4.17, we see that the presence of dead time imposes a limit on the controller gain. In other words, for the first-order process without dead time (Eq. 7-4.8), the gain can be increased without limit to obtain faster and faster responses ($\tau_c \rightarrow 0$). However, for the process with an effective dead time, from Eq. 7-4.17 we have the following limit on the controller gain:

$$K_{c,\max} = \lim_{\tau_c \rightarrow 0} \frac{\tau}{K(\tau_c + t_0)} = \frac{\tau}{Kt_0}$$

The closed-loop response will deviate from the specified first-order response as the controller gain is increased. That is, increasing the gain will eventually result in overshoot and even instability of the closed-loop response. This is because the error of the first-order Padé approximation increases with the speed of response, as s increases with speed. (Recall that s, the Laplace transform variable, has units of reciprocal time or frequency. Thus higher speeds of response, or frequencies, correspond to a higher magnitude of s.)

Integrating Process

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

From Eq. 7-4.5,

$$G_c(s) = \frac{s}{K} \cdot \frac{1}{\tau_c s} = \frac{1}{K\tau_c} \qquad (7\text{-}4.20)$$

This a proportional controller with an adjustable gain, which agrees with the discussion of Section 7-3 on tuning controllers for integrating processes.

We have now synthesized controllers for the most common process transfer functions. The synthesis of a controller for processes with inverse response is left as an exercise at the end of this chapter.

The same results obtained here by controller synthesis have been obtained via the technique of internal model control (IMC) by Rivera *et al.* (1986). In some articles, the tuning formulas we have developed by synthesis are called IMC Tuning Rules.

7-4.4 Summary of Controller Synthesis Results

Table 7-4.1 summarizes the controller modes and tuning parameters that result from the synthesis procedure for Dahlin's response. The fact that the controller gain is a function of the tuning parameter τ_c is both an advantage and a disadvantage of the tuning formulas derived by the synthesis procedure. It is an advantage in that it allows the engineer to achieve a specified response by adjusting a single parameter, the gain, regardless of the controller modes involved. The tunable gain is a disadvantage, how-

Table 7-4.1 Controller Modes and Tuning Formulas for Dahlin Synthesis

Process	Controller	Tuning Parameters
$G(s) = K$	I	$K_c = \dfrac{1}{K\tau_c}$
$G(s) = \dfrac{K}{\tau s + 1}$	PI	$K_c = \dfrac{\tau}{K\tau_c}$
		$\tau_I = \tau$
$G(s) = \dfrac{K}{(\tau_1 s + 1)(\tau_2 s + 1)}, \ \tau_1 > \tau_2$	PID	$K'_c = \dfrac{\tau_1}{K\tau_c}$
		$\tau'_I = \tau_1$
		$\tau'_D = \tau_2$
$G(s) = \dfrac{Ke^{-t_0 s}}{\tau s + 1}$	PID"	$K'_c = \dfrac{\tau}{K(t_0 + \tau_c)}$
		$\tau'_I = \tau$
		$\tau'_D = \dfrac{\tau_0}{2}$
$G(s) = \dfrac{K}{s}$	P	$K_c = \dfrac{1}{K\tau_c}$

[a] These formulas apply to both PID and PI ($\tau_D = 0$) controllers. PID is recommended when $t_0 > \tau/4$. The PID formulas are for the actual PID controller, **Eq.** 5-3.19.

ever, because the formulas do not provide a "ball-park" value for it. The following guidelines are given in order to remedy this situation.

Minimum IAE. For **disturbance inputs,** $\tau_c = 0$ approximately minimizes the IAE when t_0/τ is in the range 0.1 to 0.5 for PI controllers ($\tau_D = 0$) and 0.1 to 1.5 for PID controllers. For **set point changes,** the following formulas result in approximately minimum IAE when t_0/τ is in the range 0.1 to 1.5.

PI controller $(\tau_D = 0)$: $\qquad \tau_c = \dfrac{2}{3} t_0$ $\qquad\qquad\qquad$ (7-4.21)

PID controller: $\qquad\qquad \tau_c = \dfrac{1}{5} t_0$ $\qquad\qquad\qquad$ (7-4.22)

These formulas are to be used with the next to last entry of Table 7-4.1.

5% Overshoot. For set point inputs, a response having an overshoot of 5% of the change in set point is highly desirable. For this type of response, Martin **et al.** (1976)

recommend that τ_c be set equal to the effective dead time of the FOPDT model. This results in the following formula for the controller gain that produces 5% overshoot on *set-point changes:*

$$K_c = \frac{0.5}{K}\left(\frac{\tau}{t_0}\right) \qquad (7\text{-}4.23)$$

Comparison of this formula with the one in Table 7-2.1 shows that this is about 40% of the PID gain required for quarter decay ratio (50% overshoot).

One interesting point about the controller synthesis method is that if controllers had been designed this way from the start, then the evolution of controller modes would have followed the pattern I, PI, PID. This pattern reflects evolution from the simplest *process model* to the more complex. Contrast this to the actual evolution of industrial controllers: P, PI, PID-that is, from the simplest *controller* to the more complex.

An important insight that we can gain from the controller synthesis procedure is that the main effect of the proportional mode, when added to the basic integral mode, is to compensate for the longest or dominant process lag, whereas that of the derivative mode is to compensate for the second-longest lag or for the effective process dead time. The entire synthesis procedure is based on the assumption that the primary closed-loop response specification is the elimination of offset or steady-state error. This is what makes the integral mode the basic controller mode.

EXAMPLE 7-4.1

Determine the tuning parameters for the heat exchanger of Example 6-2.1 using the formulas derived by the controller synthesis method. Compare these results with those obtained from the minimum IAE tuning formulas for set point inputs.

SOLUTION

The FOPDT parameters obtained from the heat exchanger by fit 3 in Example 7-2.1 are

$$K = 0.80 \ \%TO/\%CO \qquad \tau = 33.8 \ s \qquad t_0 = 11.2s$$

Because the dead time in this case is greater than one-fourth of the time constant, a PID controller is indicated. The gain for minimum IAE on disturbance input is obtained with $\tau_c = 0$. From the next to last entry in Table 7-4.1,

$$K_c = \frac{33.8}{(0.80)(11.2)} = 3.8 \ \%/\%$$

For minimum IAE set point input, from Eq. 7-4.22,

$$\tau_c = \frac{1}{5} (11.2) = 2.24 \text{ s}$$

$$K_c = \frac{33.8}{(0.80)(2.24 + 11.2)} = 3.1 \%/\%$$

For 5% overshoot on set point input, from Eq. 7-4.23,

$$K_c = \frac{(0.5)(33.8)}{(0.8)(11.2)} = 1.9 \%/\%$$

The integral and derivative times are, from Table 7-4.1,

$$\tau_I = \tau = 34 \text{ s} \ (0.56 \text{ min})$$

$$\tau_D = \frac{t_0}{2} = 5.6 \text{ s} \ (0.093 \text{ min})$$

For comparison, the minimum IAE parameters for set point inputs as calculated by the PID formulas in Table 7-2.3 are

$$K_c = \frac{1.086}{0.80} \left(\frac{11.2}{33.8} \right)^{-0.869} = 3.5 \frac{\%CO}{\%TO}$$

$$\tau_I = \frac{33.8}{0.740 - 0.130(11.2/33.8)} = 48 \text{ s} \ (0.81 \text{ min})$$

$$\tau_D = 0.348(33.8) \left(\frac{11.2}{33.8} \right)^{0.914} = 4.3 \text{ s} \ (0.071 \text{ min})$$

These two sets of parameters are in the same ball park.

EXAMPLE 7-4.2

A second-order-plus-dead-time process has the following transfer function:

$$G(s) = \frac{1.0e^{-0.26s}}{s^2 + 4s + 1}$$

Compare the response to a step change in set point of a PI controller tuned by Ziegler–Nichols quarter decay ratio, minimum IAE for set-point changes, and controller syn-

thesis with the gain adjustment for a 5% overshoot. Use Fig. 7-2.8 to obtain the first-order-plus-dead-time (FOPDT) model parameters.

SOLUTION

The first step is to approximate the second-order transfer function by a FOPDT model. We start by factoring the denominator into two effective time constants.

$$G(s) = \frac{1.0e^{-0.26s}}{(3.73s + 1)(0.27s + 1)}$$

The second step is to approximate the second-order lag with a FOPDT model, using Fig. 7-2.8.

$$\frac{\tau_2}{\tau_1} = 0.072 \qquad \tau' = 1.07, \ = 3.73 \text{ min}$$

$$t_0' = 0.0727, \ = 0.27 \text{ min}$$

Note that the second time constant is small enough for the simple rule of thumb given in Section 7-2 to apply-that is, $\tau' = \tau_1$, $t_0' = \tau_2$. The effective dead time of the FOPDT model must be added to the actual-process dead time to obtain the total dead time.

$$t_0 = 0.26 + t_0' = 0.53 \text{ min}$$

The FOPDT parameters are then

$$K = 1.0 \qquad \tau = 3.73 \text{ min} \qquad t_0 = 0.53 \text{ min}$$

The third step is to calculate the tuning parameters from the formulas specified. For a PI controller:

Quarter Decay Ratio (from Table 7-2.1)

$$K_c = \frac{0.90}{1.0} \left(\frac{0.53}{3.73}\right)^{-1} = 6.3 \ \frac{\%CO}{\%TO}$$

$$\tau_I = 3.33t_0 = 3.33(0.53) = 1.8 \text{ min}$$

Minimum IAE for Set Point Changes (from Table 7-2.3)

$$K_c = \frac{0.758}{1.0} \left(\frac{0.53}{3.73}\right)^{-0.861} = 4.1 \ \frac{\%CO}{\%TO}$$

$$\tau_I = \frac{3.73}{1.02 - 0.323(0.53/3.73)} = 3.8 \text{ min}$$

Controller Synthesis Tuned for 5% Overshoot (from Table 7-4.1)

$$\tau_I = \tau = 3.7 \text{ min}$$

For 5% overshoot (Eq. 7-4.23), we get

$$\tau_c = t_0 = 0.53 \text{ min}$$

$$K_c = \frac{3.73}{2(0.53)} = 3.6 \frac{\%CO}{\%TO}$$

The final step is to compare the responses to a unit step in set point using each of the three sets of tuning parameters. Martin *et al.* (1976) published the solution to this problem, which they obtained by computer simulation. The resulting responses are shown in Fig. 7-4.4. Comparison of the responses shows that the controller synthesis formulas for 5% overshoot yield a response that is very close to the minimum IAE set point response. These responses are superior to the quarter decay ratio response in terms of stability and settling time for changes in set point,
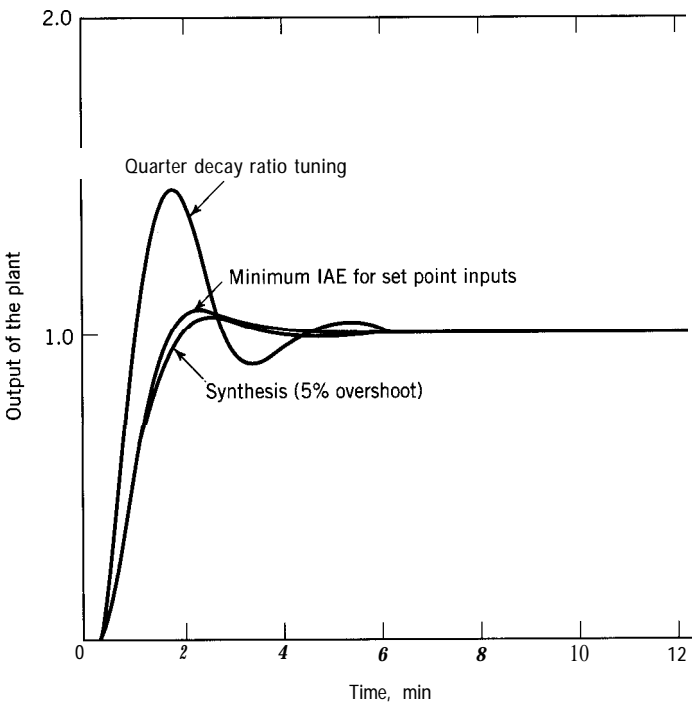
Figure 7-4.4 Closed-loop response to step change in set point for second-order process with PI controller. Model parameters: $\tau = 3.7$ min, $t_0 = 0.53$ min, and $K = 1$. (Reproduced by permission of Reference 8, copyright © ISA, 1976.)
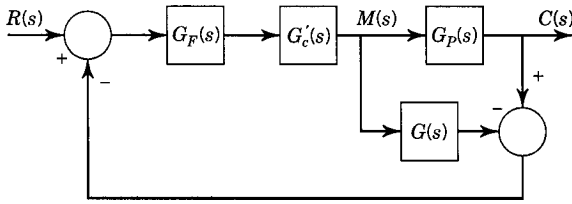
Figure 7-4.5 Block diagram for internal model control (IMC).

7-4.5 Tuning Rules by Internal Model Control (IMC)

A popular method for tuning feedback controllers is commonly known as the IMC tuning rules. Rivera *et* al. (1986) show how tuning rules for feedback controllers can be developed by internal model control (IMC). In this section we will look at the IMC method and see that the controller tuning rules are basically those we have developed in this section by the closed-loop synthesis method.

Consider the IMC feedback loop of Fig. 7-4.5. In this block diagram, G(s) is the transfer function of an *internal model* of the true process transfer function $G_P(s)$. The feedback controller consists of three blocks: the internal model G(s), an adjustable filter $G_F(s)$, and a dynamic compensator $G'_c(s)$. The filter is usually a first-order filter with an adjustable time constant, and the internal model and the compensator are designed for each type of process.

From the block diagram of Fig. 7-4.5, the controller output is

$$M(s) = G'_c(s)G_F(s)[R(s) - C(s) + G(s)M(s)]$$

Solving for M(s) and rearranging, we obtain the controller transfer function

$$G_c(s) = \frac{M(s)}{R(s) - C(s)} = \frac{G'_c(s)G_F(s)}{1 - G'_c(s)G(s)G_F(s)} \qquad (7\text{-}4.24)$$

This is essentially the IMC controller synthesis formula. The filter in this formula is usually a first-order filter with unity gain:

$$G_F(s) = \frac{1}{\tau_c s + 1} \qquad (7\text{-}4.25)$$

where τ_c is the adjustable filter time constant. The internal model, G(s), is selected to match the true process transfer function $G_P(s)$, and the dynamic compensator is selected as the reciprocal of the process model, excluding any terms that are unrealizable, such as dead time, or that would create instability, such as a positive zero. An important requirement is that the gain of $G'_c(s)$ be exactly the reciprocal of the gain of the process model G(s). We can see from Eq. 7-4.24 that if this is the case, because the filter gain is also unity, the controller steady-state gain will be infinite, which means it will not produce an offset.

To show that the IMC synthesis formula is equivalent to the closed-loop response synthesis formula, let us assume for the moment that it is possible to implement $G_c'(s) = 1/G(s)$. Substitute, along with Eq. 7-4.25, into Eq. 7-4.24 and simplify to obtain

$$G_c(s) = \cfrac{\cfrac{1}{G(s)} \cdot \cfrac{1}{\tau_c s + 1}}{1 - \cfrac{1}{G(s)} \cdot G(s) \cdot \cfrac{1}{\tau_c s + 1}} = \frac{1}{G(s)} \cdot \frac{1}{\tau_c s} \tag{7-4.26}$$

This formula is identical to the synthesis formula of Eq. 7-4.5. It follows that the feedback controller modes and tuning formulas of Table 7-4.1, which were derived using Eq. 7-4.5, would have also resulted from Eq. 7-4.26, except for the next to last set of formulas, which are for a process with dead time. For this case, the IMC dynamic compensator excludes the dead time.

$$G(s) = \frac{Ke^{-t_0 s}}{\tau s + 1} \qquad\qquad G_c'(s) = \frac{\tau s + 1}{K}$$

Substitute these expressions and Eq. 7-4.25 into Eq. 7-4.24 and simplify to obtain

$$G_c(s) = \frac{\tau s + 1}{K} \cdot \frac{1}{\tau_c s + 1 - e^{-t_0 s}} \tag{7-4.27}$$

This equation is identical to Eq. 7-4.14, which was used to derive the next to last set of tuning formulas in Table 7-4.1.

We have shown that the controller modes and tuning formulas of Table 7-4.1 are also the popular IMC tuning rules. In Chapter 15, we will discuss the IMC technique in more detail and use it to develop a dynamically compensated feedback controller.

7.5 TIPS FOR FEEDBACK CONTROLLER TUNING

Each of the tuning methods presented in the preceding sections requires some form of process dynamic testing to be carried out, either to find the ultimate gain and period of the loop or to find the parameters of a simple process model. Unfortunately, for many processes, formal tests cannot be carried out because of safety, product quality, or other considerations. The tips presented in this section are intended to facilitate the tuning when no formal testing can be performed. We will also look at an important consideration that is often overlooked when tuning feedback controllers: the compromise between tight control of the controlled variable and excessive movement of the manipulated variable.

The first tip is to realize that the performance of a feedback controller is relatively insensitive to the values of the tuning parameters, a point that was made earlier in this chapter. In modem control jargon, we say that feedback control is a robust technique. Approximate values of the tuning parameters can produce a "good enough," if not an "optimal," response. In fact, as discussed earlier, in tuning linear controllers for non-linear processes, there is no such thing as an optimal set of tuning parameters. Fur-

thermore, it is possible to be off by 40% or 50% on the integral and derivative times and still obtain a good response by adjusting the controller gain.

On the basis of this tip, our procedure for tuning a controller becomes the following:

1. Obtain approximate values for the integral and derivative times.

2. Adjust the proportional gain to obtain an acceptable response.

7-5.1   Estimating the Integral and Derivative Times

We can use the tuning formulas presented in this chapter as a guide in estimating the integral and derivative times. It is usually possible to estimate the ultimate period or the time constant of the process in the loop without formally testing the process. On the other hand, the process dead time is harder to estimate. From the estimated process parameters, the integral and derivative times can be estimated as follows:

- The period of oscillation can be used with the formulas of Table 7- 1.1 to estimate the integral and derivative times.
- If the process dominant time constant can be estimated, set the integral time equal to it (Table 7-4.1), and, on the basis of the quarter decay ratio formulas (Table 7-1.1 or 7-2.1), set the derivative time equal to one-fourth of the integral time. However, if we suspect that the dead time is much smaller than the time constant, we should use a PI controller (zero derivative time).
- If the process dead time can be estimated, the derivative time can be set at one-half of the dead time (Table 7-2.1 or 7-4.1). This method automatically results in a PI controller if the dead time is negligible.

Today's control systems provide convenient time trends of any measured process variable. The time scale on these trends can be adjusted from a few minutes to several hours. The process variable scale can also be blown up as needed for greater precision. These trends are invaluable for estimating the necessary process parameters.

The period of oscillation of the loop can be estimated by observing its closed-loop response when the controller is proportional-only, which is achieved by de-tuning or turning off the integral and derivative modes. Even if the oscillations are not sustained, the period between two peaks is usually about 40% to 60% higher than the ultimate period, so the ultimate period can be estimated as two-thirds of the period of the de-caying oscillations. Figure 7-5.1 shows a response plot of the outlet temperature from a furnace coil being controlled by manipulating the flow through the coil, which is also shown in the plot. This plot is from an actual process furnace. We see if we assume the controller is proportional, that the period of oscillation is about 15 min. Because the oscillations are sustained, we can assume this is the ultimate period and estimate the integral time as 7.5 min and the derivative time as 1.9 min, using the formulas from Table 7-1.1.

The time constant of the process can sometimes be estimated from simple models based on fundamental principles, such as those presented in Chapters 3 and 4. For example, from the models of mixed tanks presented in Sections 3-2 and 3-6, we can estimate that the time constant of a tank is approximately equal to its residence time (volume/product flow). Similar formulas are derived in Chapter 3 for the time constants of gas tanks, reactors, and so on.

The process time constant and dead time can also be estimated from the response of
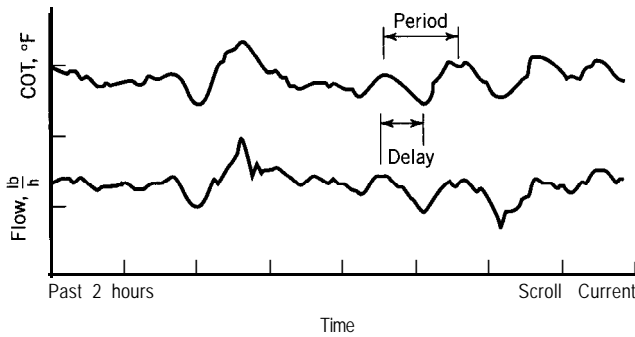
Figure 7-5.1 Trend of furnace coil outlet temperature (COT) and
the manipulated flow through the coil.

the closed loop, but this requires careful analysis. Using the concepts of Section 2-5, it
can be shown that the response of several lags in series to a ramp or slow sine wave is
eventually a ramp or sine wave that lags the input ramp by the sum of the time constants
and the dead time. Let us look, for example, at the response of the coil outlet temperature
and flow in Fig. 7-5.1. At first the response may look wrong, because the flow and the
temperature go up and down together. From basic principles we know that when the
flow through the coil goes up, the outlet temperature must come down, because there
is more fluid to heat up with the same amount of energy. The reason they go up and
down together is that the immediate action of the proportional controller causes the
flow to increase and decrease with the temperature (direct-acting controller). If we watch
carefully, we notice that the temperature does go down when the process fluid goes up,
and vice versa, but not right away! Because of the process dynamics, there is a delay
between the change in the flow and the corresponding change in temperature. This
delay is marked in the figure and seems to be about 8 min. How, then, are we going to
model the coil?

- If we model the coil as a perfectly mixed system, the delay will be equal to its
  time constant, as in Section 3-2.
- If we model the coil as a plug flow system, the delay will all be caused by pure
  dead time, as in Section 3-3.

The first of these models results, from Table 7-4.1, in a PI controller with an integral
time of 8 min. This value agrees with our earlier estimate based on the period of os-
cillation. The second model results, from Table 7-2.1, in a PID controller with an
integral time of $(2)(8) = 16$ min and a derivative time of $8/2 = 4$ min. These estimates
are about twice those obtained from the period of oscillation.

Like any real system, the coil is neither perfectly mixed nor plug flow but rather a
combination of the two, probably closer to the latter. This means that the 8-min delay
is just the sum of the time constant and the dead time. From the two extremes, we get
estimates of the integral time ranging from 8 to 16 min and estimates of the derivative
time ranging from 0 to 4 min. We could use an integral time of 12 min and a derivative
time of 3 min and be confident that these values are good enough to enable us to proceed
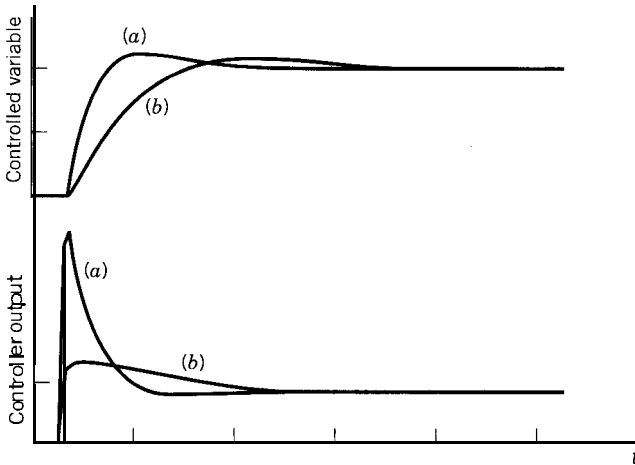with the adjustment of the proportional gain to obtain the desired response.

Figure 7-5.2 Comparison of tight (a) and reasonable (b) adjust-
ment of the controller gain for set point change.

7-5.2 Adjusting the Proportional Gain

One of the problems with many tuning methods has been their rigidity; they provide a
set of tuning formulas that leave no room for adjustment of any of the parameters. In
contrast, the controller synthesis method and IMC tuning rules of Section 7-4 provide
formulas with an adjustable gain.

When adjusting the proportional gain, we must keep in mind that very tight control
of the process variable usually requires large changes in the manipulated variable. This
is undesirable, because changes in the manipulated variable cause upsets to the process
and disturb other control loops. For example, a large decrease in the fuel flow to a
furnace could cause the flame to go out, and a large increase in the reflux flow to a
distillation column may cause the column to flood. Figure 7-5.2 shows the controlled
and manipulated variable responses for two values of the controller gain. The higher
gain results in tighter control, but it also causes a much larger upset of the process
because of the larger initial change in the controller output. When adjusting the gain,
be sure to consider both the tightness of control and the variability of the manipulated
variable.

After the integral and derivative times have been selected, tuning is reduced to the
adjustment of a single parameter, the gain. Thus no special procedure is required for
adjusting the gain.

7-6 SUMMARY

This chapter presented two methods for characterizing the process dynamic response
and for tuning feedback controllers. We also looked at the tuning of feedback controllers
for integrating processes. The controller synthesis method gave us insight into the func-
tions of the proportional, integral, and derivative modes and provided a set of simple
tuning formulas. The chapter closed with some practical tips for tuning feedback con-
trollers.

The next two chapters examine two classical methods for analyzing control loop

responses, root locus and frequency response, and a more powerful process identification method, pulse testing.

REFERENCES

1. Chiu, K. C., A. B. Corripio, and C. L. Smith. 1973. "Digital Control Algorithms. Part III: Tuning PI and PID Controllers." *Instruments* and *Control Systems* 46 (December):41–43.
2. Corripio, A. B. 1990. *Tuning of Industrial Control Systems.* Research Triangle Park, N.C.: ISA, Unit 6.
3. Dahlin, E. B. 1968. "Designing and Tuning Digital Controllers." *Instruments and Control Systems* 41 (June):77.
4. Lopez, A. M., P. W. Murrill, and C. L. Smith. 1967. "Controller Tuning Relationships Based on Integral Performance Criteria." *Instrumentation Technology* 14 (November):57.
5. Martin, Jacob, Jr. 1975. Ph.D. dissertation, Design and Comparison of Controllers Based on Simplified Models of the Process, Department of Chemical Engineering, Louisiana State University, Baton Rouge.
6. Martin, J. Jr., A. B. Corripio, and C. L. Smith. 1976. "How to Select Controller Modes and Tuning Parameters from Simple Process Models." *ISA Transactions* 15(4):314–319.
7. Moore, C. F., C. L. Smith, and P. W. Murrill. 1969. "Simplifying Digital Control Dynamics for Controller Tuning and Hardware Lag Effects. 1969." *Instrument Practice 23* (January):45.
8. Murrill, Paul W. 1991. *Fundamentals of Process Control Theory.* 2nd ed. Research Triangle Park, N.C.: Instrument Society of America.
9. Rivera, D. E., M. Morari, and S. Skogestad. 1986. "Internal Model Control, 4. PID Controller Design." *I&EC Process Design and Development* 25:252.
10. Rovira, A. A. 198 1. Ph.D. dissertation, Control Algorithm Tuning, and Adaptive Gain Tuning in Process Control, Department of Chemical Engineering, Louisiana State University, Baton Rouge.
11. Smith, Cecil L. *Digital Computer Process Control.* 1972. Scranton, Pa.: Intext Educational Publishers.
12. Ziegler, J. G., and N. B. Nichols. 1942. "Optimum Setting for Automatic Controllers." *Transactions ASME* 64 (November):759.

PROBLEMS

7-1. A feedback control loop is represented by the block diagram of Fig. 7-2.2. The process transfer function is given by

$$G_1(s) = \frac{K}{(\tau_1 s + 1)(\tau_2 s + 1)(\tau_3 s + 1)}$$

where the process gain is $K = 2.5$ %TO/%CO and the time constants are

$$\tau_1 = 5 \text{ min} \qquad \tau_2 = 0.8 \text{ min} \qquad \tau_3 = 0.2 \text{ min}$$

Determine the controller tuning parameters for quarter decay ratio response by the ultimate gain method for

(a) A proportional (P) controller.

(b) A proportional-integral (PI) controller.

(c) A proportional-integral-derivative (PID) controller.

7-2. Using the tuning parameters determined for the loop of Problem 7-1, find the roots of the characteristic equation, identify the dominant pair of roots, and calculate the damping ratio and the decay ratio of the response.

7-3. Given the feedback control loop of Fig. 7-2.2 and the following process transfer function

$$G_1(s) = \frac{Ke^{-t_0s}}{(\tau_1 s + 1)(\tau_2 s + 1)}$$

where the process gain, time constants, and dead time are $K = 1.25$ %TO/%CO, $\tau_1 = 1$ min, $\tau_2 = 0.6$ min, and $t_0 = 0.20$ min, use Fig. 7-2.8 to estimate the first-order-plus-dead-time (FOPDT) model parameters. Then use these parameters to compare the tuning parameters for a proportional-integral (PI) controller using the following formulas:

(a) Quarter decay ratio response

(b) Minimum IAE for disturbance inputs

(c) Minimum IAE for set point inputs

(d) Controller synthesis for 5% overshoot on a set point change

7-4. Do Problem 7-3 for a proportional-integral-derivative (PID) controller.

7-5. Do Problem 7-3 for a sampled-data (computer) controller with a sample time $T = 0.10$ min.

7-6. For the control loop of Problem 7-3, derive the tuning formulas for an actual PID controller using the controller synthesis procedure. Consider two cases:

(a) No dead time, $t_0 = 0$.

(b) The dead time given in the problem.

Check your answers with the entries in Table 7-4.1.

7-7. Prepare a computer program to obtain the responses to step changes in set point and in disturbance of the control loop in Problem 7-3. Use the controller tuning parameters determined there. Can you improve on the control performance by trial-and-error adjustment of the tuning parameters? (Have the program print the integral of the absolute error, IAE, and use it for the measure of control performance.)

7-8. From the results of Problem 6- 11, calculate the quarter decay ratio tuning parameters of a PI controller for the composition from the blending tank.

7-9. From the results of Problem 6-12, calculate the quarter decay ratio tuning parameters of a PID controller for the reactor temperature controller.

7-10. From the results of Problem 6-14, calculate the quarter decay ratio tuning parameters of a PI controller for the composition from the third tank.

7-11. For the composition controller for the three isothermal reactors in series of Problem 6-17, calculate the quarter decay ratio tuning parameters for a PI controller. Using these parameters, find the roots of the characteristic equation, identify the dominant root, and estimate the damping ratio and the actual decay ratio of the response.
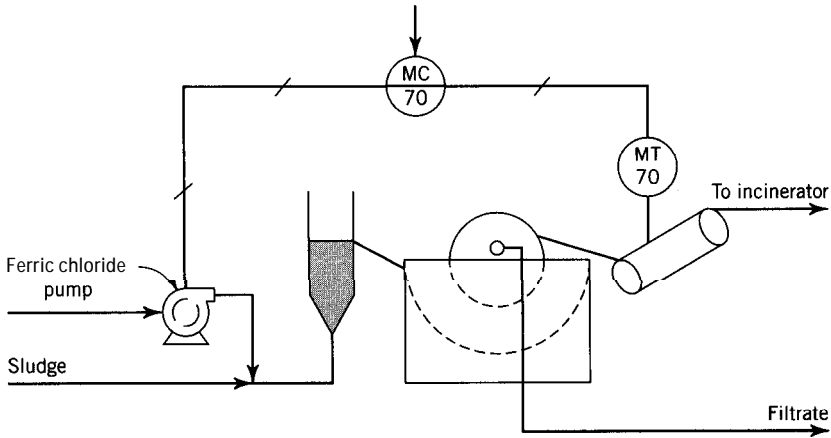
Figure P7-1 Vacuum filter for Problem 7-1.5.

7-12. Do Problem 7-11 for the compressor suction pressure controller of Problem 6-18.

7-13. Do Problem 7-11 for the temperature controller of the stirred tank cooler of Problem 6-19.

7-14. Do Problem 7-11 for the composition control of the reactors in series of Problem 6-22.

7-15. Consider the vacuum filter shown in Fig. P7-1. This process is part of a waste treatment plant. The sludge enters the filter at about 5% solids. In the vacuum filter, the sludge is de-watered to about 25% solids. The filterability of the sludge in the rotating filter depends on the pH of the sludge entering the filter. One way to control the moisture of the sludge to the incinerator is by adding chemicals (ferric chloride) to the sludge feed to maintain the necessary pH. Figure P7-1 shows a proposed control scheme. The moisture transmitter has a range of 55% to 95%.

The following data have been obtained from a step test on the output of the controller (MC70) of + 12.5 %CO.

Time, min	Moisture, %	Time, min	Moisture, %
0	75.0	10.5	70.9
1	75.0	11.5	70.3
1.5	75.0	13.5	69.3
2.5	75.0	15.5	68.6
3.5	74.9	17.5	68.0
4.5	74.6	19.5	67.6
5.5	74.3	21.5	67.4
6.5	73.6	25.5	67.1
7.5	73.0	29.5	67.0
8.5	72.3	33.5	67.0
9.5	71.6		

When the input moisture to the filter was changed by 2.5%, the following data were obtained.

Time, min	Moisture, %	Time, min	Moisture, %
0	75	11	75.9
1	75	12	76.1
2	75	13	76.2
3	75	14	76.3
4	75.0	15	76.4
5	75.0	17	76.6
6	75.1	19	76.7
7	75.3	21	76.8
8	75.4	25	76.9
9	75.6	29	77.0
10	75.7	33	77.0

(a) Draw a block diagram for the moisture control loop. Include the possible disturbances.

(b) Use fit 3 to estimate the parameters of first-order-plus-dead-time models of the two transfer functions. Redraw the block diagram showing the transfer function for each block.

(c) Give an idea of the controllability of the output moisture. What is the correct controller action?

(d) Obtain the gain of a proportional controller for minimum IAE response. Calculate the offset for a 5% change in inlet moisture.

(e) Tune a PI controller for quarter decay ratio response.

7-16. Consider the absorber shown in Fig. P7-2. There is a gas flow entering the absorber with a composition of 90 mole % air and 10 mole % ammonia (NH_3).

Air, NH_3

H_2O

Air, NH_3

H_2O
NH_3
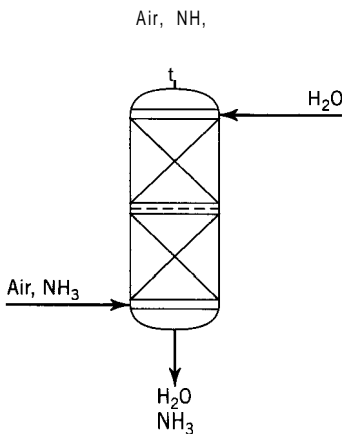
Figure P7-2 Absorber for Problem 7-16.

Before this gas is vented to the atmosphere, it is necessary to remove most of the NH, from it. This will be done by absorbing it with water. The NH, concentration in the exit gas stream cannot be above 200 ppm. The absorber has been designed so that the outlet NH, concentration in the vapor is 50 ppm. From dynamic simulations of the absorber, the following data were obtained.

Response to a Step Change in Water Flow
to the Absorber

Time, s	Water Flow, gpm	Outlet NH, Concentration
0	250	50.00
0	200	50.00
20	200	50.00
30	200	50.12
40	200	50.30
50	200	50.60
60	200	50.77
70	200	50.90
80	200	51.05
90	200	51.20
100	200	51.26
110	200	51.35
120	200	51.48
130	200	51.55
140	200	51.63
160	200	51.76
180	200	51.77
250	200	51.77

(a) Design a control loop for maintaining the outlet NH, concentration at a set point of 50 ppm. Draw the instrument diagram for the loop. There are some instruments in the stock room that you can use for this purpose. There is an analyzer transmitter calibrated for 0 to 200 ppm. This instrument has a negligible time lag. There is also a control valve that, at full opening and for the lo-psi pressure drop that is available, will pass 500 gpm. The time constant of the valve actuator is negligible. You may need more instrumentation to complete the design, so go ahead and use anything you need. Specify the fail position of the control valve and the action of the controller.

(b) Draw a block diagram for the closed loop and obtain the transfer function for each block. Approximate the response of the absorber with a first-order-plus-dead-time model using fit 3.

(c) Tune a proportional-only controller for quarter decay ratio response and calculate the offset when the set point is changed to 60 ppm.

(d) Repeat part (c) using a PID controller.

7-17. Consider the furnace shown in Fig. P7-3, which is used to heat the supply air to a catalyst regenerator. The temperature transmitter is calibrated for 300°F to
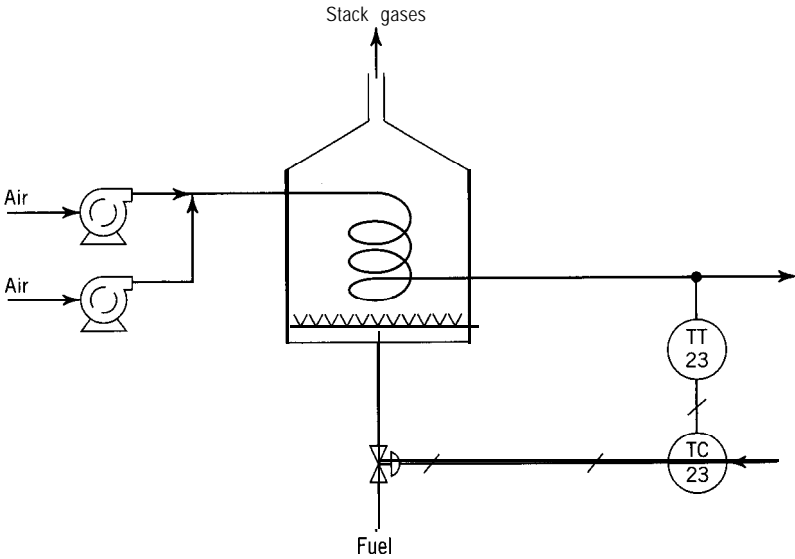
Figure **P7-3** Furnace for Problem 7-17.

500°F. The following response data were obtained for a step change of + 5% in the output of the controller.

Time, min	$T(t)$, °F	Time, min	$T(t)$, °F
0	425	5.5	436.6
0.5	425	6.0	437.6
1.0	425	7.0	439.4
2.0	425	8.0	440.7
2.5	426.4	9.0	441.7
3.0	428.5	10.0	442.5
3.5	430.6	11.0	443.0
4.0	432.4	12.0	443.5
4.5	434.0	14.0	444.1
5.0	435.3	20.0	445.0

(a) Draw the complete block diagram, specifying the units of each signal to or from each block. Identify each block and specify the fail-safe position of the valve and the correct action of the controller.
(b) Fit the process data by a first-order-plus-dead-time model using fit 3. Redraw the block diagram, showing the transfer function for each block.
(c) Tune a PID controller for quarter decay ratio response.
(d) Tune a PID controller by the controller synthesis method for 5% overshoot.

7-18. Calculate the tuning parameters for a PID temperature controller tuned for quarter decay ratio response for the oil heater of Problem 6-24.

7-19. Consider the chemical reactor system shown in Fig. P7-4. An exothermic catalytic reaction takes place inside the reactor tubes. The reactor is cooled by an oil flowing through the shell of the reactor. As the oil flows out of the reactor, it
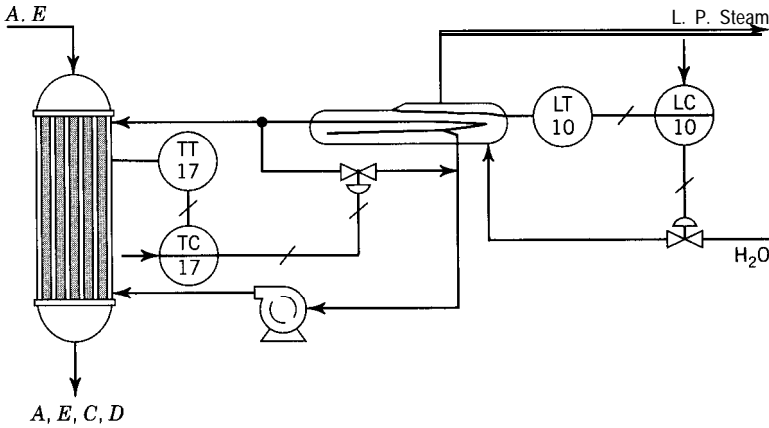
Figure P7-4 Catalytic reactor for Problem 7-19.

goes to a boiler where it is cooled by producing low-pressure steam. The temperature in the reactor is controlled by manipulating the bypass flow around the boiler. The following process conditions are known:

Reactor design temperature at point of measurement: 275°F
Oil flow the pump can deliver: 400 gpm (constant)
Control valve pressure drop at design flow: 10 psi
Valve flow at design conditions: 200 gpm
Range of the temperature transmitter: 150°F to 350°F
Density of the oil: 55 lb/ft^3
Open-loop test: A 5% change in the valve position results in a temperature change of 4.4°F after a very long time.
Closed-loop test: At a controller gain of 16 %CO/%TO, the temperature oscillates with constant amplitude and a period of 24 min.

(a) Size the temperature control valve for 100% overcapacity. Specify the fail-safe position of the control valve and the required controller action.
(b) If the pressure drop across the boiler tubes varies with the square of the flow and the valve is equal percentage with a rangeability parameter of 50, what is the valve position at design conditions? What is the flow through the valve when it is fully open?
(c) Draw a block diagram for the temperature loop. What are your recommended valve fail-safe position and controller action?
(d) Calculate the process gain, at design conditions, including the control valve and the temperature transmitter.
(e) Calculate the tuning parameters for a PID controller for quarter decay ratio response. Report them as proportional band, repeats per minute, and minutes.
(f) Tune a proportional controller for quarter decay ratio response and calculate the offset for a step change in set point of ➡ 10°F.

7-20. Consider the typical control system for the double-effect evaporator shown in Fig. P7-5. Evaporators are characterized by slow dynamics. The composition of the product out of the last effect is controlled by manipulating the steam to the first effect. The design feed rate and composition are 50,000 lb/h and 5.0 weight
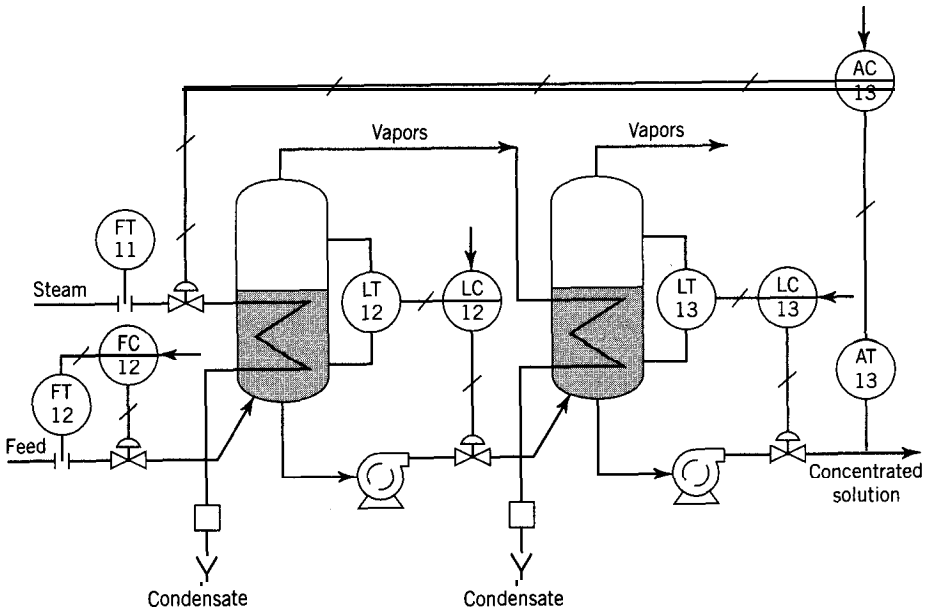
Figure P7-5 Double-effect evaporator for Problem 7-20.

%, respectively. Figure P7-6 shows the open-loop step response of the product composition for a change of 0.75% by weight in the composition of the solution entering the first effect. Figure P7-7 shows the response of the product composition to a change of 2.5% in controller output. The composition sensor/transmitter has a range of 10 to 35 weight %.

(a) Draw a complete block diagram with the transfer function of each block. What should be the fail-safe position of the control valve? What is the correct controller action?

(b) Tune a proportional-integral controller for quarter decay ratio response.
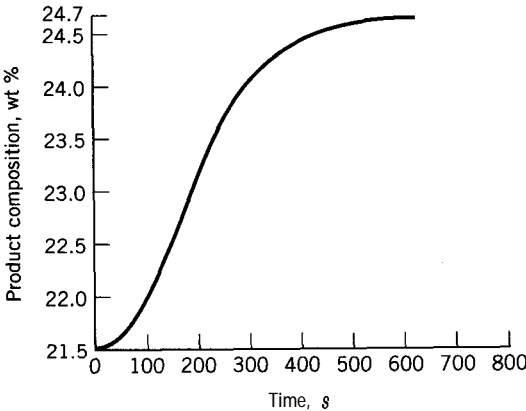
Figure P7-6 Response to step change in inlet composition for Problem 7-20.
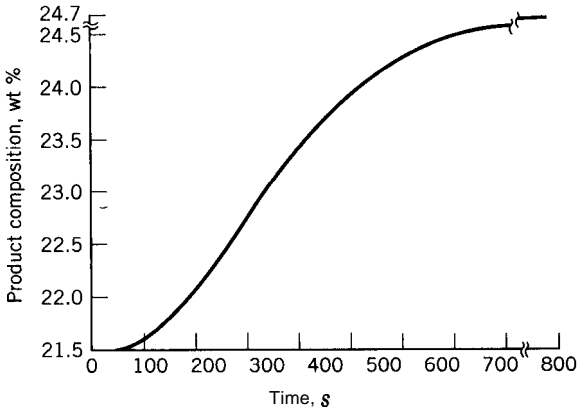
Figure P7-7 Response to step change in controller output
for Problem 7-20.

(c) Tune a PI controller for 5% overshoot, using the controller synthesis method.
7-21. The temperature in a continuous stirred tank exothermic chemical reactor is con-
trolled by manipulating the cooling water rate through a coil, as shown in Fig.
P7-8. The process design conditions are

Reactor temperature: 210°F
Cooling water flow: 350 gal/min
Pressure drop across the coil at design flow: 10 psi
Range of temperature transmitter: 190°F to 230°F
Control valve trim: Equal percentage with $\alpha = 50$.
Open-loop test: A 10 gal/min increase in water rate results in a temperature
change of 5.2°F after a long time.
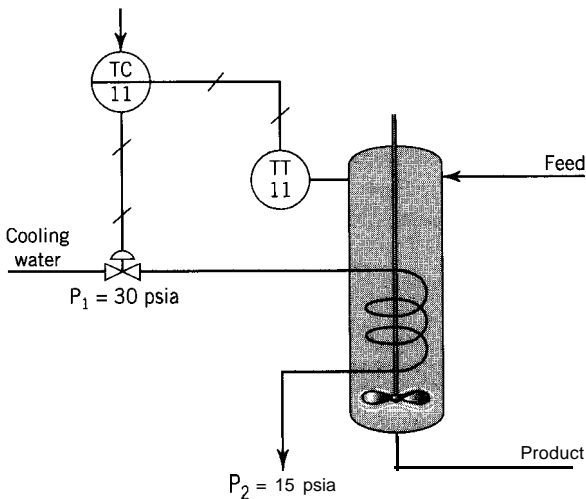
Figure P7-8 Reactor for Problem 7-21.

Closed-loop test: At a controller gain of 8.0 %CO/%TO, the temperature oscil-
lates with constant amplitude and a period of 14 min.

(a) Size the valve for 100% overcapacity, calculate the gain of the valve at design
flow, and specify the fail-safe position of the valve.
(b) Draw the block diagram for the control loop and determine the total process
gain, including the transmitter and the control valve.
(c) Calculate the PID controller tuning parameters for quarter decay ratio re-
sponse. Report them as proportional band, repeats per minute, and derivative
minutes. What is the required controller action?

7-22. Consider the process shown in Fig. P7-9 for drying phosphate pebbles. A table
feeder transports the pebble-water slurry into the bed of the dryer. In this bed
the pebbles are dried by direct contact with hot combustion gases. From the dryer,
the pebbles are conveyed to a silo for storage. It is *most* important to control the
moisture of the pebbles leaving the dryer. If the pebbles are too dry, they may
fracture into fine dust, resulting in possible loss of material. If they are too wet,
they may form large chunks, or clinkers, in the silo.

It is proposed to control the moisture of the exiting pebbles by the speed of
the table feeder, as shown in Fig. P7-9. The speed of the feeder is directly pro-
portional to its input signal. The moisture of the inlet pebbles is usually about
15% and is reduced to 3% in the dryer. The transmitter has a range of 1% to 5%
moisture. An important disturbance to this process is the moisture of the inlet
pebbles.

(a) Draw a complete block diagram of the control loop, showing all units. In-
clude the disturbances.
(b) Figure P7-10 shows the response of the outlet moisture to an increase of 8%
in controller output, and Fig. P7-11 shows the response of the outlet moisture
to an increase of 3% in inlet moisture. Approximate each process curve by
a first-order-plus-dead-time model. Use fit 2. Redraw the block diagram,
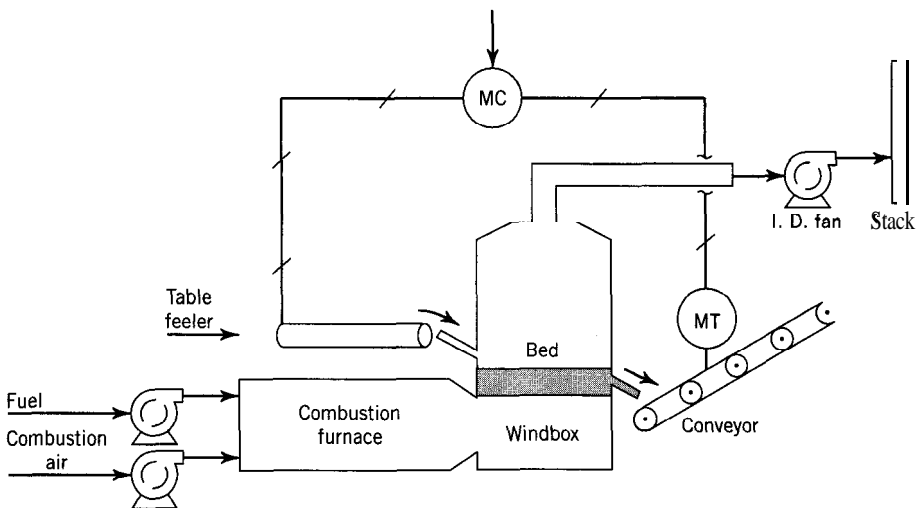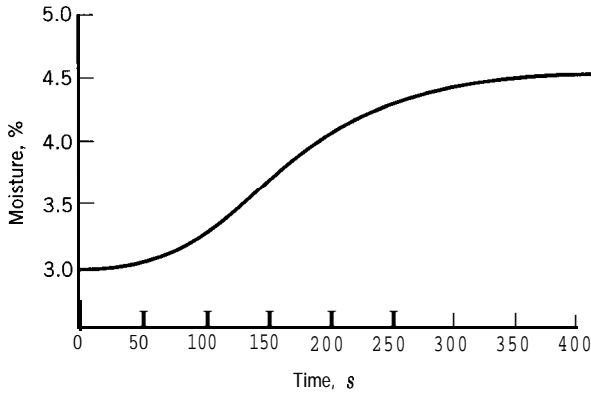showing the transfer functions of these approximate models.

Figure P7-9 Dryer of phosphate pebbles for Problem 7-22.

Figure **P7-10** Response to step change in controller output for Problem 7-22.

(c) Determine the tuning of a PID controller for minimum ISE response on disturbance inputs. Report the controller gain as proportional band. What is the correct controller action?

(d) If the inlet moisture of the pebbles drops by 2%, what is the new steady-state value of the outlet moisture? Assume that the controller is proportional-only tuned for quarter decay ratio response from the information determined in part (b).

(e) What is the controller output required to avoid offset for the disturbance of part (d)?

7-23. In the liquid level control system of Fig. 7-3.1, assume that the level control valve is installed on the inlet line to manipulate the inlet flow, $f_i(t)$, and that the outlet flow is the disturbance. Modify the block diagrams of the loop (Fig. 7-3.2), and specify the correct action of the level controller assuming the control valve fails closed. Rewrite the closed-loop transfer function for the level. Do any of the formulas developed in Section 7-3.2 for the proportional level controller change?
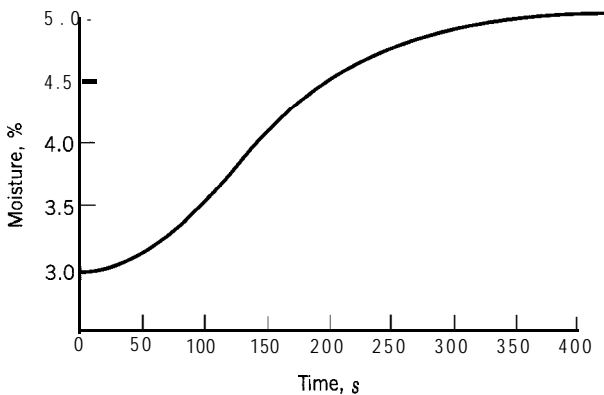
Figure **P7-11** Response to step change in inlet moisture for Problem 7-22.

7-24. Consider the first of the two tanks in series of Fig. 4-1.1. A model for this tank is derived in Section 4- 1.1, and it assumes the outlet flow depends on the level in the tank. This model results in a first-order lag transfer function for the level in the tank. A level control system is to be installed on that tank. The controller will manipulate the valve on the outlet line. Let the diameter of the tank be 3.0 m, the design level be 2.0 m, the range of the level transmitter be 1 to 3 m above the bottom of the tank, and the outlet valve be linear and sized for a maximum flow of twice the design flow of 0.003 m^3/s. The pressure drop in the line is negligible, and the valve can be represented by a first-order lag with a time constant of 5 s.

(a) Draw the block diagram for the level control loop, using the model developed for the first tank in Section 4-1.1. Specify the action of the controller and calculate the parameters of all the transfer functions. Calculate the maximum gain of a proportional level controller that will result in a non-oscillatory response. Calculate also the effective time constants of the loop at that gain and the offset caused by a 0.001 m^3/s change in inlet flow.

(b) Now model the tank as though the outlet flow is a function only of the valve position, as in Section 7-3.1 (integrating process), and carry out the calculations of part (a) using this model. Compare the answers from the two models.

7-25. Although proportional control is ideal for both tight and averaging level control, many operators prefer to have a proportional-integral (PI) controller because they are not used to controllers with offset.

(a) Obtain the closed-loop transfer function and the characteristic equation of the loop for the block diagram of Fig. 7-3.2 using a PI controller. Show that there is no offset for either a change in set point or a change in flow.

(b) Determine the roots of the characteristic equation when the integral time is equal to the valve time constant. What is the level response? Is it dependent on the controller gain?

(c) Assuming a negligible valve time constant, show that the response of the PI level controller is oscillatory at low controller gains and non-oscillatory at high gains! Show that the minimum gain for which the response is not oscillatory is $4/\tau_I$. Show also that at very high gains, the dominant time constant of the closed loop is equal to the integral time of the controller.

7-26. A level controller on a calandria-type evaporator needs to be tuned very tightly, because its operation is very sensitive to the level. A high level results in a rise in the boiling temperature due to the hydrostatic pressure of the fluid, whereas a low level results in the formation of a scale of dried-up solids at the top of the hot tubes. Because of this, most evaporator levels are controlled by manipulating the feed flow to the evaporator, which is the largest flow. Draw the schematic for the level control loop to an evaporator. The steam and product flows are the disturbances. Draw also the block diagram of the loop. Calculate the maximum proportional controller gain for non-oscillatory response, the effective time constants of the level loop at that gain, and the offset for a 10% change in flow for the following design parameters: feed flow = 800 lb/min, density of concentrated solution = 98 lb/ft^3, evaporator cross-sectional area = 10 ft^2, valve time constant = 2.0 s, and level transmitter span = 4 ft. The valve is linear and is sized for 100% overcapacity.

7-27. The transfer function for a process with inverse response (see Section 4-4.3) is

$$G(s) = \frac{K(1 - \tau_3 s)}{(\tau_1 s + 1)(\tau_2 s + 1)}$$

Show that if we tried to synthesize a feedback controller for the standard first-order response specification (Eq. 7-4.4), the controller denominator would contain a positive (unstable) root. Then synthesize a feedback controller for the following closed-loop response specification.

$$\frac{C(s)}{R(s)} = \frac{1 - \tau_3 s}{\tau_c s + 1}$$

Identify the type of controller and obtain the tuning formulas for it.

Chapter 8

Root Locus

In Chapter 6, we began the study of the stability of control systems by presenting two techniques: direct substitution and Routh's test. This chapter and the following one continue with this study. Specifically, this chapter presents the root locus technique, and Chapter 9 presents the frequency response techniques. This presentation is made from a practical point of view, stressing what the techniques indicate about stability of the processes and how the different loop terms affect the stability.

8-1 SOME DEFINITIONS

Before the root locus technique and the frequency response techniques are presented, some new terms must be defined. Consider the general closed-loop block diagram shown in Fig. 8-1.1. As we saw in Chapter 6, the closed-loop transfer functions are

$$\frac{C(s)}{R(s)} = \frac{G_c(s)G_v(s)G_{P_1}(s)}{1 + H(s)G_c(s)G_v(s)G_{P_1}(s)} \tag{8-1.1}$$

and

$$\frac{C(s)}{L(s)} = \frac{G_{P_2}(s)}{1 + H(s)G_c(s)G_v(s)G_{P_1}(s)} \tag{8-1.2}$$

with the characteristic equation

$$1 + H(s)G_c(s)G_v(s)G_{P_1}(s) = 0 \tag{8-1.3}$$

The open-loop transfer function (OLTF) is defined as the product of all the transfer functions in the control loop; that is,

$$\text{OLTF} = H(s)G_c(s)G_v(s)G_{P_1}(s) \tag{8-1.4}$$

Therefore, the characteristic equation can also be written as

$$1 + \text{OLTF} = 0 \tag{8-1.5}$$

368
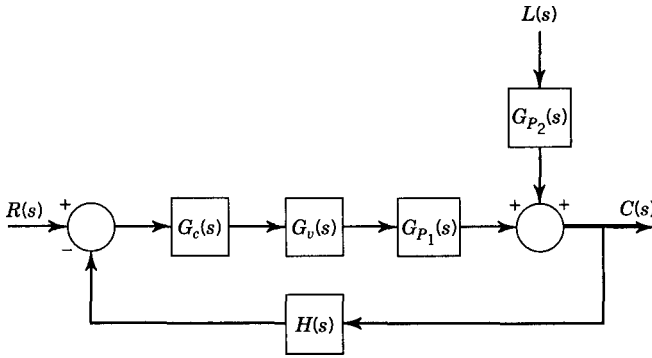
Figure **S-l.1** General closed-loop block diagram.

Now suppose that the individual transfer functions are known and that the OLTF looks like this:

$$\text{OLTF} = \frac{K_c K_v K_{P_1} K_T (1 + \tau_D s)}{(\tau_T s + 1)(\tau_1 s + 1)(\tau_2 s + 1)}$$

or

$$\text{OLTF} = \frac{K(1 + \tau_D s)}{(\tau_T s + 1)(\tau_1 s + 1)(\tau_2 s + 1)}$$

where $K = K_c K_v K_{P_1} K_T$.

The **poles** are defined as the roots of the denominator of the OLTF. For the foregoing OLTF the poles are $-1/\tau_T$, $-1/\tau_1$, and $-1/\tau_2$. The **zeros** are defined as the roots of the numerator of the OLTF, or $-1/\tau_D$ for the foregoing OLTF.

These definitions are generalized by writing the OLTF as

$$\text{OLTF} = \frac{K \prod_{i=1}^{m} (\tau_i s + 1)}{s \prod_{j=1}^{n} (\tau_j s + 1)}, \qquad n > m$$

or

$$\text{OLTF} = \frac{K' \prod_{i=1}^{m} \left(s + \dfrac{1}{\tau_i}\right)}{s \prod_{j=1}^{n} \left(s + \dfrac{1}{\tau_j}\right)}, \qquad n > m \qquad \text{(8-1.6)}$$

where

$$K' = \frac{K \prod_{i=1}^{m} \tau_i}{\prod_{j=1}^{n} \tau_j}$$

From Eq. 8-1.6, the poles are recognized as equal to $-1/\tau_j$ forj $= 1$ to n and as equal to 0 for the single s term. Similarly, the zeros are given by $-1/\tau_i$ from $i = 1$ to m. These definitions of poles and zeros will be used frequently in our study of root locus and frequency response techniques.

8-2 ANALYSIS OF FEEDBACK CONTROL SYSTEMS BY ROOT LOCUS

Root *locus* is a graphical technique that consists of graphing the roots of the character-istic equation, also referred to as eigenvalues, as a gain or any other control loop parameter changes. The resulting graph allows to see at a glance whether a root crosses the imaginary axis to the right-hand side of the s-plane. This crossing would indicate the possibility of instability of the control loop.

Several examples of how to draw the root locus are first presented. General rules for plotting are then given. These examples also illustrate the effects of the different pa-rameters of the control loop on its stability. These effects were presented in Chapter 6, so the following examples should also serve as a review.

EXAMPLE 8-2.1

Consider the block diagram shown in Fig. 8-2.1. The characteristic equation for this system is

$$1 + \frac{K_c}{(3s + 1)(s + 1)} = 0 \qquad (8\text{-}2.1)$$

and

$$\text{OLTF} = \frac{K_c}{(3s + 1)(s + 1)}$$

Note that this OLTF contains two poles, at $\frac{1}{3}$ and -1, and no zeros. From Eq. 8-2.1, the polynomial

$$3s^2 + 4s + (1 + K_c) = 0$$

is obtained. This polynomial, being of second order, has two roots. Using the quadratic
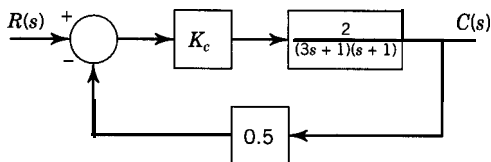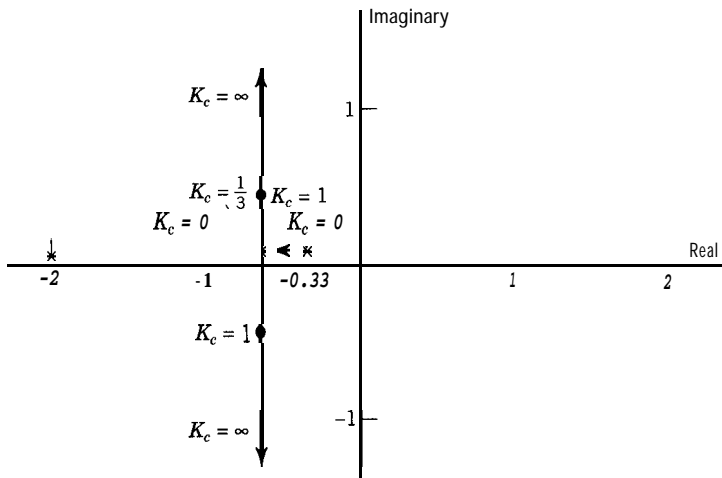

Figure 8-2.1 Block diagram of control loop-Ex-ample 8-2.1.

Figure 8-2.2 Root locus diagram for system in Figure 8-2.1.

equation to solve for the roots, we develop the expression

$$r_1, r_2 = -\frac{2}{3} \pm \frac{1}{3}\sqrt{1 - 3K_c} \tag{8-2.2}$$

Equation 8-2.2 shows that the roots of the characteristic equation depend on the value of K_c. This is the same as saying that the closed-loop response of the control loop depends on the tuning of the feedback controller! This was also shown in Chapter 6. By giving values to K_c, we can determine the loci of the roots. The graph of the roots, or root locus, is shown in Fig. 8-2.2. Several things can be learned by examining this diagram.

1. The most important point is that this particular control loop will never go unstable, no matter how high the value of K_c is set. As the value of K_c increases, the loop response becomes more oscillatory, or underdamped, but never unstable. The underdamped response is recognized because the roots of the characteristic equation move away from the real axis into the complex region as K_c increases. The fact that a control loop with a second-order (or first-order), and no dead time, characteristic equation does not go unstable, was also shown in Chapter 6 via the Routh test and direct substitution methods.
2. When $K_c = 0$, the root loci originate from the OLTF poles: $-\frac{1}{3}$ and $- 1$.
3. The number of root loci, or branches, is equal to the number of OLTF poles, $n = 2.$
4. As K_c increases, the root loci approach infinity.

EXAMPLE 8-2.2

Suppose now that the sensor/transmitter combination of the previous example has a time constant of 0.5 time unit. The block diagram is shown in Fig. 8-2.3. The new characteristic equation and open-loop transfer function are
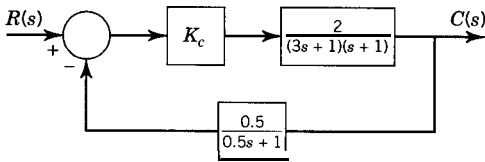
Figure 8-2.3 Block diagram of control loop—
Example 8-2.2.

Characteristic Equation

$$1 + \frac{K_c}{(3s + 1)(s + 1)(0.5s + 1)} = 0$$

or

$$1.5s^3 + 5s^2 + 4.5s + (1 + K_c) = 0$$

$$\text{OLTF} = \frac{K_c}{(3s + 1)(s + 1)(0.5s + 1)}$$

with poles: $-\frac{1}{3}, -1, -2;$ $n = 3$

zeros: none; $m = 0$

In this case the characteristic equation is a third-order polynomial, and thus the
calculation of its roots is not so straightforward. However, as we will see shortly, there
is an easier way to plot the root locus without calculating any roots.
Figure 8-2.4 shows the root locus diagram. Again, several things can be learned by
a simple glance at this diagram.

1. The most important thing is that this control system can go unstable. At some value
   of K_c, in this case K_c = 14, the root loci cross the imaginary axis. For values of K_c
   greater than 14, some roots of the characteristic equation will be on the right-hand
   side of the s-plane. The value of K_c at which the root locus crosses the imaginary
   axis, yielding a conditionally stable system, is called the ultimate gain, K_{cu}, as we
   noted in Chapter 6. The ultimate frequency, ω_u, is given by the coordinate where
   the branches cross the imaginary axis. Any loop with a characteristic equation of
   third or higher order can go unstable; first- or second-order systems, with no dead
   time, will not go unstable. Any system with dead time can go unstable, as will be
   shown in this chapter.
2. The root loci again originate, when K_c = 0, at the OLTF poles: $-\frac{1}{3}, -1, -2$.
3. The number of root loci is again equal to the number of poles of the OLTF,
   $n = 3$.
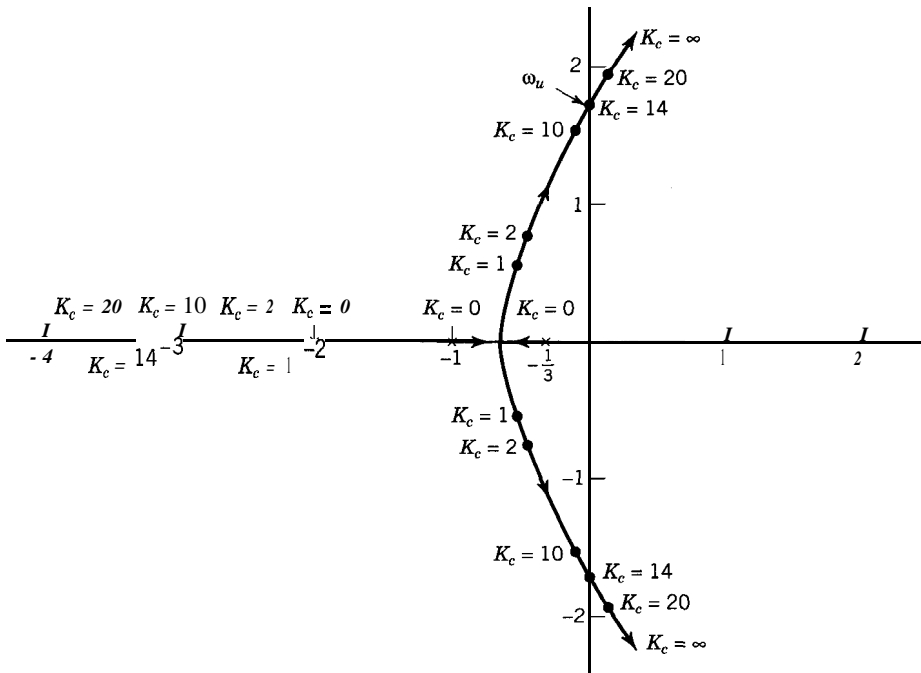4. The root loci again approach infinity as K_c increases.

Figure 8-2.4 Root locus diagram for the system in Figure 8-2.3.

EXAMPLE 8-2.3

Suppose a proportional-derivative controller, with $\tau_D = 0.2$, is now used in the original control loop of Example 8-2.1. The new characteristic equation and open-loop transfer function are

Characteristic   Equation

$$1 + \frac{K_c(1 + 0.2s)}{(3s + 1)(s + 1)} = 0$$

or

$$3s^2 + (4 + 0.2K_c)s + (1 + K_c) = 0$$

$$\text{OLTF} = \frac{K_c(1 + 0.2s)}{(3s + 1)(s + 1)}$$

$$\text{with poles:} \quad -\tfrac{1}{3}, -1; \quad n = 2$$

$$\text{zeros:} \quad -5; \quad m = 1$$

Because the characteristic equation is of second order, its roots are determined by the quadratic equation

$$r_1, r_2 = \frac{-(4 + 0.2K_c) \pm \sqrt{4 - 10.4K_c + 0.04K_c^2}}{6}$$

By giving values to K_c, we can draw the root locus for this system as shown in Fig. 8-2.5. As with the other examples, several things can be learned from this figure.

1. This control loop will never go unstable. Furthermore, as K_c increases, the root loci move away from the imaginary axis and the control loop becomes more stable. The derivative action in effect adds a "lead" term to the control loop. The addition of any lead term "adds" stability to control loops. The addition of a lag term "removes" stability from control systems, as was shown in Example 8-2.2.
2. The root loci originate at the OLTF poles: $-\frac{1}{3}$ and -1. This is similar to the previous examples.
3. The number of root loci is equal to the number of OLTF poles, $n = 2$. This is also the case for the previous examples.
4. As K_c increases, one of the root loci approaches the OLTF zero, -5, and the other root locus approaches minus infinity.
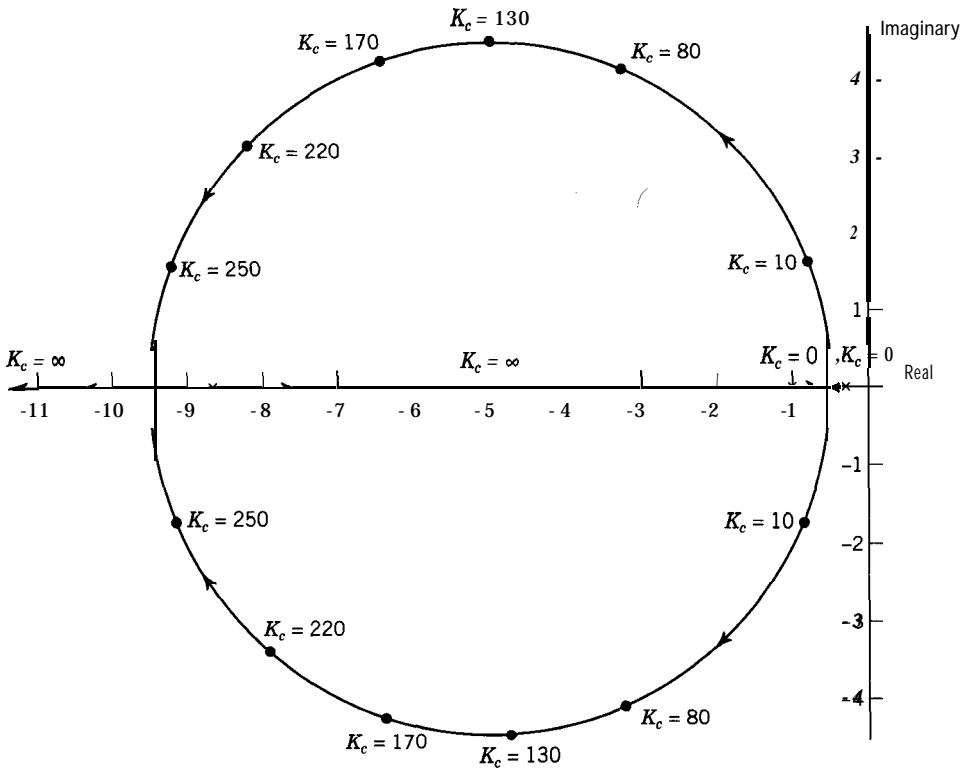
Figure 8-2.5 Root locus diagram for the control system of Example 8-2.3.

5. Figure 8-2.5 shows that as K_c increases, the loci first move from the real axis into the complex region. As the loci move away from the real axis, the response of the control system becomes more oscillatory. This continues until a value of $K_c \approx 130$. Above this value, further increases in K_c result in a less oscillatory response; the loci move closer to the real axis. Eventually there is a value of K_c, when the loci re-enter the real axis, above which the oscillations stop altogether.

8-3 RULES FOR PLOTTING ROOT LOCUS DIAGRAMS

The previous examples have shown the development of root locus diagrams. As long as the characteristic equation is of second order, it is fairly simple to develop a diagram. Several rules have been formulated to help the engineer sketch root locus diagrams without actually finding any root. To use these rules, the characteristic equation and the open-loop transfer function are first written in the following form:

Characteristic Equation

$$1 + \frac{K' \prod_{i=1}^{m} (s - z_i)}{\prod_{j=1}^{n} (s - p_j)} = 0 \tag{8-3.1}$$

and

$$\text{OLTF} = \frac{K' \prod_{i=1}^{m} (s - z_i)}{\prod_{j=1}^{n} (s - p_j)} \tag{8-3.2}$$

where

$$z_i = -\frac{1}{\tau_i} = \text{zeros}$$

$$p_j = -\frac{1}{\tau_f} = \text{poles}$$

$$K' = \frac{K \prod_{i=1}^{m} \left(-\frac{1}{z_i}\right)}{\prod_{j=1}^{n} \left(-\frac{1}{p_j}\right)}$$

K is the loop gain and is obtained by multiplying all the gains in the loop.

The rules to be presented here were developed from the fact that the root locus must satisfy what **are** called the **magnitude and angle conditions.** To understand what these conditions are, consider again the characteristic equation given by Eq. 8-3.1. This equation can also be written as

$$\frac{K' \prod_{i=1}^{m} (s - z_i)}{\prod_{j=1}^{n} (s - p_j)} = -1$$

Because this equation is complex, it can be separated into two parts: magnitude and phase angle. Carrying out the multiplication and division in polar form, as presented in Section 2-7, yields

$$\frac{K' \prod_{i=1}^{m} |s - z_i|}{\prod_{j=1}^{n} |s - p_j|} = 1 \qquad (8\text{-}3.3)$$

and

$$\sum_{i=1}^{m} \angle(s - z_i) \quad \sum_{j=1}^{n} \angle(s - p_j) = -\pi \pm 2\pi k \text{ radians} \qquad (8\text{-}3.4)$$

where k is a positive integer with values $k = 0, 1, 2, \ldots, n - m - 1$.

Equation 8-3.3 is called the **magnitude condition,** and Eq. 8-3.4 is called the **angle condition.** The roots, or eigenvalues, of the characteristic equation must satisfy both of these criteria. The angle condition is used to locate the root loci on the s-plane. The magnitude condition is then used to calculate the K' value that provides the root at a specific point in the root locus diagram.

The following example shows how the angle and magnitude conditions are used to search for a root. Consider Fig. 8-3.1, which depicts a system that has two OLTF poles, shown as (X), and one OLTF zero, shown as (0). Choose a value of s, say s_1, and determine whether it is a root and therefore a part of the root locus. The first step is to check the angle condition. To do this, form the line segments joining point s_1 with each pole and zero, as shown in the figure. Then measure the angles between each of these lines segments and the real axis. If the angle condition is satisfied, then point s_1 is part of the root locus. If the angle condition is not satisfied, another point in the s-plane is chosen and tried: a trial-and-error procedure. Once a point in the s-plane has been identified as part of the root locus, then the magnitude condition is used to calculate the value of K' that corresponds to the root. This calculation is shown later in an example.

The preceding paragraph briefly explained how the angle condition serves as the basis for sketching the root locus. The rules developed from these conditions can be
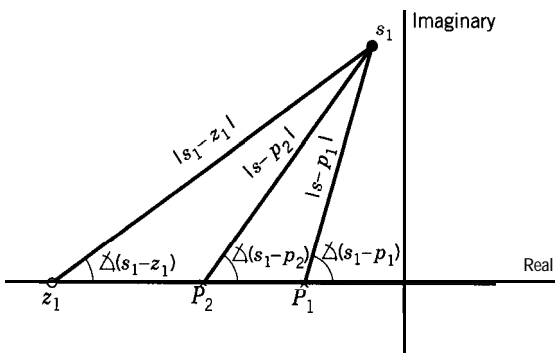

Figure 8-3.1 Search for root of characteristic equation.

used to sketch the root locus qualitatively. Certain points can then be determined from the resulting root locus, such as the ultimate gain and frequency, as well as the gain required to obtain a given damping ratio. Note that the root loci will always be symmetrical with respect to the real axis. This is a consequence of the fact that the roots of the characteristic equation are either real, or complex conjugate pairs. Here are the rules (Coughanowr,   1991):

Rule 1. On the real axis, the locus exists at a point where an odd number of poles and zeros is found to the right of the point.

Rule 2. The root loci will always originate, $K' = 0$, at the poles of the OLTF. Repeated poles originate repeated loci. That is, a qth-order pole originates q loci or branches.

Rule 3. The number of loci or branches is equal to the number of poles of the OLTF, n.

Rule 4. As the total gain of the loop increases, the loci or branches approach either the zeros of the OLTF or infinity. The number of loci that approach infinity is given by $n - m$. Repeated zeros attract repeated loci. That is, a qth-order zero attracts q loci or branches.

Rule 5. Those loci that approach infinity do so along straight-line asymptotes. All asymptotes must pass through the "center of gravity" of the poles and zeros of the OLTF. The location of this center of gravity (CG) is computed as follows

$$CG = \frac{\sum_{j=1}^{n} p_j - \sum_{i=1}^{m} z_i}{n - m} \tag{8-3.5}$$

These asymptotes make the following angle with the positive real axis:

$$\phi = \frac{180" + 360°k}{n - m} \tag{8-3.6}$$

where $k = 0, \ldots, n - m - 1$. Note that for $n - m = 1$, there is only one asymptote and it makes an angle of 180". For $n - m = 2$, there are two asymptotes with angles of ± 90"; for $n - m = 3$ there are three asymptotes with angles of ± 60" and 180"; and so on.

Rule 6. The points on the real axis where the loci meet and leave, or enter the real axis from the complex region of the s-plane, are called *breakaway points. These* breakaway points are determined, most often by trial and error, from the solution to the equation

$$\sum_{i=1}^{m} \frac{1}{s - z_i} = \sum_{j=1}^{n} \frac{1}{s - p_j} \tag{8-3.7}$$

The loci always leave or enter the real axis at the breakaway points at angles of ± 90".

When a locus leaves from a complex conjugate pole, p_k, the angle of departure relative to the real axis is found from

$$\text{Angle of departure} = 180\text{"} + \sum_{i=1}^{m} \measuredangle(p_k - z_j) - \sum_{j=1}^{n} \measuredangle(p_k - p_i) \qquad \text{with } j \neq k$$

$$(8\text{-}3.8)$$

The symbol $\measuredangle(p_k - z_i)$ signifies the angle, relative to the real axis, between polep, and zero z_i.

When a locus arrives at a complex conjugate zero, z_k, the angle of arrival relative to the real axis is found from

$$\text{Angle of arrival} = -180\text{"} + \sum_{j=1}^{n} \measuredangle(z_k - p_j) - \sum_{i=1}^{m} \measuredangle(z_k - z_i) \qquad \text{with } i \neq k$$

$$(8\text{-}3.9)$$

Let us now examine the use of these rules for plotting root locus diagrams.

EXAMPLE 8-3.1

Consider the heat-exchanger temperature control loop presented in Chapter 6. The block diagram, has been redrawn in Fig. 8-3.2 to show each transfer function.

Characteristic Equation

$$1 + \frac{0.8K_c}{(10s + 1)(30s + 1)(3s + 1)} = 0$$

$$\text{OLTF} = \frac{0.8K_c}{(10s + 1)(30s + 1)(3s + 1)}$$
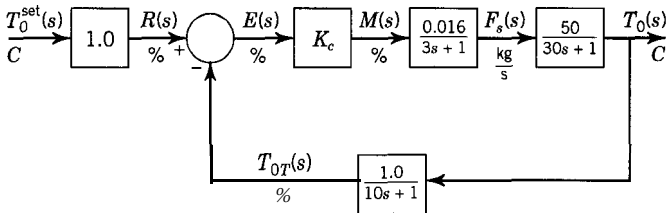
Figure 8-3.2 Block diagram of heat exchanger temperature control— P controller.
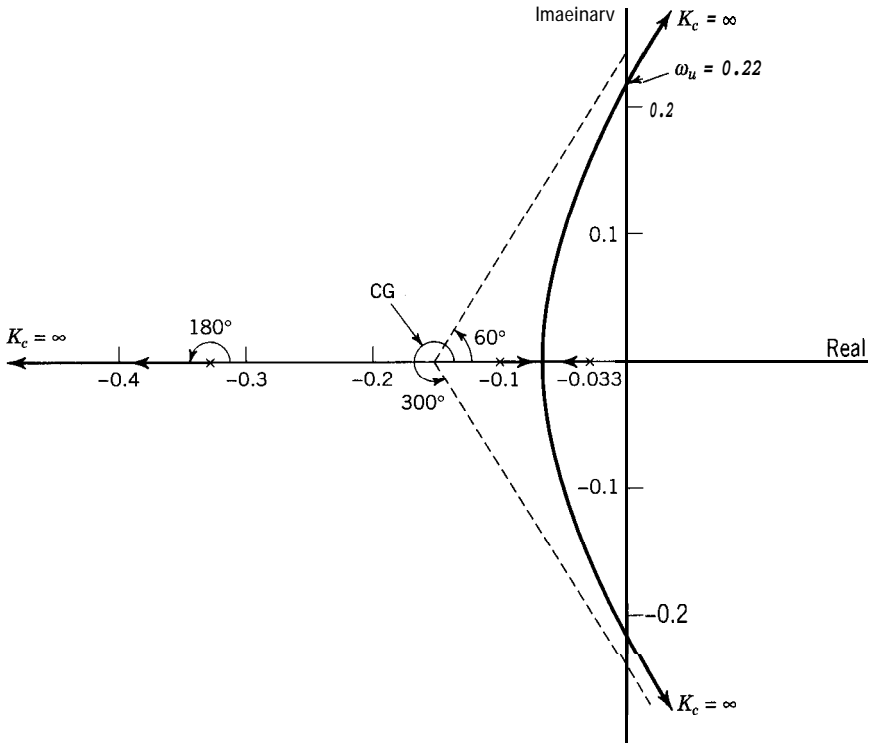
Figure 8-3.3 Root locus diagram of heat exchanger control loop.

As indicated by Eq. 8-3.2, the OLTF can also be written as

$$\text{OLTF} = \frac{K'}{\left(s + \dfrac{1}{10}\right)\left(s + \dfrac{1}{30}\right)\left(s + \dfrac{1}{3}\right)}$$

where

$$K' = \frac{0.8K_c}{(10)(30)(3)} = 0.000888K_c$$

with poles: $-\frac{1}{10}, -\frac{1}{30}, -\frac{1}{3};$ $n = 3$

zeros: none; $m = O$

Fig. 8-3.3 shows the locations of the poles (X) in the s-plane.

- Rule 1 indicates that the negative real axis between poles $-\frac{1}{30}$ and $-\frac{1}{10}$ and from pole $-\frac{1}{3}$ to $-\infty$ is part of the root locus.
- From rule 2 we know that the root loci will originate at the poles of the OLTF: $-\frac{1}{10}, -\frac{1}{30},$ and $-\frac{1}{3}.$
- Because there are three poles, $n = 3$, rule 3 indicates that there are three loci or branches.

- Because there are no zeros, $m = 0$, **rule** 4 indicates that all loci approach infinity as K_c increases.
- Rule 5 allows us to obtain the center of gravity through which the asymptotes must pass and to obtain their angles with the positive real axis. Because there are three branches that approach infinity, there must also be three asymptotes. From Eq. 8-3.5, we obtain

$$CG = \frac{-\frac{1}{10} - \frac{1}{30} - \frac{1}{3}}{3} \quad -0.155$$

and from Eq. 8-3.6, we find

$$\phi = \frac{180'' + 360°(0)}{3}, \qquad \frac{180'' + 360°(1)}{3}, \qquad \frac{180'' + 360°(2)}{3}$$

$$= 60°, \ 180°, \ 3\,0\,0\,''$$

These asymptotes and angles are shown in Fig. 8-3.3. One of the asymptotes lies on the real axis, $\phi = 180°$, and moves from the center of gravity to minus infinity. The other two asymptotes move away from the real axis into the complex region of the s-plane. These asymptotes cross the imaginary axis, indicating the possibility of instability because the loci will approach infinity along these asymptotes.

- We use rule 6 to calculate the breakaway points. Applying Eq. 8-3.7 gives us

$$-\frac{1}{s + \frac{1}{30}} + \frac{1}{s + \frac{1}{10}} + \frac{1}{s + \frac{1}{3}} = 0$$

which yields the two possibilities are -0.247 and -0.063. The only valid breakaway point is -0.063; it is the only one that lies in the region of the real axis where two loci move toward each other.

Before the final root locus diagram is drawn, it is convenient to know where the loci cross the imaginary axis. This provides one more point to draw the root loci through and enhances the accuracy of the diagram. The point is the ultimate frequency, ω_u, and is easily found by applying the direct substitution method introduced in Chapter 6. Applying this method to the present problem yields $\omega_u = 0.22$. The controller gain that produces this state of conditional stability can also be calculated by the direct substitution method; this value is $K_{cu} = 24$. Figure 8-3.3 shows the complete root locus diagram.

Example 8-3.1 demonstrates that it is fairly simple to sketch the root locus diagram and that it is not necessary to find any root of the system. The loci between the breakaway point and the crossover frequency have been sketched by drawing a smooth curve between the known points. This is usually good enough for most process control work. For convenience, a drawing instrument called the spirule can be used. The spirule is commonly used by electrical engineers. Obviously (and their use is more important and convenient), there are several software packages, such as **MATLAB** and Program CC, that draw the root locus directly from the transfer functions. These programs allow the

engineer to develop more precise root loci to get the big picture. There are also several software packages, such as TK Solver and **MATHCAD,** that make the computation and drawing of the roots very convenient. (The sources of all these packages are given in the References section of this chapter.)

The diagram shown in Fig. 8-3.3 helps to illustrate another use of root locus. Suppose it is desired to tune the feedback controller so that the closed-loop control system response will be oscillatory with a damping ratio of $0.707, \zeta = 0.707$. Chapter 2 defined damping ratio as a parameter of a second-order system, and Fig. 2-5.3 shows the responses for different damping ratios. In using this performance specification, we assume either that the process is of second order or that there are two time constants that are much longer than the others. These two time constants will "dominate" the dynamics of the process. The roots, a pair of complex conjugate roots, associated with these time constants were referred to in Chapter 2 as *the dominant roots*. Tuning a feedback controller for the foregoing specification means that the two dominant roots of the characteristic equation must satisfy the equation

$$\tau^2 s_1 + 2\tau\zeta s_1 + 1 = 0$$

where τ is the characteristic time of the closed-loop. The dominant roots are

$$s_1, s_1^* = -\frac{\zeta}{\tau} \pm \frac{1}{\tau}\sqrt{\zeta^2 - 1}$$

These roots are shown graphically in Fig. 8-3.4. From this figure, we can determine that

$$\theta = \cos^{-1}\zeta \qquad\qquad\qquad \textbf{(8-3.10)}$$
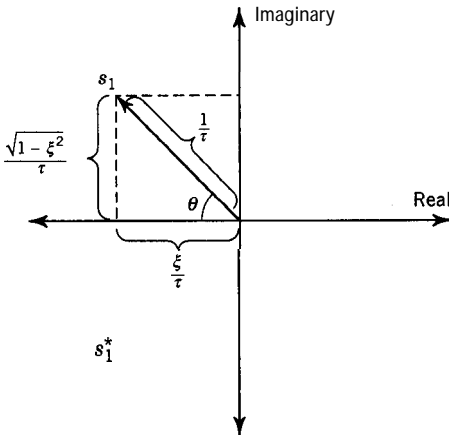
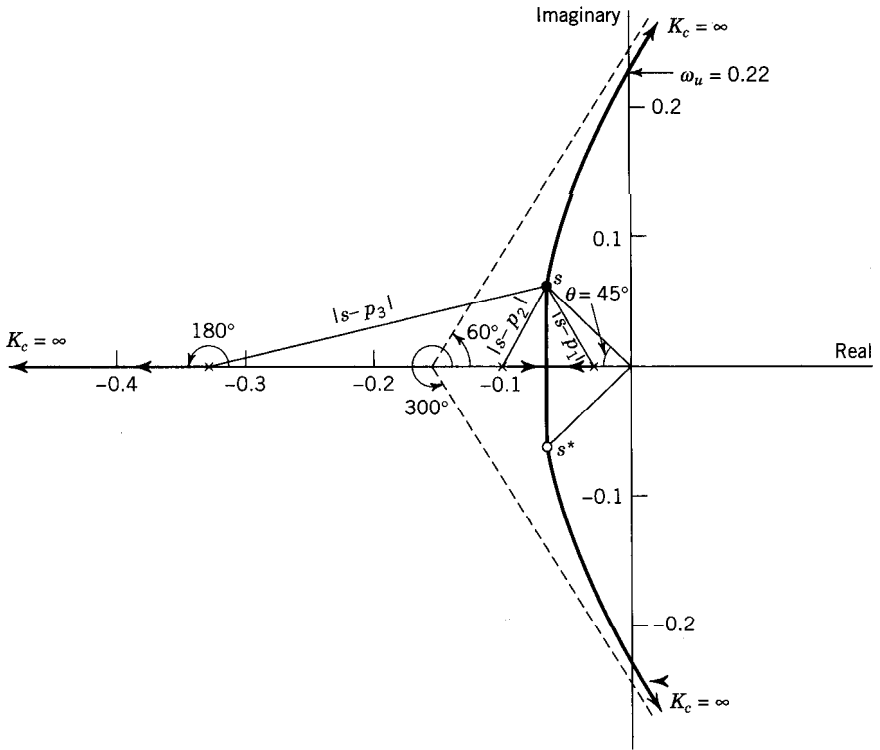**Figure 8-3.4** Roots, s_1 and s_1^*, of second-order system.

Figure 8-3.5 Use of root locus diagram for tuning controller.

This equation applies for any damping ratio. Then, for $\zeta = 0.707$,

$$\theta = \cos^{-1} 0.707 = 45"$$

Figure 8-3.5 shows how to find the roots, s_1 and s_1^*, of the system with this damping factor. In this case the roots are roughly located at $-0.06 \pm 0.06i$. The gain of the controller that yields this closed-loop behavior must now be calculated, the magnitude criterion, Eq. 8-3.3, is used. To make this calculation, we must measure the distance between the root s_1 and each pole and zero. This measurement is simply done with a ruler (using the same magnitude scale as the axis) or by using the Pythagorean theorem. The latter method is preferred because it minimizes measurement errors. For this system, the magnitude criterion is

$$\frac{K'}{|s - p_1||s - p_2||s - p_3|} = 1$$

Because the first pole occurs at $p_1 = -0.033 + 0i$, the distance $|s - p_1|$ is calculated by the Pythagorean theorem to be

$$\text{I}s - p_1| = \sqrt{(0.060 - 0.033)^2 + (0.06 - 0)'} = 0.066$$

Similarly,

$$Is - p_2| = \sqrt{(0.060 - 0.1)^2 + (0.06 - 0)^2} = 0.072$$

$$Is - p_3| = \sqrt{(0.060 - 0.33)^2 + (0.06 - 0)^2} = 0.276$$

Then

$$\frac{K'}{(0.066)(0.072)(0.276)} = 1$$

$$K' = 0.00131$$

and because

$$K' = 0.000888K_c$$

the controller gain is

$$K_c = 1.475$$

This controller gain yields an oscillatory response of the control loop with a damping factor of 0.707.

Let's look at one more example.

EXAMPLE 8-3.2

Suppose it is desired to use a PI controller to control the heat exchanger of Example 8-3.1. A reset time of 1 min (60 s) is used, so the transfer function for this controller is $G_c(s) = K_c(1 + \dfrac{1}{60s})$. Plot the root locus diagram for this new control system. What is the effect of adding reset action to the controller?

Characteristic Equation

$$1 + \frac{0.8K_c\left(1 + \dfrac{1}{60s}\right)}{(10s + 1)(30s + 1)(3s + 1)} = 0$$

$$OLTF = \frac{0.8K_c\left(1 + \dfrac{1}{60s}\right)}{(10s + 1)(30s + 1)(3s + 1)}$$

or

$$\text{OLTF} = \frac{K'\left(s + \dfrac{1}{60}\right)}{s\left(s + \dfrac{1}{10}\right)\left(s + \dfrac{1}{30}\right)\left(s + \dfrac{1}{3}\right)}$$

where

$$K' = \frac{0.8K_c}{(10)(30)(3)} = 0.000888K_c$$

with poles: $0, -\dfrac{1}{10}, -\dfrac{1}{30}, -\dfrac{1}{3};$ $n = 4$

zeros: $-\dfrac{1}{60};$ $m = 1$

Figure 8-3.6 shows the locations of the poles (×) and zeros (0) in the s-plane.
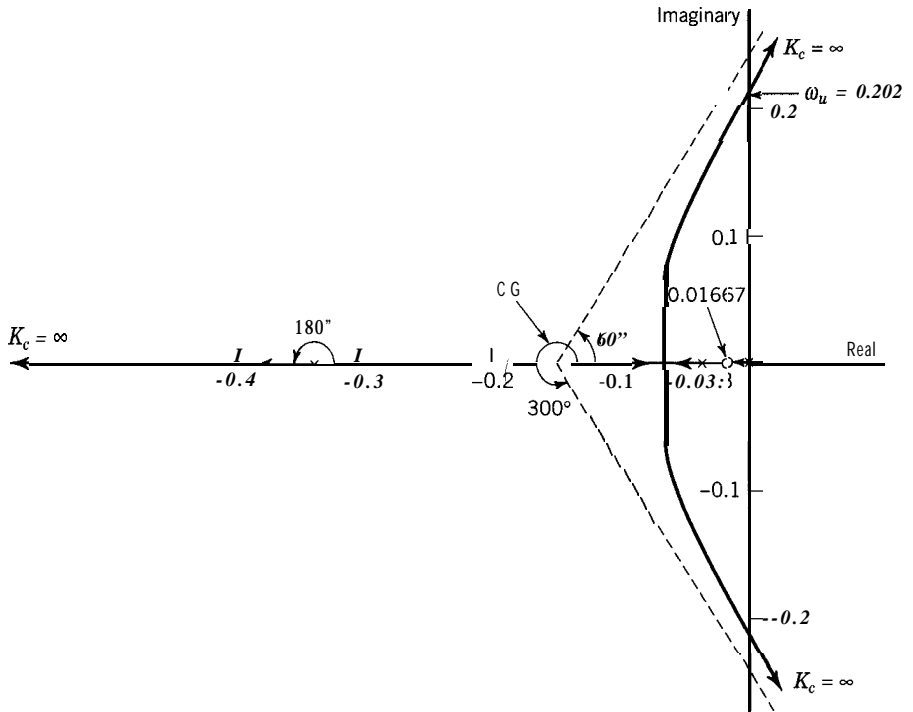
Figure 8-3.6 Root locus diagram of heat exchanger temperature control loop-PI controller.

- Rule 1 indicates that the negative real axis between the pole at 0 and the zero at $-\frac{1}{60}$ will be part of the root locus. This is also the case between the poles at $\frac{1}{30}$, and at $-\frac{1}{10}$, and from the pole at $-\frac{1}{3}$ to $-\infty$.
- Rule 2 indicates that the root loci originate at 0, $-\frac{1}{10}$, $-\frac{1}{30}$, and $-\frac{1}{3}$, the poles of the OLTF.
- Because there are four poles, $n = 4$, rule 3 indicates that there are four branches or loci.
- Because there is one zero, $m = 1$, rule 4 indicates that one of the branches will terminate at this zero. In this case, this is the branch that originates at the pole equal to zero. The other three branches, $n - m = 3$, will approach infinity as K_c increases.
- Using rule 5, we can determine the center of gravity as well as the angles the asymptotes make with the positive real axis. Because there are three branches that approach infinity, there must be three asymptotes. According to Eq. 8-3.5,

$$CG = \frac{-10 - 30 - \frac{1}{3} + \frac{1}{60}}{4 - 1} = -0.15$$

and because $n - m = 3$, the asymptotes make the angles of $60°$, $180°$, and $300°$ with the positive real axis. These asymptotes and angles are shown in Fig. 8-3.6. One of the asymptotes lies on the real axis and moves from the center of gravity to minus infinity. The other two asymptotes move away from the real axis into the complex region of the s-plane, crossing the imaginary axis and thus indicating the possibility of instability.

- Rule 6 provides the breakaway points. Applying Eq. 8-3.7 gives

$$\frac{1}{s + \frac{1}{60}} = \frac{1}{s + 0} + \frac{1}{s + \frac{1}{10}} + \frac{1}{s + \frac{1}{30}} + \frac{1}{s + \frac{1}{3}}$$

Four possibilities are obtained: $-0.0137 \pm 0.0149i$, -0.0609, and -0.245. The only valid breakaway point is at -0.0609, because it is the only point that lies in the region of the real axis where two loci move toward each other.

Applying the direct substitution method to the characteristic equation of this system, we find the crossover or ultimate frequency to be $\omega_u = 0.202$. The controller gain that yields this condition is $K_{cu} = 20.12$.

Comparing Figs. 8-3.3 and 8-3.6, we see that adding the reset action to the proportional-only controller does not significantly change the shape of the root locus. The most significant effect is the decrease in the ultimate gain and ultimate frequency.

8-4 SUMMARY

This chapter has presented the root locus technique for process control analysis and design. The development of the root locus diagram has been shown to be quite simple, without the need for sophisticated mathematics. Probably the major advantage of the

method is its graphical nature. Its major disadvantage is that it cannot be applied directly to processes with dead time. In this respect, root locus is similar to Routh's test and direct substitution.

REFERENCES

1. Coughanowr, D. R. 1991. *Process Systems Analysis and Control.* New York: McGraw-Hill, Chapter 15.
2. MATHCAD, MathSoft, Inc., 101 Main Street, Cambridge, Mass. 02142.
3. MATLAB, The MathWorks, Inc., Cochituate Place, 24 Prime Park Way, Natick, Mass. 01760-1520.
4. Program CC, Systems Technology, Inc., Hawthorne, Calif.
5. TK Solver, Universal Technical Systems, 1220 Rock Street, Rockford, Ill. 61101.

PROBLEMS

8-1. Draw the root locus diagram for each of the following open-loop transfer functions.

(a) $G(s) = \dfrac{K}{s(s + 1)(4s + 1)}$

(b) $G(s) = \dfrac{K(3s + 1)}{(2s + 1)}$

(c) $G(s) = K\left[1 + \dfrac{1}{\tau_I s}\right]\left[\dfrac{\tau_D s + 1}{\alpha \tau_D s + 1}\right]$

8-2. Sketch the root locus diagram for each of the following open-loop transfer functions (use a first-order Padé approximation, Eq. 6-2.6, for the dead-time term).

(a) $G(s) = \dfrac{K}{(s + 1)(2s + 1)(10s + 1)}$

(b) $G(s) = \dfrac{K(3s + 1)}{(s + 1)(2s + 1)(10s + 1)}$

(c) $G(s) = \dfrac{Ke^{-s}}{(s + 1)(2s + 1)(10s + 1)}$

(d) $G(s) = \dfrac{K(3s + 1)e^{-s}}{(s + 1)(2s + 1)(10s + 1)}$

8-3. Consider the following transfer function of a certain process.

$$G(s)H(s) = \dfrac{1.5}{(s + 1)(5s + 1)(10s + 1)}$$

(a) Tune a proportional-only controller, using as the gain half the ultimate gain.

(b) Determine the damping factor, ζ, of the dominant roots of the control loop, using the tuning parameter obtained in part (a). Sketch these roots in the root locus for the loop.

8-4. Sketch the root locus diagram for the following two open-loop transfer functions.
(a) System with inverse response

$$G(s) = \frac{K(1 - 0.25s)}{(2s + 1)(s + 1)}$$

(b) Open-loop unstable system

$$G(s) = \frac{K}{(\tau_1 s + 1)(1 - \tau_2 s)}$$

for two cases: $\tau_1 = 2$, $\tau_2 = 1$ and $\tau_1 = 1$, $\tau_2 = 1$

8-5. Consider the pressure control system shown in Fig. P8-1. The pressure in the tank can be described by

$$\frac{P(s)}{F(s)} = \frac{0.4}{(0.15s + 1)(0.8s + 1)}, \text{ psi/scfm}$$

The valve can be described by the following transfer function:

$$\frac{F(s)}{M(s)} = \frac{0.6}{0.1s + 1}, \text{ scfm/\%CO}$$

The pressure transmitter has a range of 0 to 185 psig. The dynamics of the transmitter are negligible.

(a) Draw the block diagram for this system, including all the transfer functions.
(b) Sketch the root locus diagram.
(c) Determine the gain of the controller at the breakaway point.
(d) Determine the ultimate gain and ultimate period.
(e) Calculate the tuning of a P controller so as to obtain a damping factor of 0.707 for **the** dominant roots.
(f) Explain graphically how adding reset action to the controller affects the sta-bility of the control loop. Use $\tau_I = 0.8$ min.
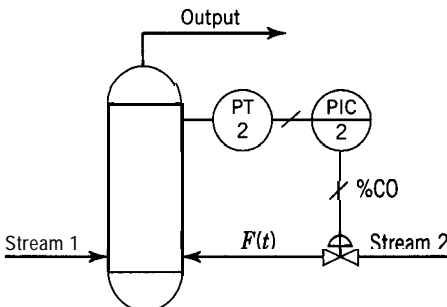
Figure P8-1 Pressure control system for Prob-lem 8-5.

(g) Explain graphically how adding rate action to the PI controller of part (f) affects the stability of the control loop.

8-6. Do Problem 8-3 for the following transfer function:

$$G(s)H(s) = \frac{0.8e^{-1}}{10s + 1}$$

Approximate the dead time with a first-order Padt approximation, Eq. 6-2.6.

8-7. Sketch the root locus diagram for the loop of Problem 6-3.

8-8. The open-loop transfer function for Problem 7-25 on the control of the level in a tank using a PI controller is

$$\text{OLTF} = K_c \left(1 + \frac{1}{\tau_I s} \right) \left(\frac{K}{s(\tau_v s + 1)} \right)$$

where τ_I is the reset time and τ_v is the valve time constant. Sketch the root locus for each of the following cases:
(a) $\tau_I > \tau_v$
(b) $\tau_I = \tau_v$
(c) $\tau_I < \tau_v$
Discuss, on the basis of your sketches, how the stability of the loop is affected by the relationship between the controller integral time and the process time constant.

8-9. Example 8-2.3 shows the effect on the stability of a system of adding the derivative mode to a controller. In that example, $\tau_D = 0.2$ was used. Discuss the effects of increasing the value used for τ_D, for example, using $\tau_D = 0.5$.

8-10. Sketch the root locus diagram for the gas flow control loop of Problem 6-6.

8-11. Sketch the root locus diagram for the steam flow control loop of Problem 6-7.

8-12. Sketch the root locus diagram for the loops of Problem 6-8. Use the first-order Padt approximation to the dead time when necessary.

8-13. Sketch the root locus diagram for the analyzer control loop of Problem 6-11.

8-14. Sketch the root locus diagram for the reactor temperature control loop of Problem 6-12.

8-15. Sketch the root locus diagram for the composition control loop of Problem 6-14.

8-16. Sketch the root locus diagram for the reactor composition control loop of Problem 6-15.

8-17. Sketch the root locus diagram for the pressure control loop of Problem 6-18.

8-18. Sketch the root locus diagram for the temperature control loop of Problem 6-19.

8-19. Sketch the root locus diagram for the pressure control loop of Problem 6-20.

8-20. Sketch the root locus diagram for each of the control loops of Problem 6-24.

Chapter 9

Frequency Response Techniques

Frequency response techniques are some of the most popular techniques for the analysis and design of control of linear systems. This chapter describes what is meant by frequency response and how to use these techniques to analyze and synthesize control systems. The chapter also explains how to use these techniques as another "tool" for process identification.

9-1 FREQUENCY RESPONSE

9-1.1 Experimental Determination of Frequency Response

Consider the general block diagram shown in Fig. 9-1.1. The control loop has been opened before the valve and after the transmitter. A variable-frequency generator provides the input signal to the valve, $x(t) = X_0 \sin \omega t$, and a recorder records the output signal from the transmitter and the input signal to the valve. Figure 9-1.2 shows the two recordings. After the transients have died out, the transmitter output reaches a sinusoidal response, $y(t) = Y_0 \sin(\omega t + \theta)$. This experiment is referred to as sinusoidal testing.

Let us now perform the same "experiment" using the transfer function that describes the process. Assume the following simple transfer function:

$$G(s) = \frac{Y(s)}{X(s)} = \frac{K}{\tau s + 1} \tag{9-1.1}$$

This transfer function describes the valve, process, and sensor/transmitter combination. The input signal to the valve is

$$x(t) = X_0 \sin(\omega t)$$
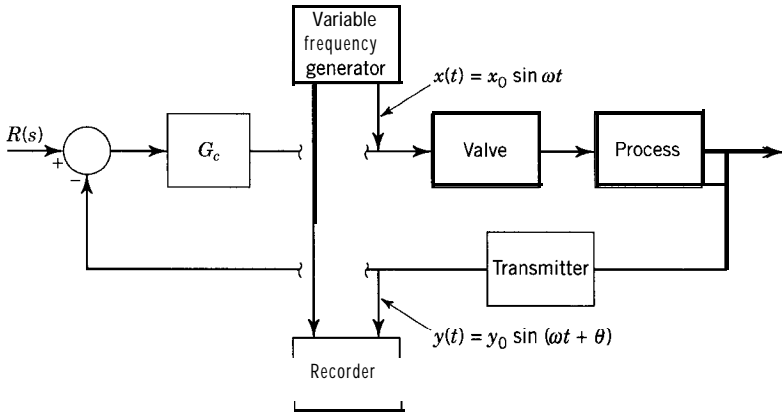
Figure **9-1.1** Block diagram showing variable frequency generator and recorder.

and its Laplace transform, from Table 2-1 .1, is

$$X(s) = \frac{X_0 \omega}{s^2 + \omega^2}$$
(9-1.2)

Therefore,

$$Y(s) = \frac{KX_0 \omega}{(\tau s + 1)(s^2 + \omega^2)}$$

The time domain expression for Y(t) can be obtained using the techniques learned in Chapter 2.

$$Y(t) = \frac{KX_0 \omega \tau}{1 + \omega^2 \tau^2} e^{-t/\tau} + \frac{KX_0}{\sqrt{1 + \omega^2 \tau^2}} \sin(\omega t + \theta)$$
(9-1.3)

with

$$\theta = \tan^{-1}(-\omega\tau) = -\tan^{-1}(\omega\tau)$$
(9-1.4)
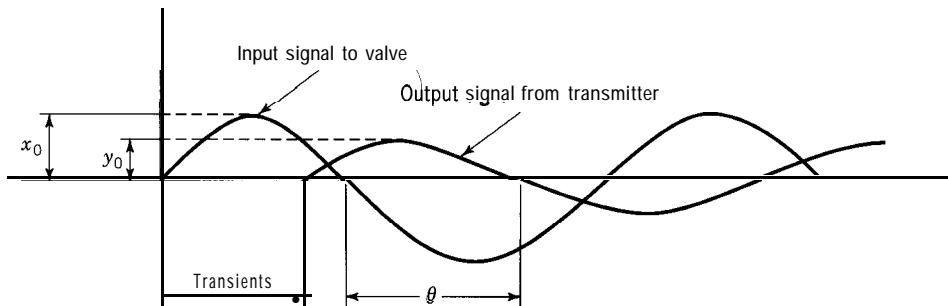
Figure 9-1.2 Recordings from sinusoidal testing.

Equation 9-1.3 was presented in Chapter 2 as Eq. 2-4.9. As time increases, the exponential term of Eq. 9-1.3 goes to zero; that is, the transient term dies out. When this happens, the output expression becomes

$$Y(t)\Big|_{t\text{ very large}} = \frac{KX_0}{\sqrt{1 + \omega^2\tau^2}} \sin(\omega t + \theta) \tag{9-1.5}$$

which constitutes the sinusoidal behavior of the output signal, and the one shown in Fig. 9-1.2. The amplitude of this output signal is

$$Y_0 = \frac{KX_0}{\sqrt{1 + \omega^2\tau^2}}$$

The minus sign in Eq. 9-1.4 indicates that the output signal "lags" the input signal by the amount θ calculated from the equation.

A word of advice is necessary here. Care should be taken when calculating the sine term in Eq. 9-1.5. The term ω is in radians/time, and the term ωt is in radians. Thus, for the operation $(\omega t + \theta)$ to be in the correct units, θ must be in radians. If degrees are to be used, then the term must be written as

$$\left(\frac{180}{\pi}\omega t + \theta\right)$$

In short, be careful with the units.

Some terms often used in frequency response studies will now be defined.

Amplitude ratio (AR) is the ratio of the amplitude of the output signal to the amplitude of the input signal. That is,

$$AR = \frac{Y_0}{X_\theta} \tag{9-1.6}$$

Magnitude ratio (MR) is the amplitude ratio divided by the steady-state gain.

$$MR = \frac{AR}{K} \tag{9-1.7}$$

Phase angle (θ) is the amount, in radians or degrees, by which the output signal lags or leads the input signal. When θ is positive, it is a lead angle; when θ is negative, it is a lag angle.

For the foregoing first-order transfer function,

$$AR = \frac{K}{\sqrt{1 + \omega^2\tau^2}} \qquad MR = \frac{1}{\sqrt{1 + \omega^2\tau^2}} \qquad \theta = \tan^{-1}(-\omega\tau)$$

Note that all three terms are functions of the input frequency. Different processes have different AR (MR) and θ dependence on ω.

   Frequency response is essentially the study of how the AR (MR) and θ of different components or systems behave as the input frequency changes. The following paragraphs show that frequency response is a powerful technique for analyzing and synthesizing control systems. We will discuss the development of the frequency response of process systems first and then its use for analysis and synthesis.

   There are in general two different methods for generating the frequency response.

1.   ***Experimental   methods.*** These methods consist essentially of the experiment with the variable-frequency generator and the recorder. The idea is to run the experiment at different frequencies in order to obtain a table of AR vs. ω and θ vs. ω. These experimental methods are reviewed again later in this chapter; they also provide a way to identify process systems.

2. **Transforming the open-loop transfer function after a sinusoidal input.** This method consists of using the open-loop transfer function to obtain the expression that describes the response of the system to a sinusoidal input. From the expression, the amplitude and phase angle of the output can then be determined. This method is the mathematical manipulations previously shown that resulted in Eqs. 9-1.4 and 9-1.5.

   Fortunately, operational mathematics provide a very simple way to determine AR (MR) and θ without having to obtain the inverse Laplace transforms. The necessary mathematics have already been presented in Chapter 2. Consider

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

for $x(t) = X_0 \sin(\omega t)$. From Table 2-1.1,

$$X(s) = \frac{X_0 \omega}{s^2 + \omega^2}$$

Then

$$Y(s) = G(s)\,\frac{X_0 \omega}{s^2 + \omega^2}$$

Expansion by partial fractions yields

$$Y(s) = \frac{A}{s + i\omega} + \frac{B}{s - i\omega} + [\text{terms for the poles of } G(s)] \qquad \textbf{(9-1.8)}$$

To obtain *A. we* use

$$A = \lim_{s \to -i\omega} \frac{[(s + i\omega)X_0 \omega G(s)]}{(s^2 + \omega^2)} = \frac{G(-i\omega)X_0 \omega}{-2iw}$$

As shown in Chapter 2, any complex number can be represented by a magnitude and an argument; then

$$A = \frac{X_0|G(i\omega)|e^{-i\theta}}{-2i}$$

where $\theta = \measuredangle G(i\omega)$. To obtain B, we use

$$B = \lim_{s \to i\omega} \left[\frac{(s - i\omega)X_0\omega G(s)}{(s^2 + \omega^2)} \right] = \frac{X_0|G(i\omega)|e^{i\theta}}{2i}$$

Then, substituting the expressions for A and B into Eq. 9-1.8 yields

$$Y(s) = \frac{X_0|G(i\omega)|}{2i} \left[\frac{-}{s + i\omega} + \frac{e^{i\theta}}{s - i\omega} \right] + \text{[terms of G(s)]}$$

Inverting back into the time domain, we get

$$Y(t) = \frac{X_0|G(i\omega)|}{2i} [- e^{-i\theta}e^{-i\omega t} + e^{i\theta}e^{i\omega t}]$$

$$+ \text{[transient terms resulting from the terms of } G(s)]$$

After the transient terms die out,

$$Y(t)|_{t\,very\,large} = X_0|G(i\omega)| \frac{e^{i(\omega t + \theta)} - e^{-i(\omega t + \theta)}}{2i}$$

or

$$Y(t)|_{t\,very\,large} = X_0|G(i\omega)| \sin(\omega t + \theta)$$

The amplitude ratio is then

$$\boxed{AR = \frac{Y_0}{X_0} = \frac{X_0|G(i\omega)|}{X_0} = |G(i\omega)|} \tag{9-1.9}$$

and the phase angle is

$$\boxed{\theta = \measuredangle G(i\omega)} \tag{9-1.10}$$

Thus, in order to obtain AR *and* θ, *one simply substitutes* $i\omega$ *for* s *in the transfer function and then calculates the magnitude and argument of the resulting complex-number expression. The magnitude is equal to the amplitude ratio* (AR), *and the argument equals the phase angle* (0).

This manipulation simplifies the required calculations.

Let us apply these results to the first-order system used earlier.

$$G(s) = \frac{K}{\tau s + 1}$$

Now substitute $i\omega$ for s.

$$G(i\omega) = \frac{K}{i\omega\tau + 1}$$

This results in a complex-number expression composed of the ratio of two terms: the numerator, a real number, and the denominator, a complex number. The equation can also be written as follows:

$$G(i\omega) = \frac{G_1}{G_2} = \frac{K}{i\omega\tau + 1} \qquad (9\text{-}1.11)$$

As shown by Eq. 9-1.9 the amplitude ratio is equal to the magnitude of this **complex-number** expression

$$AR = |G(i\omega)| = \frac{|G_1|}{|G_2|} = \frac{K}{\sqrt{\omega^2\tau^2 + 1}} \qquad (9\text{-}1.12)$$

which is the same AR as previously obtained.

The phase angle is equal to the angle of the complex-number expression

$$\theta = \sphericalangle G(i\omega) = \sphericalangle G_1 - \sphericalangle G_2 = 0 - \tan^{-1}(\omega\tau) = \tan^{-1}(\omega\tau) \qquad (9\text{-}1.13)$$

which is also the same θ as previously obtained.

Let us now look at several other examples.

EXAMPLE 9-1.1

Consider the second-order system

$$G(s) = \frac{K}{\tau^2 s^2 + 2\tau\zeta s + 1}$$

Determine the expressions for AR and 13.

SOLUTION

The first step is to substitute $i\omega$ for s.

$$G(i\omega) = \frac{K}{-\omega^2\tau^2 + i2\tau\zeta\omega + 1} = \frac{K}{(1 - \omega^2\tau^2) + i2\tau\zeta\omega}$$

Again, a complex-number expression results that is a ratio of two other numbers.

$$G(i\omega) = \frac{G_1}{G_2} = \frac{K}{(1 \quad \omega^2\tau^2) + i2\tau\zeta\omega}$$

The amplitude ratio is

$$AR = |G(i\omega)| = \frac{|G_1|}{|G_2|} = \frac{K}{\sqrt{(1 - \omega^2\tau^2)^2 + (2\tau\zeta\omega)^2}} \qquad (9\text{-}1.14)$$

and the phase angle is

$$\theta = \angle G(i\omega) = \angle G_1 - \angle G_2$$

$$e = 0 - \tan^{-1}\left(\frac{2\tau\zeta\omega}{1 - \omega^2\tau^2}\right)$$

$$\theta = -\tan^{-1}\left(+\frac{2\tau\zeta\omega}{- \omega^2\tau^2}\right) \qquad (9\text{-}1.15)$$

This result was presented as Eq. 2-5.21 in Chapter 2.

EXAMPLE 9-1.2

Consider the first-order lead transfer function

$$G(s) = K(1 + \tau s) \qquad (9\text{-}1.16)$$

SOLUTION

This is a transfer function composed of a gain times a first-order lead. Determine the expressions for AR and θ.

Substituting $i\omega$ for s results in the complex-number expression

$$G(i\omega) = K(1 + i\omega\tau)$$

which can also be thought of as being formed by two other numbers:

$$G(i\omega) = G_1G_2 = K(1 + i\omega\tau)$$

The amplitude ratio is

$$\text{AR} = |G(i\omega)| = |G_1| \, |G_2| = K\sqrt{1 + \omega^2\tau^2} \qquad \text{(9-1.17)}$$

and the phase angle is

$$\theta = \angle G(i\omega) = \angle G_1 + \angle G_2 = 0 + \tan^{-1}(\omega\tau)$$

$$\theta = \tan^{-1}(\omega\tau) \qquad \text{(9-1.18)}$$

The phase angles of the systems described by Eqs. 9-1.1 and 9-1.16 can be compared. Systems described by Eq. 9-1.1, which are referred to in Chapters 2 and 3 as first-order lags, provide negative phase angles, as shown by Eq. 9-1.13. Systems described by Eq. 9-1.16 provide positive phase angles, as shown by Eq. 9-1.18. This fact is important in the study of process control stability by frequency response techniques.

Equation 9-1.17 helps explain why real systems cannot be pure leads. This equation shows that the amplitude ratio increases with frequency, which means that high-frequency noise, which is always present in natural signals, would be infinitely amplified.

EXAMPLE 9-1.3

Determine the expressions for AR and θ for a pure dead time.

$$G(s) = e^{-t_0 s}$$

SOLUTION

Substituting $i\omega$ for s yields

$$G(i\omega) = e^{-it_0\omega}$$

Because this expression is already in polar form, using the principles learned in Chapter 2, we obtain

$$G(i\omega) = |G(i\omega)|e^{iG(i\omega)} = e^{-it_0\omega}$$

which means that

$$AR = |G(i\omega)| = 1 \qquad\qquad \text{(9-1.19)}$$

and

$$\theta = \sphericalangle G(i\omega) = -t_0\omega \qquad\qquad \text{(9-1.20a)}$$

Recall the discussion on the units of θ. As written in Eq. 9-1.20a, the unit of θ is the radian. If it is desired to obtain θ in degrees, then

$$\theta = \left(\frac{180°}{\pi}\right)(-t_0\omega) \qquad\qquad \text{(9-1.20b)}$$

It is important to notice that θ becomes increasingly negative as ω increases. The rate at which θ drops depends on t_0; the larger t_0 is, the faster θ drops. This fact will become important in the analysis of control systems. The amplitude ratio and magnitude ratios are independent of frequency when the transfer function is a pure dead time.

EXAMPLE 9-1.4

Determine the expressions for AR and θ for an integrator

$$G(s) = \frac{1}{s}$$

SOLUTION

Substituting $i\omega$ for s yields

$$G(i\omega) = \frac{1}{i\omega} = -\frac{1}{\omega}i$$

This is a pure imaginary number with amplitude ratio

$$AR = |G(i\omega)| = \frac{1}{\omega} \qquad\qquad \text{(9-1.21)}$$

and the phase angle, because it lies on the negative imaginary axis, is

$$\theta = \sphericalangle G(i\omega) = -\frac{\pi}{2}, \text{ or } -90° \qquad\qquad \text{(9-1.22)}$$

For an integrator, then, the amplitude ratio is inversely proportional to the frequency, whereas the phase angle remains constant at $-90°$. That is, the integrator provides a constant phase lag.

At this point, we can generalize the expressions for AR and θ. Consider the following general OLTF:

$$\text{OLTF}(s) = \frac{K \displaystyle\prod_{i=1}^{m} (\tau_i s + 1) e^{-t_0 s}}{s^k \displaystyle\prod_{j=1}^{n} (\tau_j s + 1)} \qquad (n + k) > m \qquad \textbf{(9-1.23)}$$

Substituting $i\omega$ for s yields

$$\text{OLTF}(i\omega) = \frac{K \displaystyle\prod_{i=1}^{m} (i\tau_i \omega + 1) e^{-i\omega t_0}}{(i\omega)^k \displaystyle\prod_{j=1}^{n} (i\tau_j \omega + 1)}$$

and finally we arrive at

$$\text{AR} = \frac{K \displaystyle\prod_{i=1}^{m} \sqrt{(\tau_i \omega)^2 + 1}}{\omega^k \displaystyle\prod_{j=1}^{n} \sqrt{(\tau_j \omega)^2 + 1)}} \qquad \textbf{(9-1.24)}$$

and

$$\theta = \sum_{i=1}^{m} \tan^{-1}(\tau_i \omega) - t_0 \omega - \sum_{j=1}^{n} \tan^{-1}(\tau_j \omega) \qquad k(\pi) \qquad \textbf{(9-1.25)}$$

So far, expressions for AR and θ as a function of ω have been developed. There are several ways to represent these expressions graphically. The three most common ways are Bode plots, the Nyquist plot, and the Nichols chart. Bode plots are presented in detail in the next section.

9-1.2 Bode Plots

Bode plots are common graphical representations of AR (MR) and θ functions. A Bode plot consists of two graphs: (1) log AR (or log MR) vs. log ω, and (2) θ vs. log ω. Frequently the term 20 log AR, referred to as decibels, is plotted instead of log AR. This term is used extensively in the electrical engineering field and sometimes in the process control field; this book plots log AR. Let's look at the Bode plots of some of the most common process transfer functions.

Gain Element. A pure gain element has the transfer function

$$G(s) = K$$

Substituting $i\omega$ for s gives

$$G(i\omega) = K$$

and using the mathematics previously presented yields

$$AR = |G(i\omega)| = K, \text{ or } MR = 1$$

and

$$\theta = \tan^{-1}G(i\omega) = 0$$

Fig. 9-1.3a shows the Bode plot for this element; log-log and semi-log graph papers have been used.

First-Order Lag. For a first-order lag, the AR and θ are given by Eqs. 9-1.12 and 9-1.13, respectively. From the AR equation the MR expression is obtained:

$$MR = \frac{1}{\sqrt{\omega^2\tau^2 + 1}}$$

and

$$\theta = -\tan^{-1}(\omega\tau) \tag{9-1.13}$$

The Bode plot for this system is shown in Fig. 9-1.3b. The magnitude ratio plot also shows two dotted lines. These lines are asymptotes to the frequency response at low and high frequencies. The figure shows that these asymptotes do not deviate much from the actual frequency response. Therefore, frequency response analysis is quite often done with the asymptotes; they are easier to draw, and not much error is involved in their use. Let us now see how these asymptotes are developed.

From the magnitude ratio equation we know that as $\omega \to 0$, $MR \to 1$, which results in the low-frequency asymptote. Before the high-frequency asymptote is developed, the magnitude ratio equation is written in log form:

$$\log MR = -\frac{1}{2}\log(\tau^2\omega^2 + 1)$$

Now, as $\omega \to \infty$,

$$\log MR \to -\frac{1}{2}\log \tau^2\omega^2 = -\log \tau\omega$$

$$\tag{9-1.26}$$

$$\log MR \to -\log \tau - \log \omega$$

(a) Gain element.

(b) First-order lag.
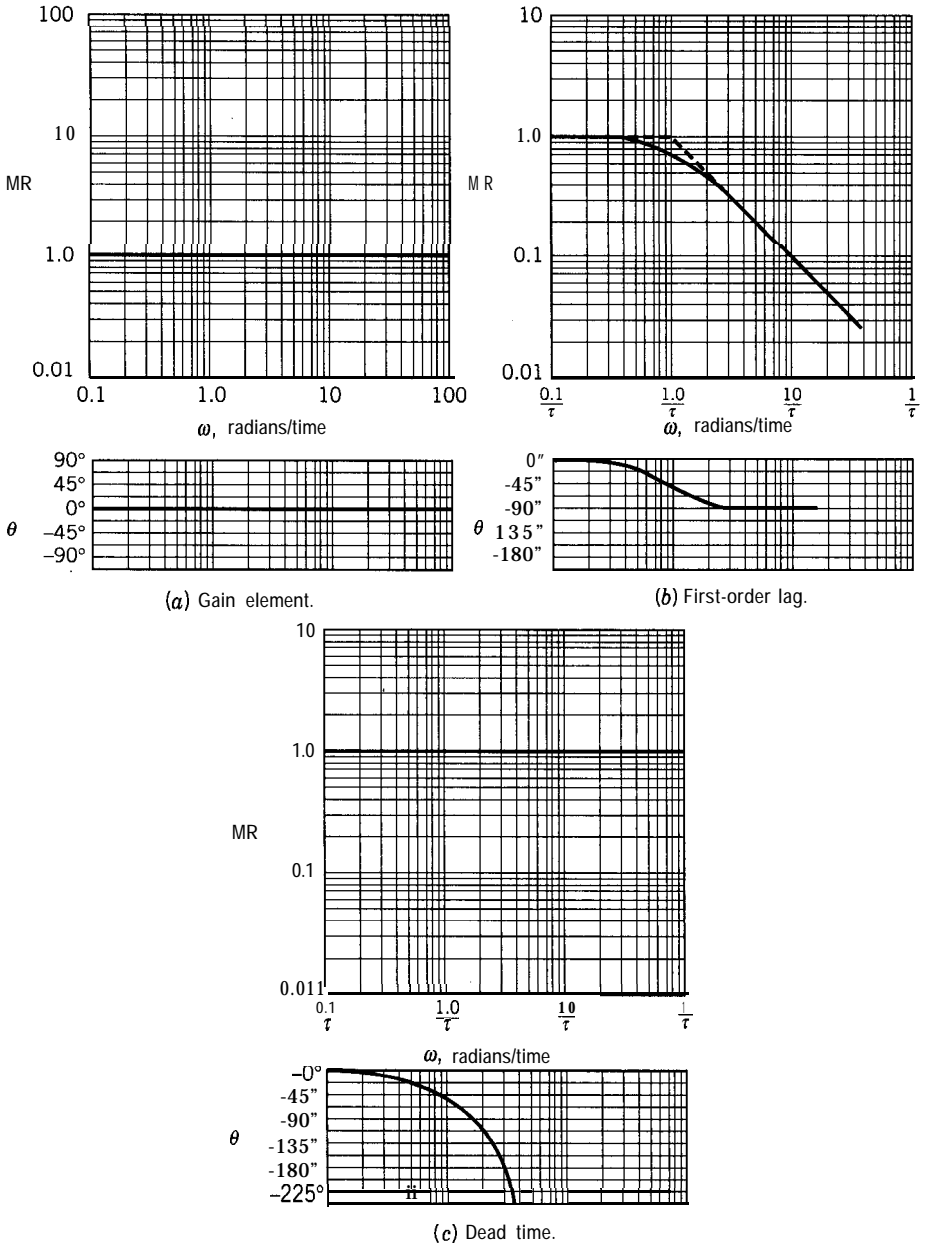
(c) Dead time.

Figure 9-1.3 Bode plots. (a) Gain element. (b) First-order lag. (c) Dead time.
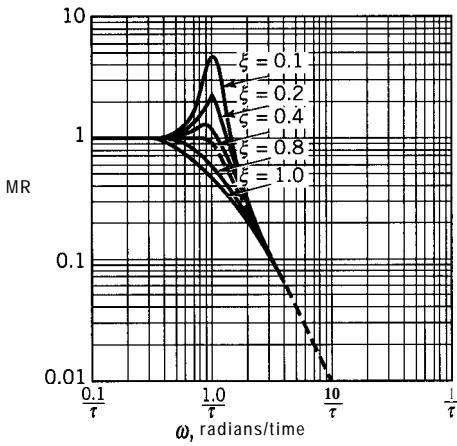
(d) Second-order lag.
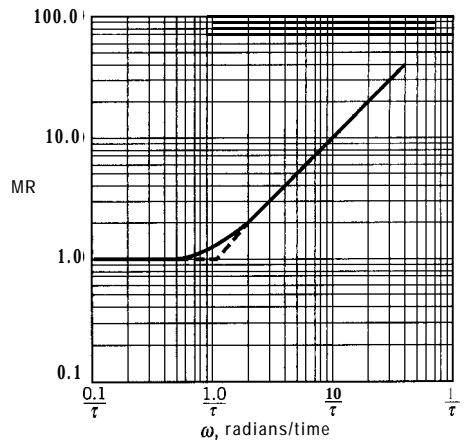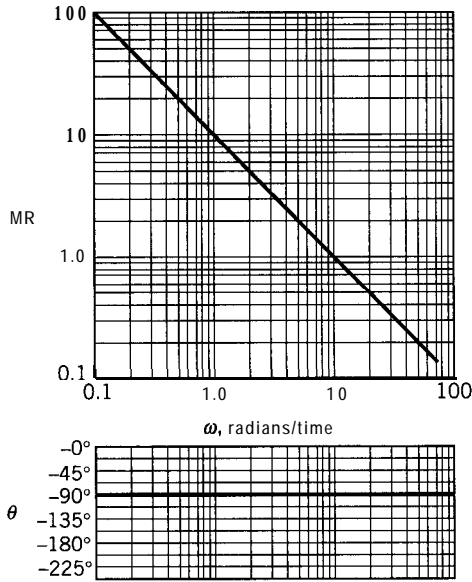
(e) First-order lead

(f) Integrator.

Figure 9-1.3 (Continued) (d) Second-order lag. (e) First-order lead. (f) Integrator.

which is the expression of a straight line in a log-log graph of MR vs. ω; this straight line has a slope of $-$ 1. The location of the line in the graph must now be determined. The simplest way to do this is to find where the high-frequency asymptote meets the low-frequency asymptote. It is known that as $\omega \rightarrow 0$, MR $\rightarrow$ 1, so

$$\log \text{MR} \rightarrow 0$$

Equating this equation to Eq. 9-1.26 yields

$$\omega = \frac{1}{\tau} \tag{9-1.27}$$

This is the frequency, referred to as the **corner frequency** (ω_c) or **breakpoint frequency,** at which the asymptotes meet, as shown in Fig. 9-1.3b. It is also at this frequency that the maximum error between the frequency response and the asymptotes exists. The actual magnitude ratio is

$$\text{MR} = \frac{1}{\sqrt{\omega^2 \tau^2 + 1}} = \frac{1}{\sqrt{2}} = 0.707$$

and not MR = 1 as given by the asymptotes.

Before leaving the Bode plot of this system, it is necessary to go over the θ at low and high frequencies. At low frequencies, $\omega \rightarrow 0$,

$$\theta \rightarrow - \text{ tar-'(wr)} = - \tan^{-1}(0) = 0$$

At high frequencies, $\omega \rightarrow \infty$,

$$\theta \rightarrow - \tan^{-1}(\infty) = -90"$$

These values of phase angle, 0" and $-$ 90°, are the asymptotes for the phase angle plot. At the corner frequency,

$$\omega_c = \frac{1}{\tau} \qquad \theta = - \tan^{-1}(1) = - 45°$$

To summarize, the important characteristics of the Bode plot of a first-order lag are the following:

1. **AR (MR) graph.** The low-frequency asymptote has a slope of 0, and the high-frequency asymptote has a slope of $-$ 1. The corner frequency, where these two asymptotes meet, occurs at $\frac{1}{\tau}$.
2. **Phase angle graph.** At low frequencies the phase angle approaches $0°$, whereas at high frequencies it approaches $-$ 90". At the comer frequency, the phase angle is $-$ 45".

Second-Order Lag. As shown in Example 9-1.1, the AR and θ expressions for a second-order lag are given by Eqs. 9-1.14 and 9-1.15, respectively. From the AR equation, we obtain

$$MR = \frac{1}{\sqrt{(1 - \omega^2\tau^2)^2 + (2\tau\zeta\omega)^2}}$$

and

$$\theta = -\tan^{-1}\left(\frac{2\tau\zeta\omega}{1 - \omega^2\tau^2}\right) \tag{9-1.15}$$

Giving values to ω for a given τ and ζ, the frequency response is determined as shown in Fig. 9-1.3*d*.

The asymptotes are obtained similarly to the first-order lag. At low frequencies, $\omega \rightarrow 0$,

$$MR \rightarrow 1$$

and

$$\theta \rightarrow 0$$

At high frequencies, $\omega \rightarrow \infty$,

$$\log MR \rightarrow \frac{1}{2}\log[(1 - \omega^2\tau^2)^2 + (2\tau\zeta\omega)^2]$$

$$\log MR \rightarrow -\frac{1}{2}\log(\omega^2\tau^2)^2 = -2\log\omega - 2\log\tau$$

which is the expression of a straight line with a slope of -2. At these high frequencies,

$$\theta \rightarrow -\pi \quad (-180°)$$

To find the comer frequency, ω_c, the same procedure as for the first-order lag is followed. It yields

$$\omega_c = \frac{1}{\tau}$$

Note from Fig. 9-1.3*d* that the transition of the frequency response from low to high frequencies depends on ζ.

At the comer frequency,

$$\theta = -\tan^{-1}(\infty) = -\frac{\pi}{2} \quad (- \quad 90°)$$

To summarize, the important characteristics of the Bode plot of a second-order lag are the following:

1. **AR (MR)** graph. The slope of the low-frequency asymptote is 0, and that of the high-frequency asymptote is -2. The comer frequency, ω_c, occurs at $\frac{1}{\tau}$. The transition of the AR from low to high frequency depends on the value of ζ.
2. **Phase angle graph.** At low frequencies the phase angle approaches $0°$, whereas at high frequencies it approaches -180". At the comer frequency, the phase angle is -90".

Dead Time. As shown in Example 9-1.3, the AR and θ expressions for a pure dead time are given by Eqs. 9-1.19 and 9-1.20, respectively.

$$AR = MR = 1 \tag{9-1.19}$$

and

$$\theta = -\omega t_0 \tag{9-1.20}$$

The Bode plot is shown in Fig. 9-1.3c. Note that as the frequency increases, the phase angle becomes more negative. The larger the value of the dead time, the faster the phase angle drops (becomes increasingly negative) without limit.

First-Order Lead. As shown in Example 9-1.2, the AR and θ expressions for a first-order lead are given by Eqs. 9-1.17 and 9-1.18, respectively. From the AR equation, we obtain

$$MR = \sqrt{1 + \omega^2 \tau^2}$$

and

$$\theta = \tan^{-1}(\omega\tau) \tag{9-1.18}$$

The Bode plot is shown in Fig. 9-1.3e. Note that the low-frequency asymptote has a slope of 0, and the high-frequency asymptote has a slope of $+1$. At low frequencies the phase angle approaches $0°$, whereas at high frequencies it approaches $+90$". At the comer frequency, the phase angle is $+4.5$". Thus a first-order lead provides "phase lead."

Integrator. As shown in Example 9-1.4, the AR and θ expressions for an integrator are given by Eqs. 9-1.21 and 9-1.22, respectively.

$$AR = MR = \frac{1}{\omega} \tag{9-1.21}$$

and

$$\theta = -90° \tag{9-1.22}$$

The Bode plot is shown in Fig. 9-1.3*f*. Note that the MR graph consists of a straight line with slope of -1. This is easily shown by taking the log of Eq. 9-1.21.

$$\log MR = -\log \omega$$

This equation also shows that MR = 1 at $\omega = 1$ radians/time.

Development of Bode Plot of Complex Systems. Most complex transfer functions of process systems are formed by the product of simpler components. Let us now look at the Bode plot of these complex transfer functions; consider Eq. 9-1.23 as an example. From the AR expression, Eq. (9-1.24), we obtain

$$\log \text{ AR} = \log K + \frac{1}{2}\sum_{i=1}^{m} \log[(\tau_i\omega)^2 + 1] - k \log \omega - \frac{1}{2}\sum_{j=1}^{n} \log[(\tau_j\omega)^2 + 1]$$

$$(9\text{-}1.28)$$

or

$$\log \text{ MR} = \frac{1}{2}\sum_{i=1}^{m} \log[(\tau_i\omega)^2 + 1] - k \log \omega - \frac{1}{2}\sum_{j=1}^{n} \log[(\tau_j\omega)^2 + 1] \quad (9\text{-}1.29)$$

These equations, together with Eq. 9-1.25, show that the Bode plot of complex systems consists of the sum of the individual components. To obtain the composite asymptote, we add the individual asymptotes.

Consider the transfer function

$$G(s) = \frac{K(s + 1)e^{-s}}{s(2s + 1)(3s + 1)}$$

Using principles we have learned, we know that

$$MR = \frac{\sqrt{\omega^2 + 1}}{\omega\sqrt{4\omega^2 + 1}\sqrt{9\omega^2 + 1}}$$

or

$$\log \text{ MR} = \frac{1}{2}\log(\omega^2 + 1) - \log(w) - \frac{1}{2}\log(4\omega^2 + 1) - \frac{1}{2}\log(9\omega^2 + 1)$$

and

$$\theta = \tan^{-1}(\omega) - \omega t_0 \quad \pi \quad \tan^{-1}(2\omega) \quad \tan^{-1}(3\omega)$$
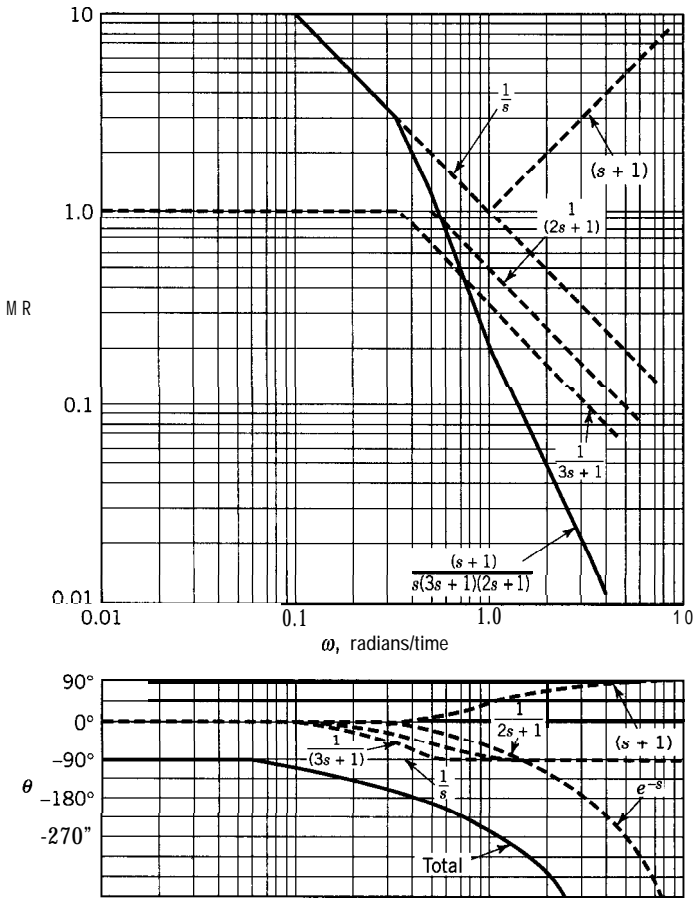
Figure 9-1.4 Bode plot of $G(s) = \dfrac{K(s + 1)e^{-s}}{s(2s + 1)(3s + 1)}$,

From these last two equations, the Bode plot is developed as shown in Fig. 9-1.4. At low frequencies, $\omega < 0.33$, the slope is $- 1$ because of the integrator term. At $\omega = 0.33$, one of the first-order lags starts to contribute to the graph, and thus the slope changes to $- 2$ at this frequency. At $\omega = 0.5$, the other first-order lag starts to contribute, changing the slope of the asymptote to $- 3$. Finally, at $\omega = 1$, the first-order lead enters with a slope of $+ 1$, and the slope of the asymptote changes back to $- 2$. Similarly, the composite phase angle plot is obtained by algebraically adding the individual angles.

One final comment can be made about the slopes of the low-frequency and **high-**frequency asymptotes (initial and final slopes) and angles of Bode plots. Consider a

general transfer function such as

$$G(s) = \frac{K(a_m s^m + a_{m-1}s^{m-1} + \ldots + 1)}{s^k(b_n s^n + b_{n-1}s^{n-1} + \ldots + 1)}, \qquad (n+k) > m$$

The slope of the low-frequency asymptote is given by

$$\text{Slope of AR } (MR)|_{\omega \to 0} \longrightarrow (-1)k$$

and the angle

$$\theta|_{\omega \to 0} \to (-90°)k$$

The slope of the high-frequency asymptote is given by

$$\text{Slope of AR } (MR)|_{\omega \to \infty} \to (n+k \quad m)(-1)$$

and the angle

$$\theta|_{\omega \to \infty} \longrightarrow (n+k-m)(-90")$$

Most systems follow these slopes and angles; these systems are called *minimalphase systems*. There are three exceptions, however, which are called *nonminimal phase systems*. These exceptions are

1. Systems with dead time: $G(s) = e^{-t_0 s}$ (phase angle decreases without limit).

2. Systems that show inverse response (positive zeros): $G(s) = \dfrac{(1 - \tau_1 s)}{(1 + \tau_2 s)}$. A process example of this type of response is given in Chapter 4.

3. Systems that are open-loop unstable (positive poles): $G(s) = \dfrac{1}{(\tau s - 1)}$, A process example of this type of response is given in Chapter 4.

In each of these cases, the magnitude ratio plot is not changed, but the phase angle plot is. The dead-time term was presented earlier; the Bode plot of the other two systems is the subject of one of the problems at the end of this chapter.

The expression for the slope of the high-frequency asymptote also serves to show why transfer functions of real systems must have at least as many lags as leads. If $(n + k - m) < 0$, then the final slope is positive and noise of high frequency is amplified with infinite gain.

Example 9-1.5 has shown how to obtain the Bode plot using mainly the asymptotes. For a more precise graph, the expressions for AR, or MR, and θ are used. There are several software programs, such as MATLAB (see the References section in this chapter), that provide the Bode plot given the transfer function.

9-2 FREQUENCY RESPONSE STABILITY CRITERION

Section 6-2 presented two ways to determine the limits of stability of a feedback control loop: direct substitution and Routh's test. However, as we saw there, neither of these
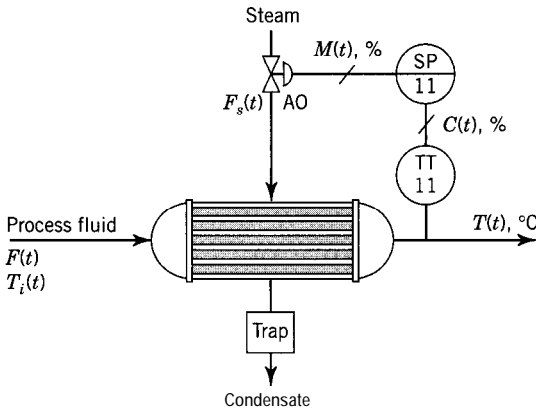
Figure 9-2.1 Feedback control loop for temperature control of heat exchanger.

methods can handle the presence of dead time in the loop except through approximation. The frequency response stability criterion presented here can determine the stability limits for feedback loops even when there is dead time in the loop. The criterion consists of determining the frequency at which the phase angle of the open-loop transfer function (OLTF) is 180° ($-\pi$ radians) and the amplitude ratio of the OLTF at that frequency.

Consider the heat exchanger temperature control loop first presented in Example 6-2.1, and then again in Example 8-3.1. For convenience, the exchanger is shown again in Fig. 9-2.1 and its block diagram in Fig. 9-2.2. The open-loop transfer function is

$$\text{OLTF} = \frac{0.8K_c}{(10s + 1)(30s + 1)(3s + 1)}$$

The MR and θ expressions are

$$\text{MR} = \frac{\text{AR}}{0.8K_c} = \frac{1}{\sqrt{(10\omega)^2 + 1}\sqrt{(30\omega)^2 + 1}\sqrt{(3\omega)^2 + 1}} \qquad (9\text{-}2.1)$$

$$\theta = -\tan^{-1}(10\omega) - \tan^{-1}(30\omega) - \tan^{-1}(3\omega) \qquad (9\text{-}2.2)$$
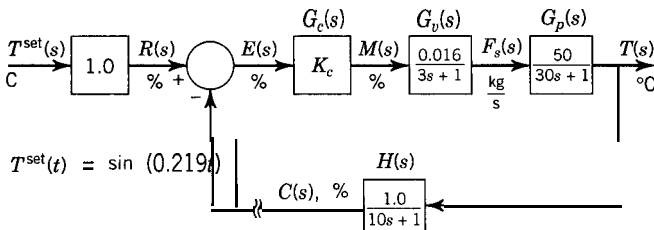
Figure 9-2.2 Block diagram of heat exchanger temperature control loop-P controller.
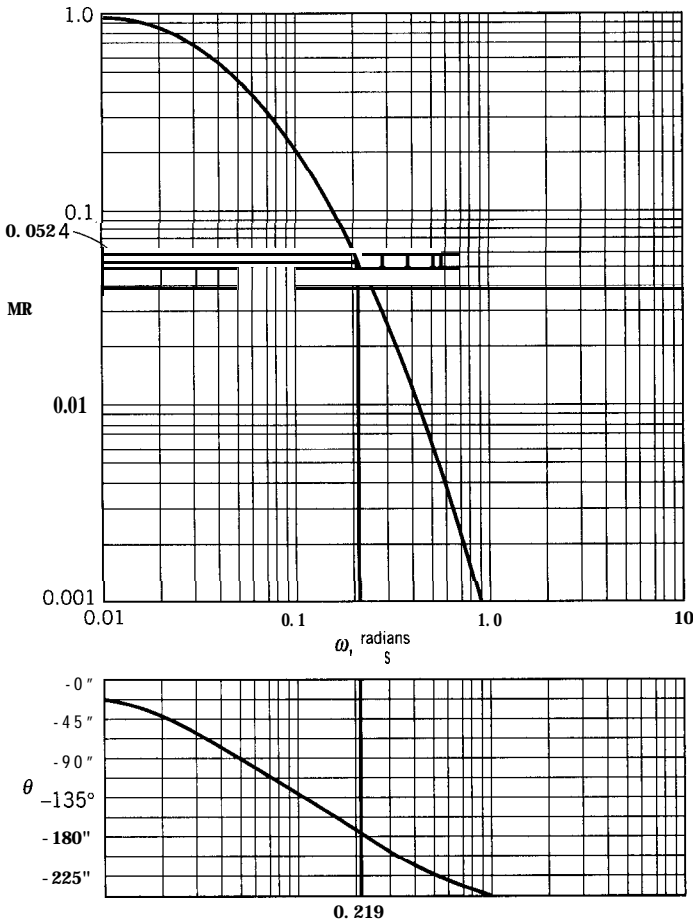
Figure 9-2.3 Bode plot of heat exchanger temperature control loop-P controller.

The Bode plot is shown in Fig. 9-2.3. From this figure, or from Eq. 9-2.2, the frequency at which $\theta = -180$" (or $-\pi$ radians) is 0.219 rad/s. At this frequency, from the Bode plot or from Eq. 9-2.1,

$$\frac{\text{AR}}{0.8K_c} = 0.0524$$

The controller gain that yields AR $= 1$ is

$$K_c = \frac{\text{AR}}{0.8(0.0524)} = \frac{1}{0.8(0.0524)} = 23.8 \frac{\%\text{CO}}{\%\text{TO}}$$

These calculations are highly significant. The value $K_c = 23.8$ %CO/%TO is the gain of the controller that yields AR $= 1$ when the phase angle is -180". Remember that AR is defined as the ratio of the amplitude of the output signal to the amplitude of the

input signal, Y_0/X_0. This means that if the input set point to the temperature controller is varied as follows

$$T^{\text{set}}(t) = \sin(0.219t)$$

then the output signal from the transmitter, after the transients disappear, will vary as

$$C(t) = \sin(0.219t - \pi) = -\sin(0.219t)$$

Note that the feedback signal is disconnected from the controller, as shown in the block diagram of Fig. 9-2.2, and that the frequency of the set point oscillation is 0.219 rad/s. This is the frequency at which $\theta = -180" = -\pi$ radians and, when the controller gain is 23.8, AR = 1. Under these conditions, the amplitude of the output signal is equal to that of the input signal.

Suppose now that at some time, $t = 0$, the set point oscillations are stopped, $T^{\text{set}}(t) = 0$, and at the same time the transmitter signal is connected to the controller. The error signal, E(s), inside the controller remains unchanged, and the oscillations are sustained. If nothing changes in the control loop, the oscillations remain indefinitely. If at some time the controller gain is slightly increased to 25, the amplitude ratio becomes 1.04.

$$AR = 0.0524(0.8)K_c$$
$$= 0.0524(0.8)(25) = 1.05$$

This means that as the signal goes through the control loop, it is amplified. After the first time, the output signal from the transmitter is $-1.05 \sin(0.219t)$. After the second time, it is $-(1.05)^2 \sin(0.219t)$, and so on. If this is not stopped, the outlet temperature will increase continuously, yielding an unstable control loop.

On the other hand, if the controller gain is slightly decreased to 23, the amplitude ratio becomes 0.96.

$$AR = 0.0524(0.8)23 = 0.96$$

This means that as the signal goes through the control loop, it decreases in amplitude. After the first time, the output signal from the transmitter is $-0.96 \sin(0.219t)$. After the second time, it is $-(0.96)^2 \sin(0.219t)$, and so on. This results in a stable control loop.

In summary, the stability criterion based on frequency response can be stated as follows:

For a control system to be stable, the amplitude ratio must be less than unity when the phase angle is $-180"$ ($-\pi$ radians). If AR < I at $\theta = -180°$, the system is stable; if AR > 1 at $\theta = -180°$, the system is unstable.

The controller gain that provides the condition of AR = 1 at $\theta = -180"$ is the

ultimate gain, K_{cu}. The frequency at which this condition happens is the ultimate frequency, ω_u. From this frequency, the ultimate period can be calculated as $T_u = 2\pi/\omega_u$. Note that the values of ultimate frequency and gain obtained in this example are the same as those obtained in Example 6-2.1 by direct substitution.

Before proceeding with more examples, it is important to stress that the ultimate frequency and ultimate gain can be obtained directly from the MR and θ equations, Eqs. 9-2.1 and 9-2.2 for this example, without the need for the Bode plot. The Bode plot was developed from these equations. Using these equations saves drawing the plot. Many years ago when hand-held calculators were not available (remember slide rules?), it was probably easier to draw the Bode plot, using the high- and low-frequency asymptotes. The use of calculators and computers makes the determination of ω_u and K_{cu} a rather easy procedure. The determination of ω_u requires a small amount of trial and error using the θ equation-that is, finding which ω yields $\theta = -180$". This ω is ω_u. Once ω_u has been determined, the equation for AR is used to calculate K_{cu}. This complete procedure is usually faster and yields more accurate results than drawing and using the Bode plot. Of course, the Bode plot is still very useful because it shows at a glance how AR and θ vary as the frequency varies.

The following examples will provide more practice with this powerful technique.

EXAMPLE 9-2.1

Consider the same heat exchanger (Fig. 9-2.1) previously used to explain the frequency response stability criterion. Suppose now that for some reason, the outlet temperature cannot be measured at the exit of the exchanger but must be measured farther down the pipe. The effect of this new sensor location is the addition of some dead time, say 2 s, to the control loop.

The new OLTF is

$$\text{OLTF} = \frac{0.8 K_c e^{-2s}}{(10s + 1)(30s + 1)(3s + 1)}$$

with

$$\text{MR} = \frac{\text{AR}}{0.8 K_c} = \frac{1}{\sqrt{(10\omega)^2 + 1}\sqrt{(30\omega)^2 + 1}\sqrt{(3\omega)^2 + 1}}$$

and

$$\theta = -2\omega - \tan^{-1}(10\omega) - \tan^{-1}(30\omega) - \tan^{-1}(3\omega)$$

The last two equations can be used to determine the ultimate frequency and ultimate gain. Performing these calculations yields, for $\theta = -\pi$ radians,
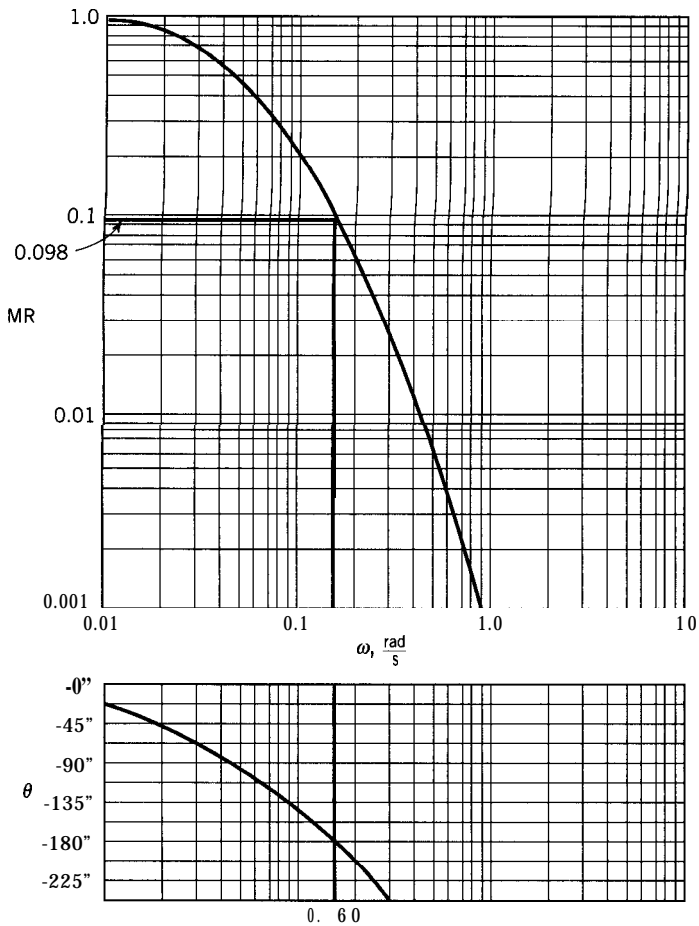
$$\omega_u = 0.160 \ \text{rad/s}$$

Figure 9-2.4 Bode plot of heat e: changer temperature control with dead time-P controller.

and

$$K_{cu} = 12.8$$

The Bode plot is shown in Fig. 9-2.4.

The results of Example 9-2.1 show the effect of dead time on the stability (and consequently on the controllability) of the control loop. The ultimate gain and ultimate period for the heat exchanger without dead time were previously found to be

$$K_{cu} = 23.8 \ \frac{\%CO}{\%TO} \quad \text{and} \quad \omega_u = 0.219 \ \text{rad/s}$$

When the dead time of 2 s was added in Example 9-2.1, the results were

$$K_{cu} = 12.8 \ \frac{\%CO}{\%TO} \quad \text{and} \quad \omega_u = 0.160 \ \text{rad/s}$$

Thus it is easier for the process with dead time to go unstable. The difference in ω_u also indicates that the closed-loop response of a process with dead time is slower than that of a process without dead time.

The preceding example demonstrates that the frequency response stability criterion can exactly analyze the effect of dead time in the control loop. As mentioned in previous chapters, dead time is the worst thing that can happen to any control loop; this example proves this point. The dead-time term "adds phase lag" to the control loop, so the phase angle crosses the 180" value at a lower frequency. The longer the dead time, the lower the ultimate frequency and ultimate gain.

Example 8-3.2 showed that the addition of reset mode to a proportional controller decreases the ultimate frequency and ultimate gain. This can be explained, from a frequency response point of view, by saying that the addition of reset mode "adds phase lag" to the control loop. A proportional-only controller has a phase angle of $0°$, as shown in Fig. 9-1.3a. Consider now a proportional-integral controller:

$$G(s) = K_c \left(1 + \frac{1}{\tau_I s} \right) = K_c \left(\frac{\tau_I s + 1}{\tau_I s} \right)$$

This transfer function is composed of a lead term, $\tau_I s + 1$, and an integrator term, $1/\tau_I s$. At low frequencies, $\omega \ll \frac{1}{\tau_I}$, the lead term does not affect the phase angle, but the integrator term contributes $-90°$, thus adding phase lag. At higher frequencies, $\omega \gg \frac{1}{\tau_I}$, the lead term cancels the integrator term with a resulting 0" phase angle. However, unless the reset time is very long, this canceling effect occurs at a frequency higher than that at which the phase angle crosses the -180" value. Thus the loop with the PI controller will have a lower ultimate frequency and gain than that with the P controller.

Remember that the reset mode in a controller is the one that removes the offset. However, as explained in the previous paragraph, its use adds phase lag to the loop. It seems that an analog of the second law of thermodynamics applies to process control: You cannot get something for nothing.

The following example demonstrates the effect of derivative mode on the stability of a control loop.

EXAMPLE 9-2.2

Consider the same heat exchanger control loop, no dead time, with a proportional derivative controller. Suppose the rate time is 0.25 min (15 s). The equation for a "real"

PD controller, as shown in Chapter 5, is

$$G_c(s) = K_c \left(\frac{1 + \tau_D s}{1 + \alpha \tau_D s} \right)$$

or, for this example, using $\alpha = 0.1$,

$$G_c(s) = K_c \left(\frac{1 + 15s}{1 + 1.5s} \right)$$

The OLTF is then

$$\text{OLTF} = \frac{0.8 K_c (1 + 15s)}{(10s + 1)(30s + 1)(3s + 1)(1 + 1.5s)}$$

with

$$\text{MR} = \frac{\text{AR}}{0.8 K_c} = \frac{\sqrt{(15\omega)^2 + 1}}{\sqrt{(10\omega)^2 + 1}\sqrt{(30\omega)^2 + 1}\sqrt{(3\omega)^2 + 1}\sqrt{(1.5\omega)^2 + 1}}$$

and

$$\theta = \tan^{-1}(15\omega) - \tan^{-1}(10\omega) - \tan^{-1}(30\omega) - \tan^{-1}(3\omega) - \tan^{-1}(1.5\omega)$$

The Bode plot for this system is shown in Fig. 9-2.5. Comparing this Bode plot to the one shown in Fig. 9-2.3, we see that the phase angle plot has been "moved up"; the derivative action "adds phase lead." In this system, the ultimate gain and period are found to be

$$K_{cu} = 33 \ \frac{\%\text{CO}}{\%\,\text{TO}} \qquad \text{and} \qquad \omega_u = 0.53 \ \text{rad/s}$$

Thus these results show that the derivative mode makes the control loop more stable and faster.

In the preceding example and discussion, we used the terms *ultimate gain* and *ultimate frequency* for controllers other than proportional controllers. However, the ultimate gain and period used for tuning are still defined only for proportional controllers.

The examples presented in this section have demonstrated the use of frequency response, and in particular Bode plots, for analysis of control loops. These examples have also shown the effect of different terms, dead time, and derivative mode on the stability of control loops.

The frequency response stability criterion confirms that control loops with a pure (no dead time) first- or second-order open-loop transfer function will never go unstable; their phase angles will never go below − 180". Once a dead time is added, no matter how small, the system can go unstable because the phase angle will always cross the − 180" value.
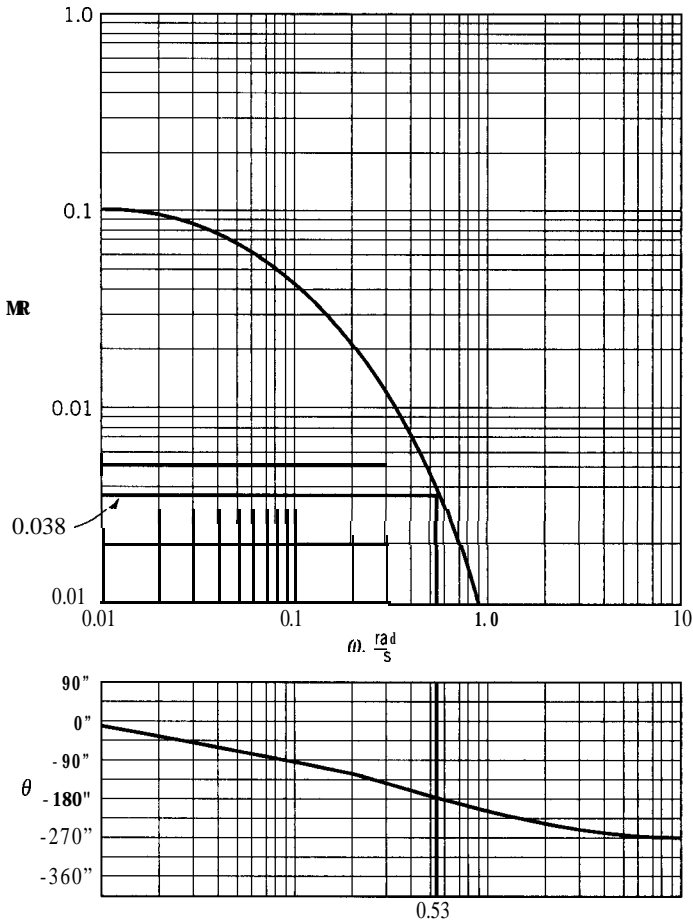
Figure 9-2.5 Bode plot of heat exchanger temperature control loop—PD controller.

Controller Performance Specifications

Chapter 7 presented several ways to tune controllers to obtain a desired loop perform-ance. The methods presented were the Ziegler-Nichols (quarter decay ratio) equations, the error-integral criteria (IAE, ISE, and ITAE), and controller synthesis. Frequency response provides a procedure by which to obtain the ultimate gain and ultimate fre-quency of a control loop. Once these terms have been determined, the Ziegler-Nichols equations can be used to tune the controller. Frequency response techniques provide still other performance specifications for tuning controllers. There are three such meth-ods: gain margin, phase margin, and closed-loop response. The gain margin and the phase margin are both based on the frequency response of the open-loop transfer func-tion (OLTF), whereas the closed-loop response is based on the frequency response of the closed-loop transfer function.

Gain Margin. The *gain margin* (GM) is a typical performance specification associated with the frequency response technique. The gain margin represents the factor by which

the total loop gain must increase to make the system just unstable. The controller gain that yields a desired gain margin is calculated as follows:

$$K_c = \frac{K_{cu}}{\text{GM}} = \frac{1}{K(\text{GM})\text{MR}|_{\theta = -180°}}$$

(9-2.3)

where **K** is the product of the gains of all other elements in the loop. A typical specification is GM $\geq$ 2. Note that the tuning of a proportional controller with GM = 2 is the same as the Ziegler-Nichols quarter decay ratio tuning presented in Chapter 7.

Phase Margin. Phase margin (PM) is another specification commonly associated with the frequency response technique. The phase margin is the difference between $-$ 180 and the phase angle at the frequency for which the amplitude ratio (AR) is unity. That is,

$$\text{PM} = 180" + \theta|_{\text{AR}=1}$$

(9-2.4)

PM represents the additional amount of phase lag required to make the system unstable. A typical specification is PM $>$ 45".

Closed-Loop Response. The gain and phase margin specifications are based on the open-loop transfer function (OLTF). Furthermore, all the Bode plots developed so far in this chapter have been based on the OLTF. The Bode plot could also be developed from a closed-loop transfer function. The *closed-loop response* (CLR) is another performance specification associated with the frequency response techniques and based on the closed-loop transfer function.

Consider the block diagram shown in Fig. 9-2.2. The set point closed-loop transfer function is

$$\frac{T(s)}{T^{\text{set}}(s)} = \frac{0.8K_c}{900s^3 + 420s^2 + 43s + (1 + 0.8K_c)}$$

(9-2.5)

The closed-loop Bode plot is shown in Fig. 9-2.6 for different values of K_c. Most significantly, the figure shows that as K_c increases, the corner frequency moves to the right, and the height of the peak also increases. As we may recall, for an OLTF the comer frequency, ω_c, is the reciprocal of the time constant, as given by Eq. 9-1.27. The same holds true for a closed-loop transfer function; that is, the closed-loop comer frequency, $\omega_{c_{CL}}$, is the reciprocal of the closed-loop time constant, τ_{CL}. Therefore, the movement of $\omega_{c_{CL}}$ to the right is desirable because the larger $\omega_{c_{CL}}$ is, the shorter τ_{CL} is, and consequently, the faster-responding the closed-loop process becomes. However, as K_c increases, the peak height also increases. This peak, also referred to as the **resonant peak,** is related to the closed-loop damping ratio, ζ_{CL}. The higher the peak, the smaller ζ_{CL} is, and consequently, the more underdamped (oscillatory) the closed-loop response becomes. This ζ_{CL} is equivalent to the open-loop ζ of a second-order system, and its effect is shown in Fig. 9-1.3d. Thus, as K_c increases, two common conflicting results develop: a faster but more oscillatory controlled process.
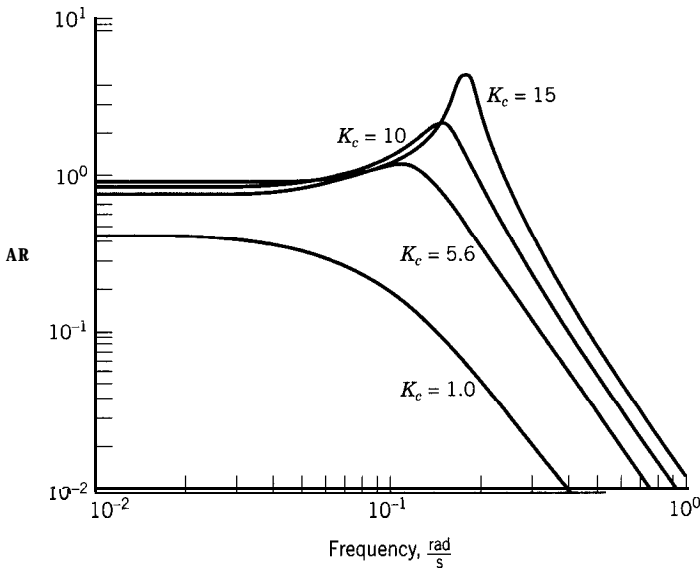
Figure 9-2.6 Bode plot of Eq. 9-2.5 for different values of K_c.

A common CLR specification is to tune the controller to obtain a *maximum peak height* (MPH) of 1.26. (Books that graph 20 log AR instead of AR specify a maximum peak height of + 2 db and refer to this specification as the *maximum log modulus.*) The MPH of 1.26 provides a ζ_{CL} of about 0.4.

Example 9-2.3 illustrates the use of these last three specifications to tune a controller.

EXAMPLE 9-2.3

Consider the heat exchanger of Example 9-2.1. Tune a proportional controller for (a) GM = 2, (b) PM = $45°$, and (c) MPH = 1.26.

SOLUTION

(a) In Example 9-2.1, the ultimate gain of the controller was determined to be $K_c =$ 12.8. To obtain GM specification of 2, the controller gain is then set to

$$K_c \bigg|_{GM=2} = \frac{K_{cu}}{2} = 6.4 \; \frac{\%CO}{\%TO}$$

(b) In Example 9-2.1, expressions for MR and θ were determined to be

$$MR = \frac{AR}{0.8K_c} = \frac{1}{\sqrt{(10\omega)^2 + 1} \; \sqrt{(30\omega)^2 + 1} \sqrt{(3\omega)^2 + 1}}$$

and

$$\theta = - (2\omega t_0) - \tan^{-1}(10\omega) - \tan^{-1}(30\omega) - \tan^{-1}(3\omega)$$

On the basis of the definition of phase margin, for PM = $45°$, $\theta = -135"$. Using the equation for θ, or the Bode plot of Fig. 9-2.4, we find that the frequency for this phase angle is

$$\omega|_{\text{PM}=45°} = 0.087 \text{ rad/s}$$

Then, substituting into the equation for the magnitude ratio, we get

$$\frac{\text{AR}}{0.8K_c} = 0.261$$

$$K_c|_{\text{PM}=45°} = \frac{\text{AR}}{0.8(0.261)} = \frac{1}{0.8(0.261)} = 4.8 \frac{\%\text{CO}}{\%\text{TO}}$$

(c) Figure 9-2.6 shows that when $K_c = 5.6 \dfrac{\%\text{CO}}{\%\text{TO}}$, the peak on the AR graph reaches 1.26.

Example 9-2.3 has shown how to obtain the tuning of the feedback controller for a certain GM, PM, and CLR. In part (a) the controller was tuned to yield a control loop with a GM of 2. This means that the overall loop gain must increase (because of process nonlinearities, or for any other reason) by a factor of 2 before instability is reached. In choosing the value of GM, the engineer must understand the process to decide how much the process gain can change over the operating range. On the basis of this understanding, the engineer can choose a realistic GM value. The larger the GM value chosen, the greater the "safety factor" designed into the control system. However, the larger this safety factor is, the smaller the controller gain that results and, therefore, the less sensitive the controller is to errors.

In part (b) of the example, the controller was tuned to yield a PM of 45". This means that 45" of phase lag must be added to the control loop before it goes unstable. Changes in phase angle of the control loop are due mainly to changes in its dynamic terms, time constants, and dead time, because of process nonlinearities.

Gain margin and phase margin are two different performance criteria. The choice of one of them as the criterion for a particular loop depends on the process being controlled. If, because of process nonlinearities and characteristics, the gain is expected to change more than the dynamic terms, then the GM may be the indicated criterion. If, on the other hand, the dynamic terms are expected to change more than the gain, then the PM may be the indicated criterion.

In part (c) the controller was tuned to yield a MPH of 1.26. A value of $K_c = 5.6 \dfrac{\%\text{CO}}{\%\text{TO}}$ was obtained. It is instructive to calculate the gain margin obtained with this MPH specification. A $K_c = 5.6 \dfrac{\%\text{CO}}{\%\text{TO}}$ yields a gain margin of $12.8/5.6 = 2.28$. Therefore, the MPH specification yields a more conservative tuning than the one provided with GM = 2.

Example 9-2.3 demonstrated how the gain of a proportional-only controller is calculated to yield the desired performance specification. If a PI or PID controller is used,

the reset and rate times must be set before K_c can be calculated. This means that more than one set of tuning parameters yields the desired performance. It is up to the engineer to choose what he or she considers "best."

In this section we discussed the meaning of gain margin, phase margin, and CLR, as well as how to tune feedback controllers on the basis of these performance specifications. In the process industries, however, the performance specifications of Chapter 7 are almost exclusively preferred.

9-3 POLAR PLOTS

The polar plot is another common way to graph the frequency response of control systems. It has the advantage of being only one graph as opposed to the two graphs of Bode plots. The polar plot is a graph of the complex-valued function $G(i\omega)$ as ω goes from 0 to ∞. For every value of ω, there will be a vector in the complex plane. The end of this vector will generate a locus as ω changes. The vector has its base at the origin and has length equal to the amplitude ratio of the $G(i\omega)$ function; its angle with the positive real axis is the phase angle. This section presents the fundamentals of polar plots and explains how to graph them. The polar plots of some of the most common process components are presented first.

First-Order Lug. The amplitude ratio and the phase angle of a first-order lag are given by Eqs. 9-1.12 and 9-1.13, respectively.

$$AR = \frac{K}{\sqrt{(\omega\tau)^2 + 1}} \qquad (9\text{-}1.12)$$

$$\theta = -\tan^{-1}(\omega\tau) \qquad (9\text{-}1.13)$$

For $\omega = 0$, $AR = K$ and $\theta = 0°$. For $\omega = \dfrac{1}{\tau}$, $AR = 0.707K$ and $\theta = -45°$. For $\omega = \infty$, $AR = 0$ and $\theta = -90"$. The polar plot for this system is shown in Fig. 9-3.1. The solid curve represents the amplitude ratio and phase angle as the frequency goes
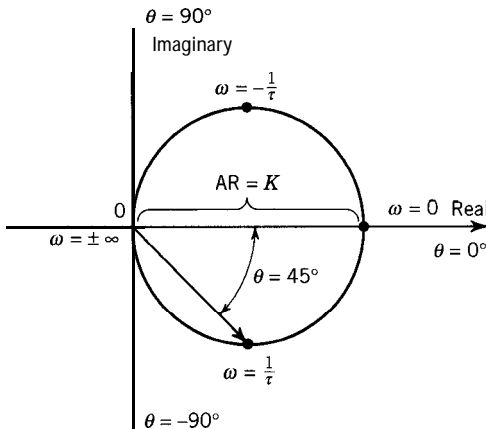
Figure 9-3.1 Polar plot of first-order lag.

from 0 to ∞. Each point in the curve represents a different ω. The length of the vector from the origin to a point on the curve is equal to the amplitude ratio at that ω. The angle that the vector makes with the positive real axis is equal to the phase angle. Figure 9-3.1 shows two vectors. The first vector represents AR and θ for $\omega = 0$. The second vector represents AR and θ for $\omega = \frac{1}{\tau}$. Note that as AR approaches zero, θ approaches -90"; this is what Eqs. 9-1.12 and 9-1.13 indicate. The dotted curve is the plot of AR and θ as ω goes from $-\infty$ to 0.

Second-Order Lag. The AR and θ equations for a second-order lag are given by Eqs. 9-1.14 and 9-1.15, respectively.

$$AR = \frac{K}{\sqrt{(1 - \omega^2 \tau^2)^2 + (2\tau\zeta\omega)^2}} \qquad (9\text{-}1.14)$$

$$\theta = -\tan^{-1}\left(\frac{2\tau\zeta\omega}{1 - \omega^2\tau^2}\right) \qquad (9\text{-}1.15)$$

For $\omega = 0$, AR=K and $\theta = 0°$. For $\omega = 1/\tau$, AR $= K/2\zeta$ and $\theta = -90$". For $\omega = \infty$, AR $= 0$ and $\theta = -180$". Figure 9-3.2 shows the polar plot for this system. In this system AR approaches zero from the negative real axis because θ approaches 180".

Dead Time. The AR and θ expressions for a pure dead-time system are given by Eqs. 9-1.19 and 9- 1.20, respectively.

$$AR = 1 \qquad (9\text{-}1.19)$$

$$\theta = -t_0\omega \qquad (9\text{-}1.20)$$

These equations indicate that the vector will always have a magnitude of unity and
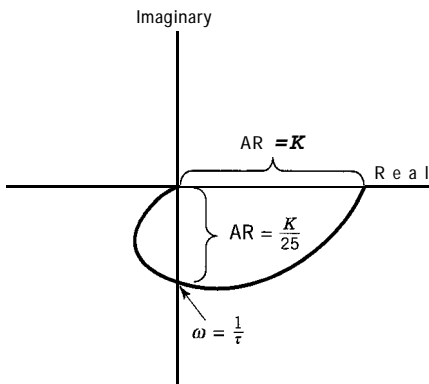
Figure 9-3.2 Polar plot of second-order system.

Imaginary

Real
$\theta = 0,\ 360''$

$\theta = 270°$
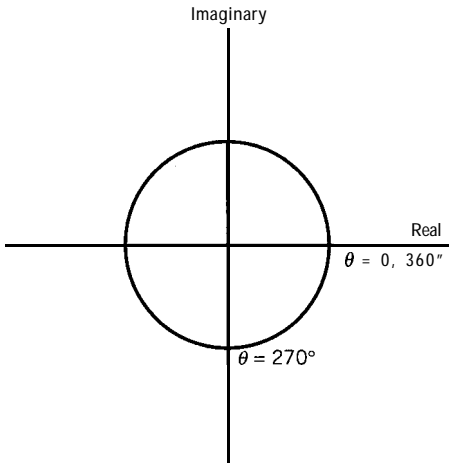
Figure 9-3.3 Polar plot of dead time.

that as ω increases, the vector will start rotating. The resulting polar plot, shown in Fig. 9-3.3. is a unit circle.

Conformal Mapping. Some examples of polar plots have been shown. However, before we continue with this subject, it is important to talk about conformal mapping, **because** the polar plots rely heavily on this theory. The following brief introduction to conformal mapping will make it easier to understand polar plots.

Consider the general transfer function G(s). As we know, the variable s is the independent variable, which can be real, imaginary, or complex; that is, in general, $s = \sigma + i\omega$. This variable s can be graphed in the s-plane, as shown in Fig. 9-3.4a. Substituting the value of s into the transfer function G(s), we can obtain the value for this function at the given s. This value of G(s), which can also be real, imaginary, or complex, G(s) = $\delta + i\gamma$, can be graphed in the G(s)-plane, as shown in Fig. 9-3.4b.

Imaginary

s-plane

ω - X

$s = \sigma + \omega i$

Real

σ

(a)

Imaginary

G(s)-plane

X

γ

$G(s) = \delta + \gamma i$

Real

δ

(b)
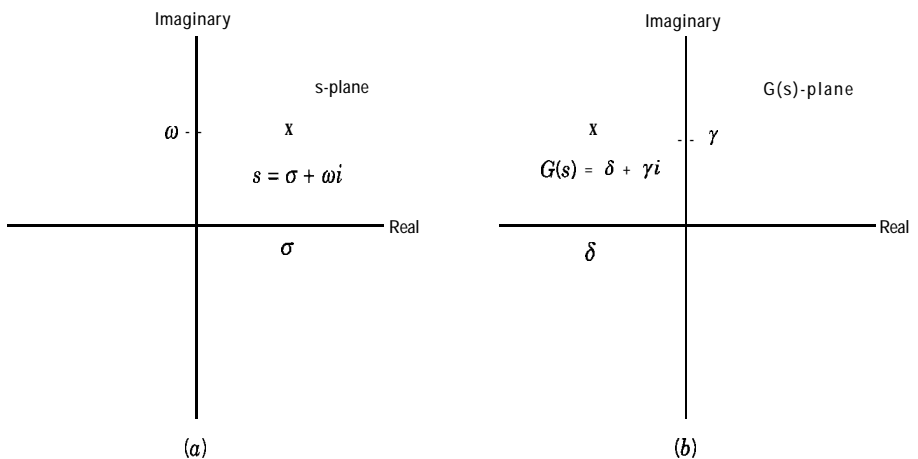
Figure 9-3.4 Mapping of a point in the s-plane onto the G(s)-plane.

For every point in the s-plane, there is a corresponding point in the G(s)-plane. It is said that the function G "maps" the s-plane onto the G(s)-plane. The function G maps not only points but also paths or regions.

The word *conformal* is used because, in mapping, the G(s)-plane "conforms" with the s-plane. To explain what is meant by this, suppose that a path from the s-plane must be mapped onto the G(s)-plane. Further, if the path in the s-plane has a sharp turn, then the mapped path in the G(s)-plane will also have a sharp turn. That is, the G(s)-plane "conforms" with the s-plane.

To be a bit more specific about this conformal mapping, consider the transfer function

$$\mathbf{G(s)} = \frac{10}{(2s + 1)(4s + 1)}$$

Figure 9-3.5 shows the mapping of a region in the s-plane, given by points 1-2-3-4-1, onto the G(s)-plane, given by points $1'$-$2'$-$3'$-$4'$- 1'. Table 9-3.1 shows the mathematical manipulations required to produce the mapping.

This has been a brief explanation of conformal mapping. As we noted earlier, polar plots are based on this theory. Conformal mapping will also be important when the use of polar plots to study process control stability is presented.
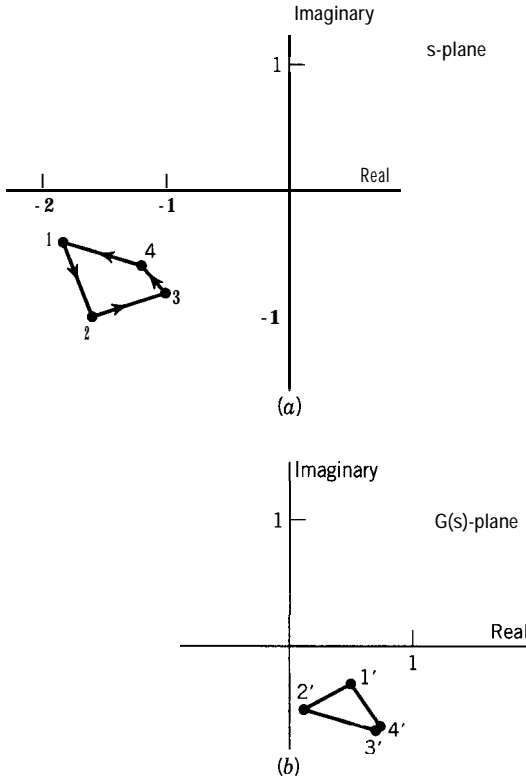
Figure 9-3.5 Mapping from the s-plane onto the G(s)-plane.

Table 9-3.1

Point	Coordinate	2s + 1	4s + 1	(2s + 1)(4s + 1)	G(s)	Point in G(s) Plane
1	$-1.8 - 0.4i$	$-2.6 \quad 0.8i$	$-6.2 - 1.6i$	$14.84 + 9.12i$	$\dfrac{10}{(14.84 + 9.12i)} = 0.49 - 0.3i$	1'
2	$-1.6i$	$-2.2 - 2i$	$-5.4 - 4i$	$3.88 + 19.6i$	$\dfrac{10}{3.88 + 19.6i} = 0.10 - 0.49i$	2'
3	$-1 - 0.8i$	$-1 - 1.6i$	$-3 \quad 3.2i$	$8.12 + 8i$	$\dfrac{10}{8.12 + 8i} = 0.62 \quad 0.61i$	3'
4	$-1.2 - 0.6i$	$-1.4 - 1.2i$	$-3.8 - 2.4i$	$8.2 + 7.921'$	$\dfrac{10}{8.2 + 7.921'} = 0.63 - 0.61i$	4'

The Nyquist Stability Criterion. Polar plots are useful in the stability analysis of process control loops. This section presents the Nyquist stability criterion (Nyquist, 1932), which makes use of these plots. No proof of the theorem is given; the reader is strongly encouraged to read the original paper. Because this criterion is based on the use of polar plots, these are also often referred to as Nyquist plots.

The Nyquist criterion may be stated as follows:

A closed-loop control system is stable if the region R (consisting of the entire right half of the s-plane, including the imaginary axis), when mapped onto the G(s)-plane, the open-loop transfer function plane, results in region R', which does not include the point (− 1 ,0).

The following example demonstrates the application of this criterion.

EXAMPLE 9-3.1

Consider the heat exchanger control loop shown in Fig. 9-2.2. The open-loop transfer function is

$$\text{OLTF} = \frac{0.8K_c}{(10s + 1)(30s + 1)(3s + 1)}$$

The Nyquist criterion requires us to map the entire right-hand plane (RHP) of the *s*-plane, as shown in Fig. 9-3.6, onto the G(s)-plane. To best show this procedure, we will divide the mapping into three steps.
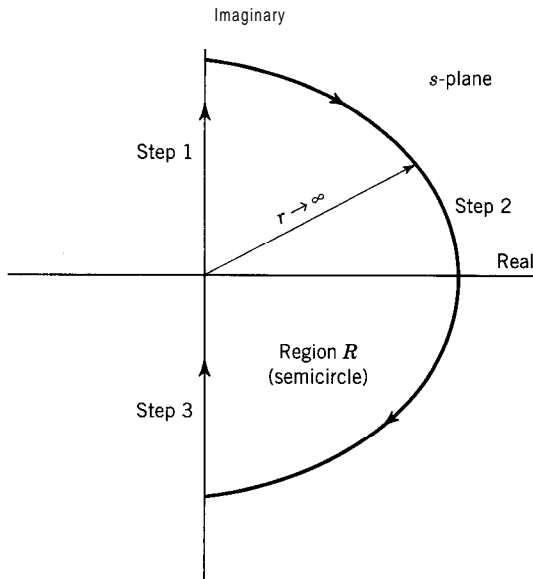
Figure 9-3.6 Region *R* of the s-plane.
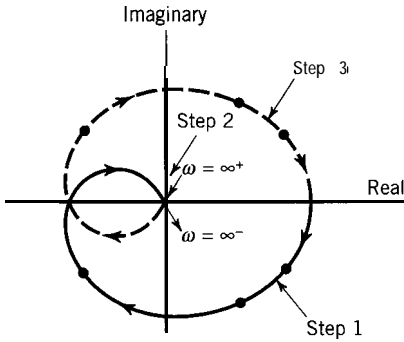
Figure 9-3.7 Polar plot of G(s) =
$$\frac{0.8K_c}{(10s + 1)(30s + 1)(3s + 1)},$$

Step 1. Frequency ω goes from 0 to CC on the positive imaginary axis, so $s = io$. Substituting this expression for s into the G(s) expression yields

$$G(s) = \frac{0.8K_c}{(10i\omega + 1)(30i\omega + 1)(3i\omega + 1)}$$

The plot starts, $\omega = 0$, on the positive real axis at the value $0.8K_c$ and terminates, $\omega \rightarrow \infty$, at the origin (AR = 0) with a phase angle of $-270°$, shown in Fig. 9-3.7. The frequency at which the locus crosses the 180" value is the ultimate frequency and can be found, as before, by solving Eq. 9-2.2 for ω. Once this value of ω is obtained, the value of AR is calculated by using Eq. 9-2.1.

Step 2. In this step the frequency moves from $\omega = \infty$ to $\omega = -\infty$ along the path shown in Fig. 9-3.6. Along this path, $s = re^{i\sigma}$, with σ going from 90" to -90". Substituting this expression for s into the G(s) expression gives

$$G(s) = \frac{0.8K_c}{(10re^{i\sigma} + 1)(30re^{i\sigma} + 1)(3re^{i\sigma} + 1)}$$

Because $r \rightarrow \infty$, the $+ 1$ term in each of the parentheses can be neglected, and

$$\lim G(i\omega) = \lim \left[\frac{0.8K_c}{(10re^{i\sigma})(30re^{i\sigma})(3re^{i\sigma})} \right]$$

$$\lim G(i\omega) = \lim \left[\frac{0.8K_c}{900r^3e^{i3\sigma}} \right] = 0$$

which says that the semicircle with $r \rightarrow \infty$ maps in the G(s)-plane as a point at the origin.

Step 3. In this step the frequency moves from $-\infty$ to 0 along the negative imaginary axis, so $S = -i\omega$. Again, substituting this expression for s into the G(s) expression yields

$$G(i\omega) = \frac{0.8K_c}{(-i10\omega + 1)(-i30\omega + 1)(-i3\omega + 1)}$$

The plot starts at the origin, $\omega = -\infty$, and terminates, $\omega = 0$, on the positive real axis at the value of $0.8K_c$ with a phase angle of 0"; the path is shown in Fig. 9-3.7.

Note that step 3 is just the "mirror image" of step 1. This is understandable if one realizes that G(s) is a complex conjugate function, which means that the map is symmetrical around the abscissa axis in the G(s)-plane.

Step 1 explained how to obtain the value of AR for a phase angle of 180". This AR is the distance from the origin at which the path crosses the negative real axis. If this cross point is to the right of -1, then the system is stable [the mapped region does not

(a) First-order lag.

(b) Dead time.
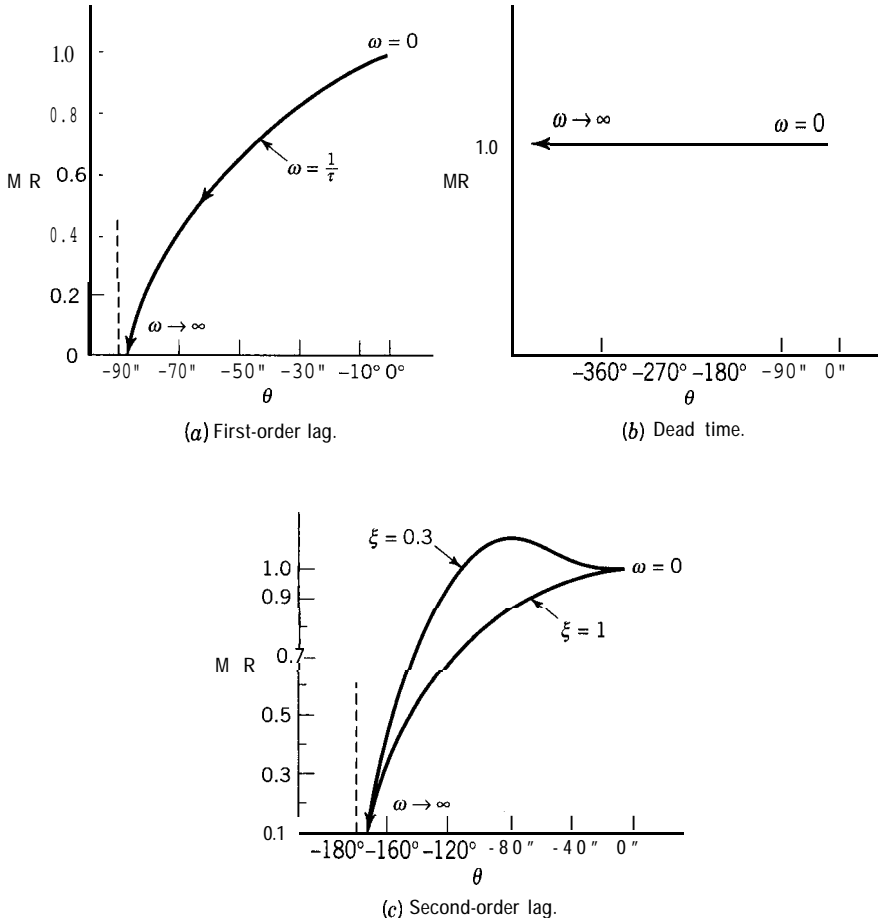
(c) Second-order lag.

Figure 9-4.1 Nichols plots. (a) First-order lag. (b) Dead time. (c) Second-order lag.

include point $(-1,0)$]. If the cross point is to the left of -1, then the system is unstable [the mapped region includes point $(-1,0)$]. As K_c increases, AR also increases, resulting in a less stable system. This is the same statement as the frequency response stability criterion.

The section has presented a brief introduction to polar plots and the Nyquist stability criterion. The reader has no doubt noticed the equivalence between the Nyquist stability criterion and the frequency response stability criterion.

9-4 NICHOLS PLOTS

The Nichols plot is still another way to represent the frequency response of systems graphically. Essentially, it is a plot of the amplitude ratio (or magnitude ratio) versus phase angle. Fig. 9-4.1 shows this type of plot for some typical systems. In these plots, frequency is the parameter along the curve.

9-5 PULSE TESTING

A most practical and interesting application of frequency response is in the use of pulse testing for finding the transfer functions of existing processes, instruments, and other devices. Hougen (1964) presents several industrial applications of pulse testing. In this section, we describe the technique and derive the basic formulas required for its application.

In Chapter 7 we learned the method of step testing for finding the parameters of a first-order-plus-dead-time model of the process. The advantages of step testing are simplicity and small computation requirements. Its major disadvantage is that it is not accurate for models higher than first-order.

The technique of sinusoidal testing described in Section 9-1 can in principle find the transfer function of a process of any order. But though it is used extensively for finding the transfer functions of sensors, transmitters, and control valve actuators, sinusoidal testing is not useful for identifying actual processes. This is because most processes are too slow for sinusoidal testing. Consider a process in which the longest time constant is 1 min. The breakpoint frequency on the Bode plot of such a process is, from Eq. 9-1.27, at

$$\omega_c = \frac{1}{\tau} = 1.0 \ \text{rad/min}$$

To locate the low-frequency asymptote of this process, we must carry out at least two tests at lower frequencies than the breakpoint frequency, say 0.5 and 0.25 rad/min. For this latter test, the period of the sine wave is

$$T = \frac{2\pi}{\omega} = \frac{6.28 \ \text{rad}}{0.25 \ \text{rad/min}} = 25.1 \ \text{min!}$$

This period is not practical for several reasons. First, it is difficult to find a sine wave

generator that can consistently generate such a slow signal. Second, the test would take at least a couple of hours, given that four or five cycles are required for a complete test. Finally, such a test would give us only one point on the Bode plot. If our process time constant were 10 min, then a sine wave with a period of over 4 h would have to be applied to the process.

Pulse testing produces a complete Bode plot of the process from a single test lasting considerably less time than the test described in the preceding paragraph. Given that we cannot get something for nothing, we must pay for the savings in testing effort with additional computational effort.

9-5.1   Performing the Pulse Test

The diagram for the pulse test is similar to the one for the sinusoidal test in Fig. 9-1.1. However, instead of the sinusoidal signal and response of Fig. 9-1.2, the input signal is a pulse such as the one shown in Fig. 9-5.1. Note that the duration of the response, T_F, is longer than that of the pulse, T_D. The three parameters to be selected to carry out a pulse test are the shape of the pulse, its amplitude, and its duration.

Although the rectangular pulse of Fig. 9-5.1 is the easiest to generate and to analyze, we can use other pulse shapes, such as those shown in Fig. 9-5.2. The rectangular pulse is predictably the most popular, followed by the rectangular doublet of Fig. 9-5.2b. A most important requirement for the pulse shape is that it return to its initial steady-state value.

As in the case of step and sinusoidal testing, the amplitude of the pulse, X_0, must be large enough for the measurements of the response to be accurate. However, it must not be so large as to force the response outside the range within which the linear transfer function is a valid approximation of the process response. This requirement usually necessitates a very sensitive recorder or an on-line digital computer for recording the response.

The duration T_D of the pulse depends entirely on the time constants of the process being tested. It should be long enough to allow the process to react but not so long as to let the response reach steady state before the pulse is completed. Such a long pulse represents a waste of test time. It also results in a reduction of the highest frequency for which the test results are useful, as we shall see shortly.
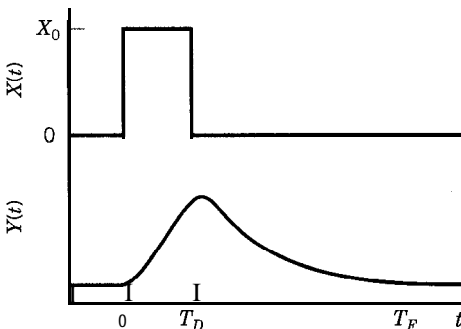
Figure 9-5.1 Rectangular pulse input and output   response.
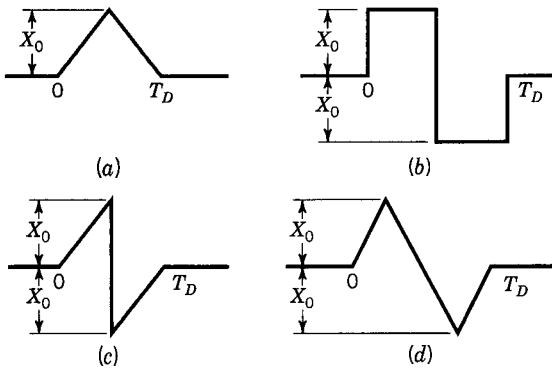
Figure 9-5.2 Pulse shapes for process dynamic testing.
(a) Triangular pulse. (b) Rectangular doublet. (c) Ramp.
(d) Triangular doublet.

9-5.2 Derivation of the Working Equation

In pulse testing, we find the frequency response of the process by calculating the complex transfer function $G(i\omega)$, as a function of frequency, from the process response to the pulse input. To do this, we must use the definition of the Fourier transform of a signal, which is

$$F(i\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t}\, dt \qquad (9\text{-}5.1)$$

Comparing Eq. 9-5.1 with Eq. 2-1.1, we see that except for the lower limit on the integral, we can obtain the Fourier transform by substituting $s = i\omega$ in the definition of the **Laplace** transform. Because the signals of interest in process control are deviations from the initial steady-state value, they are zero for negative time. Consequently, we can change the lower limit on the integral of Eq. 9-5.1 to zero. The Fourier transform was developed before the **Laplace** transform as an extension of the Fourier series to nonperiodic signals.

By definition of the process transfer function (see Chapter 3),

$$\mathbf{G(s)} = \frac{Y(s)}{X(s)} \qquad (9\text{-}5.2)$$

where

$Y(s) =$ the **Laplace** transform of the process response (as a deviation
from its initial steady-state value)
$X(s) =$ the **Laplace** transform of the pulse

Substitute $s = i\omega$ to get

$$G(i\omega) = \frac{Y(i\omega)}{X(i\omega)} \tag{9-5.3}$$

and apply Eq. 9-5.1 to both signals to obtain

$$G(i\omega) = \frac{\displaystyle\int_0^\infty Y(t)e^{-i\omega t}\,dt}{\displaystyle\int_0^\infty X(t)e^{-i\omega t}\,dt} \tag{9-5.4}$$

Equation (9-5.4) is our working equation for calculating the frequency response of the process being tested. From the response, Y(t), to a single pulse, X(t), we can calculate, for each value of frequency ω of interest, the integrals in the numerator and the denominator of Eq. 9-5.1. The result of the calculation is a single complex number $G(i\omega)$. The magnitude of this number is then, from Eq. 9-1.9, the amplitude ratio, and its argument is, from Eq. 9-1.10, the phase angle, at frequency ω. By repeating the calculations for several values of ω, we can generate the entire Bode plot from the results of a single test.

We must numerically calculate the Fourier transform integral of the process output signal, Y(t), as we will see shortly. On the other hand, for the integral of the input pulse, X(t), we can derive an analytical formula starting from Eq. 9-5.1, as illustrated in the following   example.

EXAMPLE 9-5.1

Derive the Fourier transform of a rectangular pulse of amplitude X_0 and duration T_D (see Fig. 9-5.1).

SOLUTION

From Eq. 9-5.1, we know that

$$X(i\omega) = \int_0^\infty X(t)e^{-i\omega t}\,dt$$

Because the pulse is zero at all times except between zero and T_D, we can say that

$$X(i\omega) = \int_0^{T_D} X_0 e^{-i\omega t}\,dt = -\frac{X_0}{i\omega}[e^{-i\omega t}]_0^{T_D}$$

$$X(i\omega) = \frac{X_0}{i\omega}[1 - e^{-i\omega T_D}]$$

$$X(i\omega) = \frac{X_0}{\omega} [\sin \omega T_D - i(1 - \cos \omega T_D)]$$

where we have made use of the identity

$$e^{-i\omega T_D} = \cos \omega T_D \quad i \sin \omega T_D$$

and $1/i = -i$. The magnitude and argument of $X(i\omega)$ are

$$|X(i\omega)| = \frac{X_0}{\omega} \sqrt{\sin^2 \omega T_D + (1 - \cos \omega T_D)^2}$$

$$= \frac{X_0}{\omega} \sqrt{2(1 - \cos \omega T_D)}$$

$$\measuredangle X(i\omega) = \tan^{-1} \left(\frac{1 - \cos \omega T_D}{\sin \omega T_D} \right)$$

At $\omega = 0$, $X(O) = X_0 T_D$. The magnitude of the pulse is a maximum at $\omega = 0$, and then it drops to zero as the frequency increases to infinity. The magnitude is also zero at values of the frequency ω that are multiples of $2\pi/T_D$ radians/time. These values occur more frequently as the pulse duration increases.

From the preceding example, we see that the highest magnitude of the Fourier transform of the rectangular pulse is proportional to the area of the pulse, $X_0 T_D$. Because the Fourier transform of the pulse appears in the denominator of Eq. 9-5.3, we want to avoid values of ω at which the transform is zero. This imposes an upper limit on the range of frequencies for which the frequency response can be calculated from the pulse test. For the rectangular pulse, the upper limit on the frequency is $2\pi/T_D$, which, as mentioned earlier, decreases with the duration of the pulse, T_D. This means that the larger T_D, the smaller the frequency range from where useful information can be obtained. Thus a small T_D provides a larger useful frequency range. However, the smaller T_D is, the less time is given to the process to react, and the less information is obtained from the process. Therefore, a compromise must be reached. It usually requires some testing before a final value of T_D is chosen.

9-5.3 Numerical Evaluation of the Fourier Transform Integral

Several computer programs are available for the numerical calculation of the Fourier transform of a function of time (e.g., MATLAB; see the References). The trapezoidal rule of integration is the basis for an efficient and accurate technique for computing the Fourier transform. In this section we present the resulting formulas without derivation.

Assume we divide the response interval, zero to T_F in Fig. 9-5.1, into N equal incre-

ments of duration At. Then we can calculate the Fourier integral of the output response to the pulse by the following formula:

$$Y(i\omega) = \frac{4 \sin^2\left(\dfrac{\omega\,\Delta t}{2}\right)}{\omega^2\,\Delta t} \sum_{k=1}^{N-1} Y_k e^{-i\omega k \Delta t} \tag{9-5.5}$$

where

Y_k = the output at time k At

$N = T_F/\Delta t$ = the number of intervals

and we have assumed that the output response returns to its initial value. Because Y_k is the deviation from the initial value, then Y_0 and Y_N are equal to zero.

Equation 9-5.5 can be programmed directly in FORTRAN or some other high-level computer language. It is efficient because it requires evaluation of the sine function only once for each frequency of interest.

When the Fourier transforms of both the output response and the input pulse are calculated numerically, substitution of Eq. 9-5.5 for both the numerator and the denominator of Eq. 9-5.3 results in the following formula for the process transfer function:

$$G(i\omega) = \left| \frac{\displaystyle\sum_{k=1}^{N-1} Y_k e^{-i\omega k \Delta t}}{\displaystyle\sum_{j=1}^{M-1} X_j e^{-i\omega j \Delta t}} \right| \tag{9-5.6}$$

where M is the number of increments used to integrate the input pulse $(M = T_D/\Delta t)$. We have assumed the same integration interval for the numerator as for the denominator in Eq. 9-5.6. When the Fourier transform of the pulse can be calculated from an analytical formula such as the one derived in Example 9-5.1, it is usually more efficient to do so than to use Eq. 9-5.6.

To generate the Bode plot for the process, the calculation of $G(i\omega)$ must be repeated at several values of the frequency ω. By selecting each frequency as a constant factor of the previous one, we can make the points on the Bode plot equally spaced in the logarithmic frequency axis.

$$\omega_1 = \beta\omega_{i-1} \qquad \text{for } i = 1, 2, \ldots, N_\omega \tag{9-5.7}$$

where

$$\beta = \left(\frac{\omega_{max}}{\omega_{min}}\right)^{1/N_\omega}$$

ω_{max} = the upper limit of the frequency range on the Bode plot, rad/time

ω_{min} = the lower limit of the frequency range on the Bode plot, rad/time

N_ω = the number of increments into which the frequency range is divided
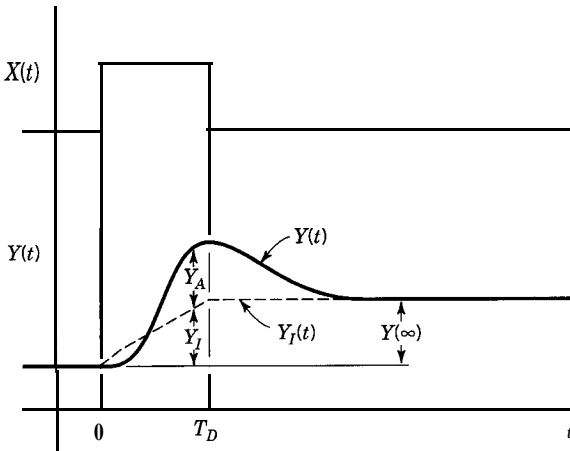
Figure 9-5.3 Handling pulse response of integrating processes.

The procedure just described enables us to obtain the complete Bode plot from a single pulse test. However, as mentioned earlier, the response must return to its initial value. It does not do so when the process contains an integration (pole at the origin). We will consider this case next.

Processes with Integration

Figure 9-5.3 shows the response of a process with integration to a pulse input. The final steady-state value of the output deviation variable is proportional to the integral of the input pulse:

$$Y_\infty = K_I \int_0^{T_D} X(t) \, dt \tag{9-5.8}$$

Luyben (1990) proposes that the process be postulated to consist of a pure integrator with gain K_I in parallel with a fictitious process with transfer function $G_A(s)$. The process transfer function is

$$G(s) = G_A(s) + \frac{K_I}{s} \tag{9-5.9}$$

and the output of the integrator is

$$Y_I(t) = K_I \int_0^t X(t) \, dt \qquad \text{for } 0 \le t < T_D \tag{9-5.10}$$

$$Y_I(t) = Y_\infty \qquad \text{for } t \ge T_D \tag{9-5.11}$$

We then obtain the output of the fictitious process by

$$Y_A(t) = Y(t) - Y_I(t) \qquad (9\text{-}5.12)$$

This makes the signal $Y_A(t)$ zero at both the initial and the final times, as illustrated in Fig. 9-5.3. We calculate $Y_A(t)$ from Eq. 9-5.12 and then use the result in Eqs. 9-5.5 and 9-5.3 to calculate $G_A(i\omega)$. Then, from Eq. 9-5.9, the process transfer function is

$$G(i\omega) = G_A(i\omega) + \frac{K_I}{i\omega} \qquad (9\text{-}5.13)$$

where the gain K, of the integrator is calculated from Eq. 9-5.8.

In this section we have outlined the pulse testing method for experimentally obtaining the transfer function of a process. Luyben (1990) presents a computer program for generating the Bode plot from pulse response data.

9-6 SUMMARY

This chapter presented frequency response techniques for the analysis and design of feedback control systems. These techniques can handle the presence of dead time directly, without approximation. Frequency response is also the basis for obtaining process dynamic parameters through the pulse testing method. As we saw in the last section, we can obtain the entire frequency response of a process or instrument from the results of a single pulse test.

Having studied the design and analysis of feedback control systems, we will next look into other important control strategies that are commonly used in industry. This is the subject of our next chapters.

REFERENCES

1. Hougen, J. 0.1964. "Experiences and Experiments with Process Dynamics." *CEP Monograph Series* 60(4).
2. Luyben, W. L. 1990. *Process Modeling, Simulation, and Control for Chemical Engineers,* 2nd ed. New York: McGraw-Hill, Section 14.3.
3. MATLAB, The MathWorks, Inc., Natick, Mass.
4. Nyquist, H. 1932. "Regeneration Theory." *Bell System Technical Journal* 11: 126– 147.

PROBLEMS

9-1. Draw the asymptotes of the Bode plot magnitude ratio (or amplitude ratio) and roughly sketch the phase angle plot for the transfer functions given in Problem 8-1.

9-2. Repeat Problem 9-1 for the transfer functions given in Problem 8-2.

9-3. Consider the following transfer functions:

$$G(s) = \frac{4s + 1}{s + 1} \quad \text{and} \quad G(s) = \frac{1 + 2s}{(2s + 1)(4s + 1)}$$

(a) Sketch the asymptotes of the magnitude ratio part of the Bode plot, marking the breakpoint frequencies.
(b) Indicate the phase lag (or lead) at high frequencies ($\omega \to \infty$).

9-4. Sketch the Bode plot of the transfer functions given in Problem 8-4.

9-5. Figure P9-1 shows the Bode plot of an open-loop system. Obtain the transfer function for this system. What controller gain can be tolerated if a gain margin of 2 is desired? What is the phase margin with a controller gain of 0.6? What controller gain is required for MPH = 1.26?

9-6. Consider the vacuum filter process of Problem 7-15. Using frequency response techniques:
(a) Sketch the asymptotes of the Bode plot and the phase angle plot.
(b) Obtain the ultimate gain, K_{cu}, and the ultimate period, T_u.
(c) Tune the reset time of a PI controller by the controller synthesis method and determine the controller gain that would provide a gain margin of 2.

9-7. Consider the absorber of Problem 7-16. In that problem, a feedback control loop was designed to control the exit concentration of ammonia. For this control loop:
(a) Sketch the asymptotes of the Bode plot and the phase angle plot.
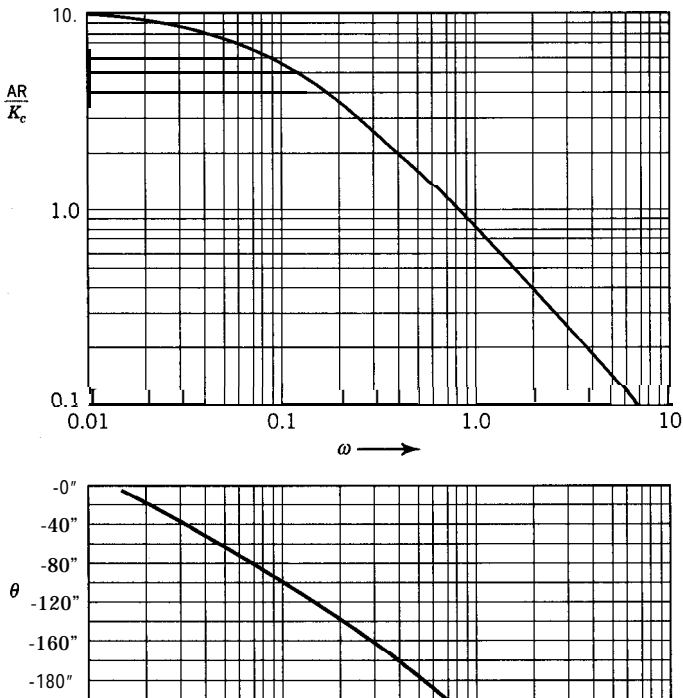
Figure P9-1 Bode diagram for Problem 9-5.

(b) Obtain the ultimate gain, K_{cu}, and the ultimate period, T_u.

(c) Tune a P controller for a phase margin of 45".

(d) Tune a P controller for MPH = 1.26.

9-8. Consider a process with the following transfer function:

$$G(s) = \frac{4.0}{(1s + 1)(0.8s + 1)(0.2s + 1)}$$

(a) Tune a P controller to obtain MPH = 1.26.

(b) If a PI controller is used with Ziegler-Nichols setting, what gain margin, phase margin, and MPH are obtained?

9-9. Do Problem 9-8 for the process transfer function

$$G(s) = \frac{6(1 - s)}{(s + 1)(0.5s + 1)}$$

9-10. Consider the block diagram shown in Fig. P9-2a. The input N(s) represents noise that corrupts the output signal. If this process noise is significant, the control of the process may be difficult. To improve the control of noisy processes, filtering the feedback signal is usually done. A typical way to filter signals is by a filter device with a first-order transfer function. This device-either pneumatic, electronic, or digital-is installed between the transmitter and the controller as shown in Fig. P9-2b. The gain of the filter is 1, and its time constant, called the filter time constant, is τ_F. Using frequency response techniques, explain how τ_F affects the filtering of the noisy signal and the performance of the control loop. Specifically, plot the gain margin as a function of τ_F.

(a)

(b)
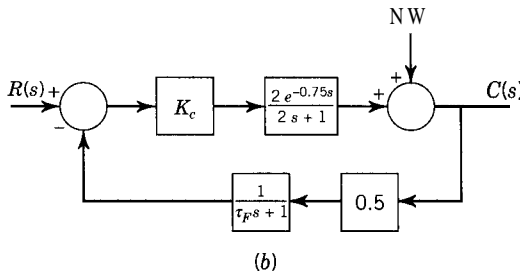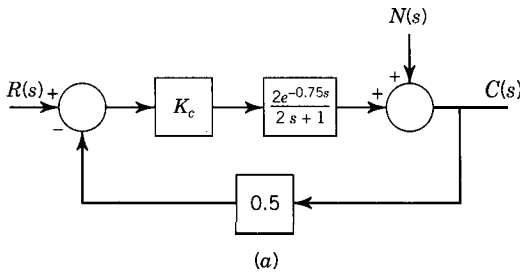
Figure P9-2 Block diagrams for Problem 9-10.

9-11. Consider a thermal process with the following transfer function for the process output versus controller output signal:

$$\frac{C(s)}{M(s)} = \frac{0\ 65e^{-0.35s}}{(5.1s + 1)(1.2s + 1)}$$

A sine wave of unity amplitude and a frequency of 0.80 rad/min is applied to the process (time constants and dead time are in minutes). Calculate the amplitude and phase lag of the sine wave out of the process (after the transient response dies out).

9-12. The symmetrical rectangular pulse shown in Fig. 9-5.2b has the advantage of averaging out the effect of nonlinearities on the result of the dynamic test.
(a) Derive the Fourier transform of the pulse, $X(i\omega)$.
(b) Write the formulas for the magnitude, $|X(i\omega)|$, and phase angle, $\angle X(i\omega)$, as functions of frequency.

9-13. A ramp pulse of duration T_D and final amplitude X_0 is to be used in a pulse test of a process. Determine the Fourier transform integral of the pulse. The pulse is sketched in Fig. 9-5.2c.

9-14. The asymptotes of the amplitude ratio versus frequency plot for a process results in the sketch given in Fig. P9-3. The phase angle plot does not reach a high-frequency asymptote but becomes more negative as the frequency increases. At a frequency of 1 .O rad/min, the phase angle is − 246". Postulate a transfer function for the process and estimate the gain, the time constants, and the dead time (if any).

9-15. The Bode diagram shown in Fig. P9-4 is obtained for the transfer function of a tubular reactor temperature to the cooling water rate by the pulse testing method. Determine:
(a) The steady-state process gain.
(b) The time constant.
(c) An estimate of the dead time of the system.
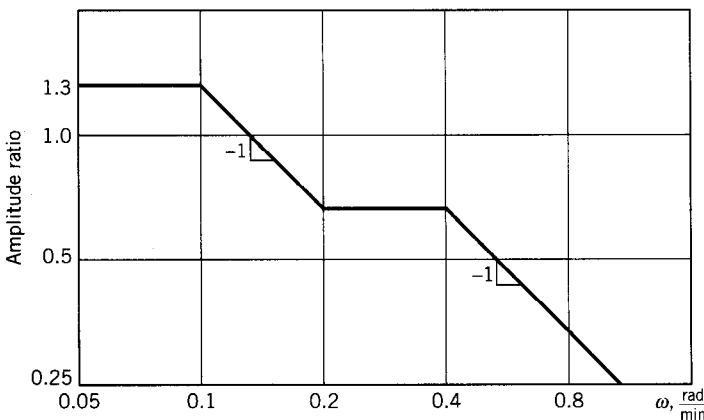
Figure P9-3 Amplitude ratio versus frequency plot for Problem 9-14.
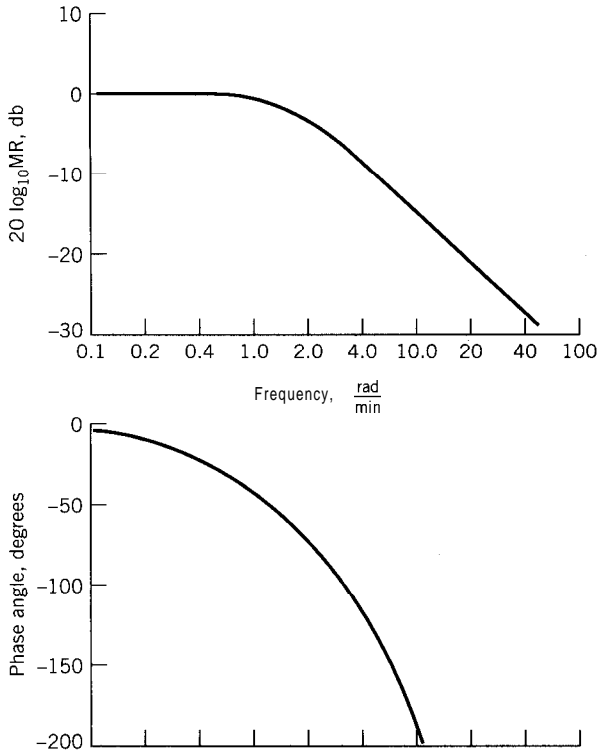
Figure **P9-4** Bode diagram for Problem 9-15.

9-16. Derive the Fourier transforms and their magnitude ratio and phase angle for the triangular pulse sketched in Fig. 9-5.2a.

9-17. In Example 6-2.4 a feedback control loop with a first-order-plus-dead-time process was considered. The direct substitution method, coupled with an approximation of the dead-time term, resulted in approximate formulas for the ultimate gain and frequency of the loop. Compute the ultimate gain and frequency for that system using the frequency response stability criterion, and compare your results with those of the approximate formulas. Use $t_0/\tau = 0.1$, 0.2, 0.5, 1.0, and 2.0. You may assume $\tau = 1$ min without loss of generality.

9-18. Sketch the Bode plot for the feedback loop of Problem 6-3. Use reset times of 0.5, 1.0, and 2.0. How does the Bode plot help you visualize the answer to part (b) of that problem?

9-19. Sketch the Bode plot and verify the ultimate gains and periods for the loops of Problem 6-8.

9-20. Sketch the Bode plot for the open-loop unstable process of Problem 6-10. Verify the range of controller gains for which the closed loop is stable.

9-21. Sketch the Bode plot for the concentration control loop of the three isothermal reactors in series of Problem 6-17. Verify the ultimate gain and period.

9-22. Sketch the Bode plot of the compressor suction pressure control loop of Problem 6-18. Verify the ultimate gain and period.

Chapter **10**

Cascade Control

Chapters 6 and 7 presented the design of feedback control. Feedback control is the simplest form of automatic process control that compensates for process upsets. However, the disadvantage of feedback control is that it reacts only after the process has been upset. That is, when a disturbance enters the process, it has to propagate through the process and make the controlled variable deviate from set point before feedback takes corrective action. Thus a deviation in the controlled variable is needed to initiate corrective action. Even with this disadvantage, probably 80% of all control strategies used in industrial practice are simple feedback control. In these cases, the control performance provided by feedback is satisfactory for safety, product quality, and production rate.

As the processes requirements tighten, however, and in processes with slow dynamics and processes with too many, or frequently occurring, upsets, the control performance provided by feedback control often becomes unacceptable. It is necessary to use other strategies to provide the required performance. These additional strategies are the subject of this and subsequent chapters. The strategies presented complement feedback control; they do not replace it. Remember that it is always necessary to provide some feedback from the controlled variable.

Cascade control is a strategy that improves, in some applications significantly, the performance provided by feedback control. This strategy has long been well known. Computers provide a simpler, safer, and less costly implementation of cascade control than is obtained by the use of analog instrumentation. Therefore, cascade control is implemented more often now, with computers available, than it was when analog instrumentation alone was used. This chapter explains in detail the fundamentals and benefits of cascade control.

10-1 A PROCESS EXAMPLE

Consider the furnace/preheater and reactor process shown in Fig. 10-1.1. In this process, reaction A $\rightarrow$ B occurs in the reactor. Reactant A is usually available at a low temperature, so it must be heated somewhat before being fed to the reactor. The reaction is exothermic, and to remove the heat of reaction, a cooling jacket surrounds the reactor.

The important controlled variable is the temperature in the reactor, T_R. The original control strategy called for controlling this temperature by manipulating the flow of
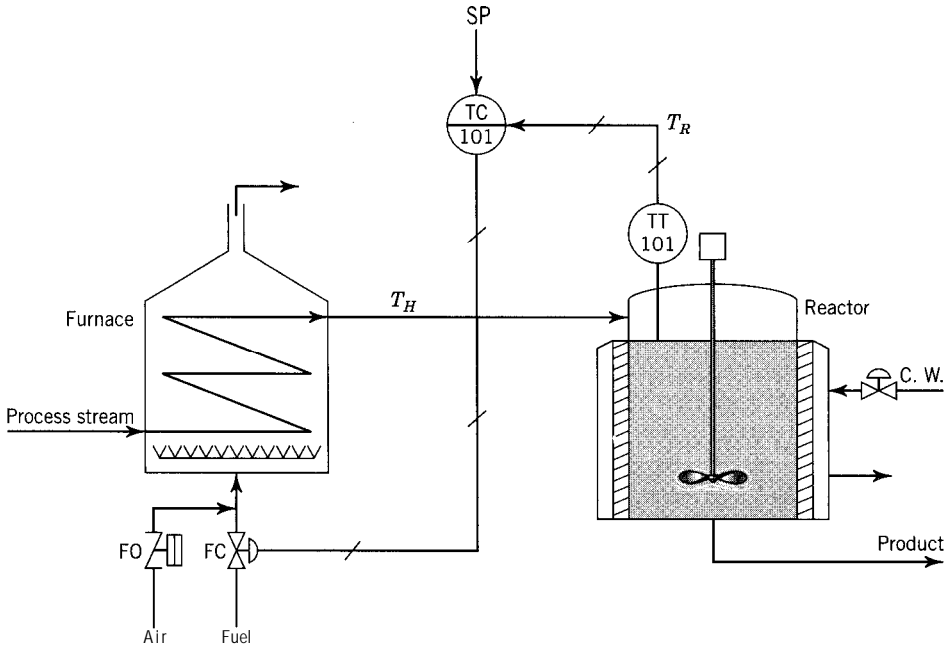
Figure 10-1.1 Preheater/reactor process-feedback control.

coolant to the jacket. The inlet reactant temperature to the reactor was controlled by manipulating the fuel valve. It was noticed during the start-up of this process that the cooling jacket could not provide the cooling capacity required; the cooling valve was open almost all the time. Thus it was decided to open the cooling valve completely and control the reactor temperature by manipulating the fuel to the preheater, as shown in the figure. This strategy worked well enough, providing automatic control during start-up.

Once the process was "lined-out," the process engineer noticed that every so often the reactor temperature would move from set point enough to make off-spec product. After checking the feedback controller tuning to be sure that the performance obtained was the best possible, the engineer started to look for possible process disturbances. Several upsets were found around the reactor itself-cooling fluid temperature and fluid flow variations, etc. -and others around the furnace-variations in inlet temperature of reactant A, in heating value of fuel, in inlet temperature of combustion air, etc. Furthermore, the engineer noticed that every once in a while the inlet reactant temperature to the heater would vary by as much as $25°C$, certainly a major upset.

It is fairly simple to realize that the effect of an upset in the furnace results first in a change in the reactant exit temperature from the furnace, T_H and that this then affects the reactor temperature, T_R. Once the controller senses the error in T_R, it manipulates the signal to the fuel valve. However, with so many lags in the process, furnace plus reactor, it will take a considerable amount of time to bring the reactor temperature back to set point. Because of these lags, the simple feedback control shown in the figure will result in cycling and in general sluggish control.

A superior control strategy can be designed by making use of the fact that the upsets in the furnace first affect T_H. Thus it is logical to start manipulating the fuel valve as

soon as a variation in T_H is sensed, before T_R starts to change. That is, the idea is not to wait for an error in T_R to start changing the manipulated variable. This control strategy uses an intermediate variable, T_H in this case, to reduce the effect of some dynamics in the process. This is the idea behind cascade control, and it is shown in Fig. 10-1.2.

This strategy consists of two sensors, two transmitters, two controllers, and one control valve. One sensor measures the intermediate, or secondary, variable, T_H in this case, and the other sensor measures the primary controlled variable, T_R. Thus this strategy results in two control loops: one loop controlling T_H, and the other loop controlling T_R. To repeat, the furnace exit temperature is used only as an intermediate variable to improve the control of the reactor temperature, which is the important control objective (controlled variable).

The strategy works as follows: Controller TClOl looks at the reactor temperature and decides how to manipulate the furnace outlet temperature to satisfy its set point. This decision is passed on to TC102 in the form of a set point. TC102 in turn manipulates the signal to the fuel valve to maintain T_H at the set point required by TC101. If one of the upsets mentioned earlier enters the furnace, T_H deviates from set point and TC102 takes corrective action right away, *before* T_R changes. Thus the dynamic elements of the process have been separated to compensate for upsets in the furnace before they affect the primary controlled variable.

In general, the controller that keeps the primary variable at set point is referred to as the master controller, outer controller, or primary controller. The controller used to maintain the secondary variable at the set point required by the master controller is usually referred to as the slave controller, inner controller, or secondary controller. The terminology primary/secondary is commonly preferred for systems with more than two cascaded loops, because it extends naturally.
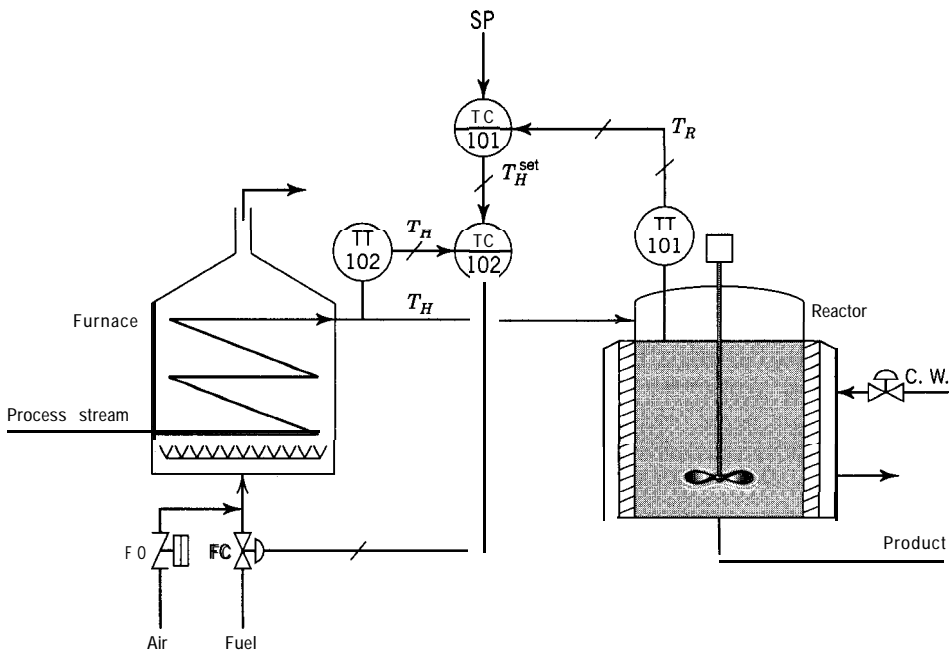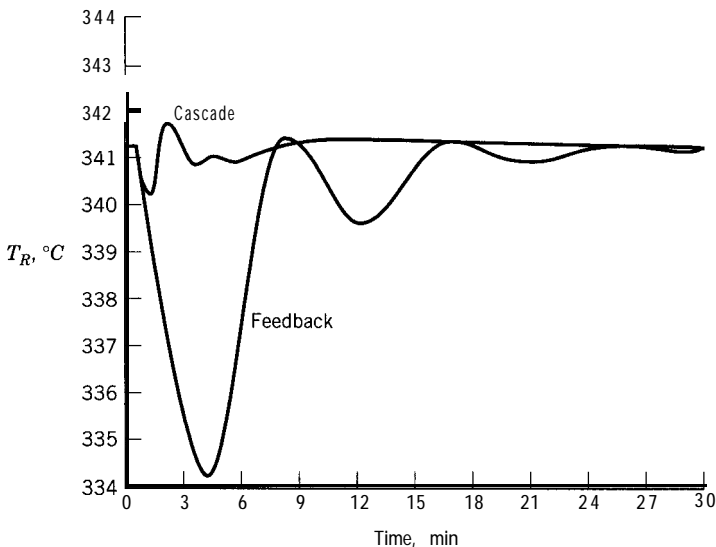

Figure 10-1.2 Preheater/reactor process-cascade control.

Figure 10-1.3 Response of reactor temperature to a change of -25°C in feed temperature to heater.

Figure 10-1.3 shows the response of the process to a — 25°C change in inlet reactant temperature under simple feedback control and under cascade control. The improvement is very significant and probably pays for the added expenses in no time.

The following must be stressed: In designing cascade control strategies, *the most important consideration is that the inner loop must be faster than the outer loop and the faster the better.* This requirement makes sense, and it is extended to any number of cascade loops. In a system with three cascaded loops, as shown in Section 10-3.2, the tertiary loop must be faster than the secondary loop, which must be faster than the primary loop.

Note that the innermost controller is the one that sends its output to the valve. The outputs of all other controllers are used as set points to other controllers; the final control element of these controllers is the set point of another controller.

As this example illustrates, we are starting to develop more complex control schemes than simple feedback. It is helpful, in developing these schemes and others shown in the following chapters, to remember that every signal must have a physical significance. In Figs. 10-1.1 and 10-1.2, we labeled each signal with its significance. For example, in Fig. 10-1.2 the output signal from TT101 indicates the temperature in the reactor, T_R; the output signal from TT102 indicates the outlet temperature from the heater, T_H; and the output signal from TC101 indicates the required temperature from the heater, T_H^{set}. Even though indicating the significance of the signals in control diagrams is not standard practice, we will continue to do so. This practice helps in understanding control schemes, and we recommend that the reader do the same.

10-2   STABILITY   CONSIDERATIONS

We will now look at how the implementation of cascade control affects the stability of the control system. Figure 10-2.1 is the block diagram of the simple feedback control
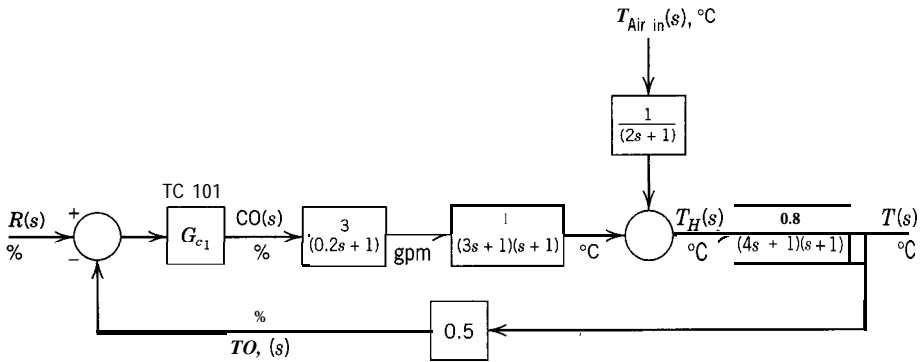
Figure **10-2.1** Block diagram of the process shown in Figure 10-1.1.

strategy shown in Fig. 10-1.1, and Fig.lO-2.2 is the block diagram of the cascade strategy shown in Fig. 10-1.2. Simple transfer functions have been selected to represent the process.

The block diagram of Fig. 10-2.2 clearly shows why the secondary loop starts to compensate for any disturbance that affects the secondary controlled variable before its effect is felt by the primary controlled variable. The diagram also shows why the secondary loop is sometimes referred to as the inner loop. This loop is imbedded inside the primary loop, or outer loop.

The characteristic equation for the simple feedback control system, Fig. 10-2.1, is

$$1 + \frac{1.2G_{c_1}}{(0.2s + 1)(3s + 1)(s + 1)(4s + 1)(s + 1)} = 0 \qquad (10\text{-}2.1)$$

Using the block diagram algebra techniques learned in Chapter 3, we obtain the characteristic equation for the cascade control strategy, Fig. 10-2.2, as

$$1 + \frac{\dfrac{1.2G_{c_1}G_{c_2}}{(0.2s + 1)(3s + 1)(s + 1)(4s + 1)(s + 1)}}{1 + \dfrac{1.5G_{c_2}}{(0.2s + 1)(3s + 1)(s + 1)}} = 0 \qquad (10\text{-}2.2)$$

Applying the direct substitution method (Chapter 6) or the frequency response technique (Chapter 9) to the feedback control system, Eq. 10-2.1, we can calculate the ultimate gain and ultimate frequency.

$$K_{cu} = 4.33 \ \frac{\%CO}{\% TO} \qquad \text{and} \qquad \omega_u = 0.507 \ \frac{rad}{min}$$

To determine the ultimate gain and frequency of the primary controller of the cascade strategy, the tuning of the secondary controller must first be obtained. This tuning can
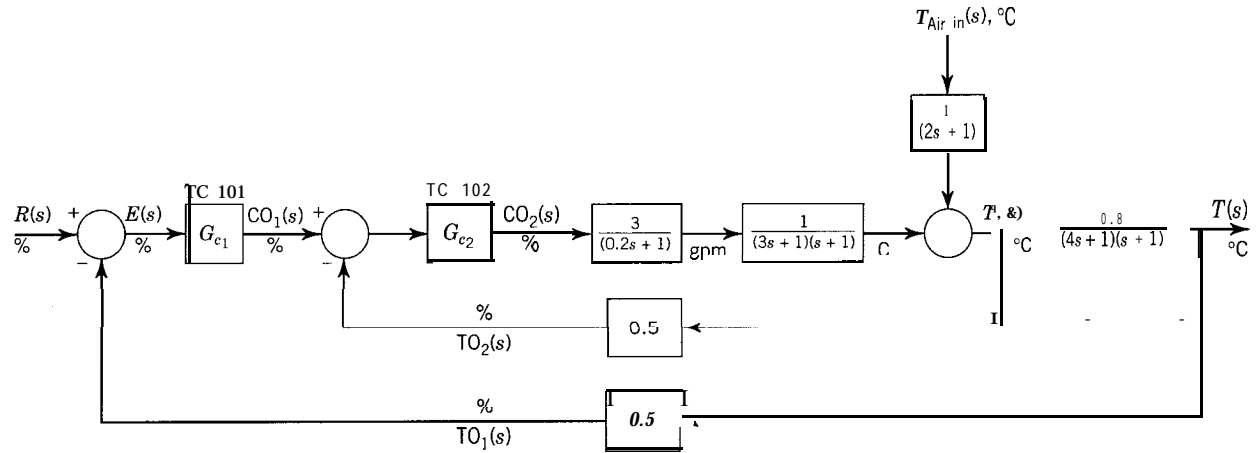
Figure 10-2.2 Block diagram of the process shown in Figure 10-1.2.

be obtained by determining the ultimate gain of the secondary loop in Fig. 10-2.2. The characteristic equation for this secondary process is

$$1 + \frac{1.5G_{c_2}}{(0.2s + 1)(3s + 1)(s + 1)} = 0$$

and it yields

$$K_{cu_2} = 17.06 \frac{\%CO_2}{\%TO_2}$$

Assuming a proportional-only controller and using the Ziegler-Nichols suggestion, we find that the tuning for the secondary controller becomes

$$K_{c_2} = 0.5K_{cu_2} = 8.53 \frac{\%CO_2}{\%TO_2}$$

When we use this tuning value for the secondary controller, the characteristic equation for the cascade strategy (Eq. 10-2.2) yields, for the primary controller,

$$K_{cu_1} = 7.2 \frac{\%CO_1}{\%TO_1} \quad \text{and} \quad \omega_{u_1} = 1.54 \frac{\text{rad}}{\text{min}}$$

Comparing the results, we note that the cascade strategy yields a greater ultimate gain, or limit of stability, 7.2 %CO/%TO vs. 4.33 %CO/%TO, than the simple feedback control loop. The value of the ultimate frequency is also greater for the cascade strategy, 1.54 rad/min vs. 0.507 rad/min, indicating faster process response.

The use of cascade control makes the overall control faster and most times increases the ultimate gain of the primary controller. The methods of analysis are the same as for simple feedback loops.

10-3 IMPLEMENTATION AND TUNING OF CONTROLLERS

Two important questions still remain: how to put the cascade strategy into full automatic operation and how to tune the controllers. The answer to both questions is the same: from inside out. That is, the inner controller is the first to be tuned and put into auto state while the other loops are in manual. As the inner controller is set in cascade, it is good practice to check how it performs before proceeding to the next controller. For the process shown in Fig. 10-1.2, TC102 is tuned first and set in cascade while TClOl is in manual. The control performance of TC102 is then checked before proceeding to TClOl. This checking can usually be done very simply by varying the set point to TC102. Remember, it is desired to make TC102 as fast as possible, even if it oscillates a bit, to minimize the effect of the upsets. Once this is done, TClOl can be tuned and set in automatic. However, before TClOl is set in automatic, TC102 must be set to the cascade state, which means that it will accept the output from TClOl as its set point.

Tuning cascade control systems is more complex than tuning simple feedback systems, if only because there is more than one controller to tune. However, this does not

mean it is difficult. We will first present the tuning methods available for two-level cascade systems and then discuss the methods available for three-level cascade systems.

10-3.1 Two-Level Cascade Systems

The control system shown in Fig. 10-1.2 is referred to as a two-level cascade system. Because the inner loop by itself is a simple feedback loop, TC102 can be tuned by any of the techniques discussed in Chapter 7. As we have said, this controller should be tuned as fast as possible-avoiding instability, of course. The objective is to make the inner loop fast and responsive in order to minimize the effect of upsets on the primary controlled variable. Tuning this system then comes down to tuning the primary controller.

There are several ways to obtain a first-guess tuning for the primary controller. Trial and error is often used by experienced personnel. The other methods available follow a "recipe" to obtain the first tuning values. One such method is the Ziegler-Nichols oscillatory technique presented in Chapter 7. After the secondary controller is tuned and set in cascade, the integral and derivative modes present in the primary controller are removed, and its gain is increased cautiously until the controlled variable oscillates with sustained oscillations. The controller gain that provides these oscillations is called the ultimate gain, K_{cu}, and the period of the oscillations is the ultimate period, T_u. The Ziegler-Nichols equations presented in Chapter 7 are then used.

The second method is the use of direct substitution or frequency response techniques. That is, if all the transfer functions are known, direct substitution or frequency response can be used to tune the controller either by obtaining K_{cu} and T_u and using the Ziegler–Nichols equations, or by applying any other criterion such as the gain or phase margin or the closed-loop response.

The third method available is the one presented by Pressler (see the References at the end of this chapter). Pressler's method was developed for systems where the secondary controller is proportional-only and that the primary controller is proportional-integral. This P/PI combination is usually quite convenient, and the method works well. However, it assumes that the inner loop does not contain dead time, and this assumption limits its application to cascade systems with very fast inner loops, such as flow or liquid pressure loops. For processes with some amount of dead time in the inner loop, such as the one shown in Fig. 10-2.2, applying Pressler's method would yield an unstable response if the master controller were ever set in manual.

The fourth method available to tune cascade systems is the one developed by Austin (1986). The method provides a way to tune both the primary and the secondary controllers with only one test, the step test presented in Chapter 7. Tuning equations are provided for the primary controller, PI or PID, when the secondary controller is either P or PI. The method consists of generating a step change in signal to the control valve, $\%CO_2$, as explained in Chapter 7, and recording the response of the secondary and primary variables. The response of the secondary variable is used to calculate the gain, K_2 in $\dfrac{\%TT102}{\%CO_2}$, the time constant, τ_2, and the dead time, t_{0_2}, of the secondary loop. The response of the primary variable is used to calculate the gain, K_1 in $\dfrac{\%TT101}{\%CO_2}$, the time constant, τ_1, and the dead time, t_{0_1}, of the primary loop. This information and the equations presented in Table 10-3.1 or Table 10-3.2 are used to obtain the tunings of

Table 10-3.1 Tuning Equations-Two-Level Cascade System-for Disturbance Changes

PRIMARY $G_{c_1}(s)$	PI	PID
	$\tau_{I_1} = \tau_1$	$\tau_{I_1} = \tau_1; \ \tau_{D_1} = \dfrac{t_{0_1} - \tau_2}{2}$
SECONDARY $G_{c_2}(s)$	K_{c_1}	K_{c_1}
P	$1.4 \left[\dfrac{1 + K_{c_2}K_2}{K_{c_2}K_1}\right]\left[\dfrac{t_{0_1}}{\tau_1}\right]^{-1.14}\left[\dfrac{\tau_2}{\tau_1}\right]^{0.1}$	$1.4 \left[\dfrac{1 + K_{c_2}K_2}{K_{c_2}K_1}\right]\left[\dfrac{t_{0_1}}{\tau_1}\right]^{-1.14}\left[\dfrac{\tau_2}{\tau_1}\right]^{0.1}$
PI	$1.25 \left[\dfrac{K_2}{K_1}\right]\left[\dfrac{t_{0_1}}{\tau_1}\right]^{-1.07}\left[\dfrac{\tau_2}{\tau_1}\right]^{0.1}$	$1.25 \left[\dfrac{K_2}{K_1}\right]\left[\dfrac{t_{0_1}}{\tau_1}\right]^{-1.07}\left[\dfrac{\tau_2}{\tau_1}\right]^{0.1}$
Range	$0.02 \leq \left(\dfrac{\tau_2}{\tau_1}\right) \leq 0.38$	$0.02 \leq \left(\dfrac{\tau_2}{\tau_1}\right) \leq 0.38$
	$t_{0_2} \leq t_{0_1}$	$t_{0_2} \leq t_{0_1}$
		$\dfrac{t_{0_1} - \tau_2}{2} \geq 0.08$

Note: Use this table if $\tau_2/\tau_1 > 0.38$. Otherwise, use Table 10-3.2.

the primary controller. Table 10-3.1 presents the equations for tuning the primary controller when its set point is constant. However, when the set point to the primary controller is continuously changing with time, the equations in Table 10-3.2 should be used. *Note that when $\tau_2/\tau_1 > 0.38$, Table 10-3.2 should be used. Under this ratio condition, the equations in Table 10-3.2 provide better tuning. The τ_2/τ_1 ratio should always be checked first.*

The response under cascade control shown in Fig. 10-1.3 was obtained with controller tunings calculated using Austin's method. TC102 was set in manual, its output was changed by $+5\%$, and the secondary and primary variables were recorded. The output of TC102 was changed again by -5% and the variables were recorded. The following average terms were obtained:

$$K, \ = 0.738 \ \frac{\%TT101}{\%CO_2} \qquad K_2 = 0.902 \ \frac{\%TT102}{\%CO_2}$$

$$\tau_1 = 11.6 \ \text{min} \qquad \tau_2 = 3.47 \ \text{min}$$

$$t_{0_1} = 3.94 \ \text{min} \qquad t_{0_2} = 0.75 \ \text{min}$$

The inner controller was tuned using the controller synthesis method of Chapter 7, which yielded

$$K_{c_1} = 2.6 \ \frac{\%CO_2}{\%TT102} \qquad \text{and} \qquad \tau_{I_2} = 3.5 \ \text{min}$$

Table   10-3.2   Tuning   Equations-Two-Level   Cascade   System-for   Set   Point   Change

PRIMARY $G_{c_1}(s)$	PI	PID
	$\tau_{I_1} = \tau_1$	$\tau_{I_1} = \tau_1; \; \tau_{D_1} = \dfrac{t_{0_1} - \tau_2}{2}$
SECONDARY $G_{c_2}(s)$	K_{c_1}	K_{c_1}
P	$0.84 \left[\dfrac{1 + K_{c_2}K_2}{K_{c_2}K_1}\right]\left[\dfrac{t_{0_1}}{\tau_1}\right]^{-1.14}\left[\dfrac{\tau_2}{\tau_1}\right]^{0.1}$	$1.17 \left[\dfrac{1 + K_{c_2}K_2}{K_{c_2}K_1}\right]\left[\dfrac{t_{0_1}}{\tau_1}\right]^{-1.14}\left[\dfrac{\tau_2}{\tau_1}\right]^{0.1}$
PI	$0.75 \left[\dfrac{K_2}{K_1}\right]\left[\dfrac{t_{0_1}}{\tau_1}\right]^{-1.07}\left[\dfrac{\tau_2}{\tau_1}\right]^{0.1}$	$1.04 \left[\dfrac{K_2}{K_1}\right]\left[\dfrac{t_{0_1}}{\tau_1}\right]^{-1.07}\left[\dfrac{\tau_2}{\tau_1}\right]^{0.1}$
Range	$0.02 \leq \left(\dfrac{\tau_2}{\tau_1}\right) \leq 0.65$	$0.02 \leq \left(\dfrac{\tau_2}{\tau_1}\right) \leq 0.35$
	$t_{0_2} \leq t_{0_1}$	$t_{0_2} \leq t_0,$ $\dfrac{t_{0_1} - \tau_2}{2} \geq 0.08$

The primary controller was then tuned using Austin's tunings, which yielded

$$K_{c_1} = 4.3 \; \frac{\%CO_1}{\%TT101} \qquad \text{and} \qquad \tau_{I_1} = 11.6 \text{ min}$$

Austin's method provides a simple procedure for obtaining near-optimum tunings for the primary controller. The fact that both controllers can be tuned from information obtained from the same test makes the method even more useful.

Systems in which the innermost loop is very fast, such as a flow or liquid pressure, are most common and thus are worthy of further discussion. One such system is presented in Fig. 10-4.2a. There are several ways to tune the temperature controller. One way is to follow Austin's method. That is, with both controllers in manual, introduce a step change in the output from the flow controller to the valve, and record the flow and temperature responses. From the recordings calculate the respective gains, time constants, and dead times. Because flow loops are quite fast, the time constant will be in the order of seconds and the dead time very close to zero, $t_{0_2} \approx 0.0$ min. We noted in Chapter 6 that flow controllers are usually tuned with low gain, $K_c \approx 0.1$, and short reset time, $\tau_I \approx 0.1$ min. However, in the process shown in Fig. 10-4.2a, the flow controller is the inner controller in a cascade system. Because a fast-responding inner loop is desired, the recommendation in this case is to increase the controller gain close to 1, $K_c \approx 1.0$; to maintain stability, the reset time may also have to be increased (Corripio, 1990). Once the flow controller has been tuned to provide fast and stable response, the temperature controller can be tuned by following Austin's guidelines. It is important to realize that t_{0_2} is not a factor in the equations and therefore will not have an effect in the tuning of the master controller.

Another method is to reduce the two-level cascade system to a simple feedback loop

by realizing that the flow loop is very fast and thus just considering it part of the process. This is done by first tuning the flow controller as previously explained and then setting it in cascade. Once this is done, the flow controller is receiving its set point from the temperature controller. Introduce a step change in the output from the temperature controller, and record the temperature. From the recording calculate the gain, time constant, and dead time, and tune the controller by any of the methods presented in Chapter 7.

10-3.2 Three-Level Cascade Systems

Controller TC102 in the cascade system shown in Fig. 10-1.2 manipulates the valve position to maintain the furnace outlet temperature at set point. The controller manipulates the valve position, not the fuel flow. The fuel flow depends on the valve position and on the pressure drop across the valve. A change in this pressure drop, a common upset, results in a change in fuel flow. The control system, as is, will react to this upset once the outlet furnace temperature deviates from set point. If it is important to minimize the effect of this upset, tighter control can be obtained by adding one extra level of cascade, as shown in Fig. 10-3.1. The fuel flow is then manipulated by TC102, and a change in flow due to pressure drop changes would then be corrected immediately by FC103. The effect of the upset on the outlet furnace temperature would be minimal.

It is important to realize that in this new three-level cascade system, the most inner loop, the flow loop, is the fastest. Thus the necessary requirement of decreasing loop speed from "inside out" is met.

To tune this three-level cascade system, first note that controllers FC103 and TC102
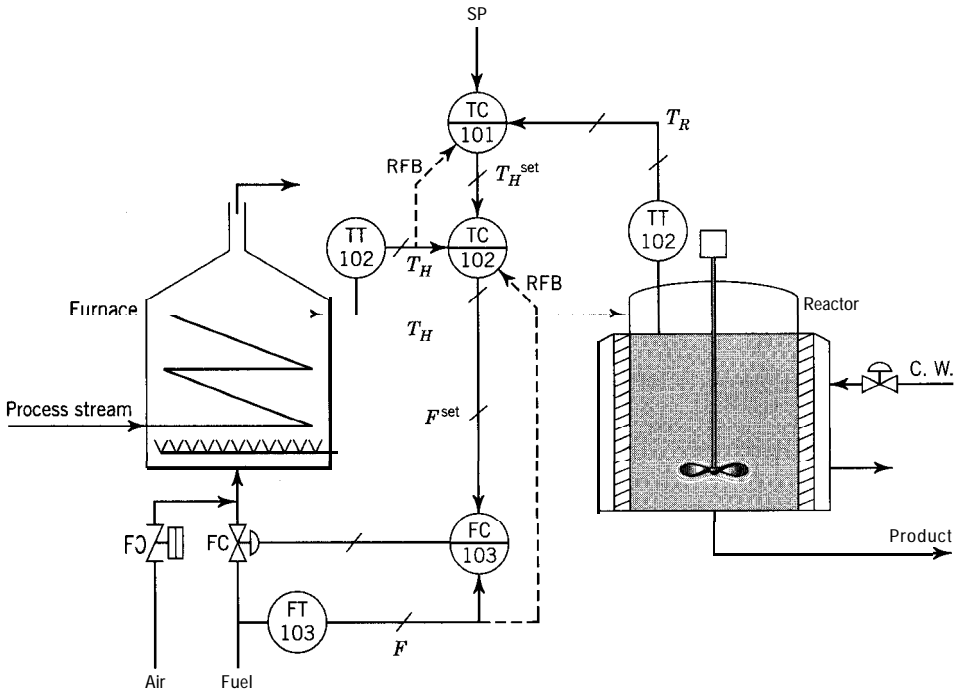
Figure 10-3.1 Preheater/reactor process-three-level cascade control.

constitute a two-level cascade "subsystem" in which the inner loop is very fast. Thus these controllers can be tuned by any of the methods previously presented. The tuning then reduces to tuning the primary controller of a two-level cascade system. Austin's method is very easily applied. With TC101 and TC102 in manual and FC103 in cascade, introduce a step change in the signal from TC102 to FC103, and record the furnace temperature and the reactor temperature. From the furnace temperature response, obtain the gain, K_2 in $\dfrac{\%TT102}{\%CO_2}$; the time constant, τ_2; and the dead time, t_{0_2}. Using the reactor temperature response, obtain the gain, K, in $\dfrac{\%TT101}{\%CO_2}$; the time constant, τ_1 ; and the dead time, t_{0_1}. Then use Table 10-3.1 or Table 10-3.2 to tune the primary controller.

10-4 OTHER PROCESS EXAMPLES

Consider the heat exchanger control system shown in Fig. 10-4.1, in which the outlet process fluid temperature is controlled by manipulating the steam valve position. In the previous section, we noted that the flow through any valve depends on the valve position **and** on the pressure drop across the valve. If a pressure surge in the steam pipe occurs, the steam flow changes. The temperature control loop shown can compensate for this disturbance only after the process temperature deviates from set point.

Two cascade schemes that improve this temperature control, when steam pressure surges are important disturbances, are illustrated in Fig. 10-4.2. Figure 10-4.2a shows a cascade scheme in which a flow loop has been added; the temperature controller provides the flow controller set point. Any flow changes are now compensated for by the flow loop. The cascade scheme shown in Fig. 10-4.2b accomplishes the same con-
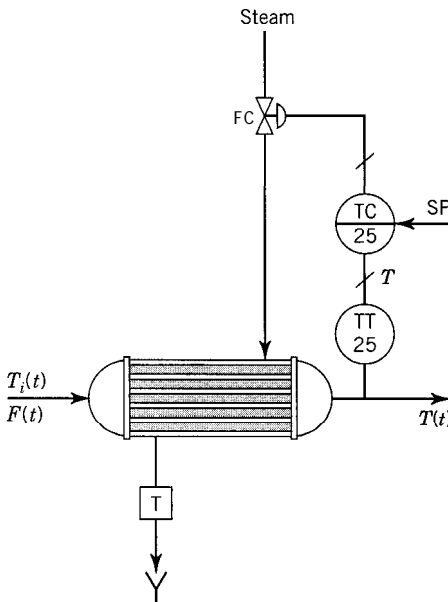
Figure 10-4.1 Heat exchanger temperature control loop.

(a)

(b)
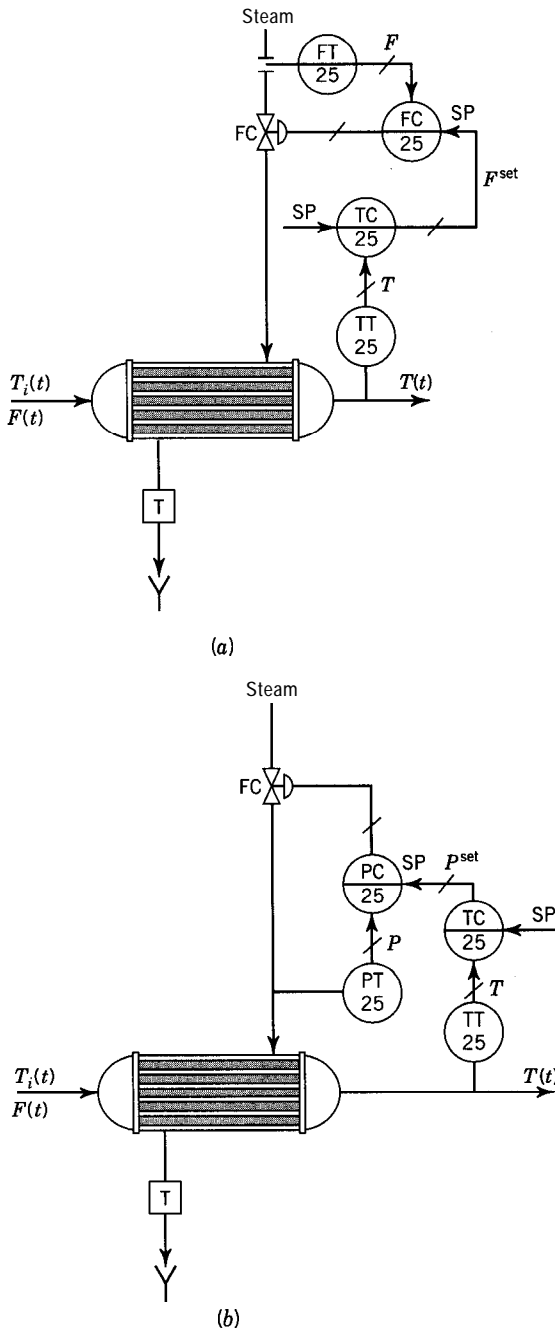
Figure 10-4.2 Cascade control schemes applied to heat
exchanger   temperature   control.

trol, but here the secondary variable is the steam pressure in the exchanger shell side. Actually, this steam pressure is usually measured in the line entering the shell; this is less expensive and safer than measuring the actual pressure in the shell. Any change in steam flow affects the pressure quite rapidly. Any pressure change is then compensated for by the pressure loop. This pressure loop also compensates for disturbances in the heat content (superheat and latent heat) of the steam, because the pressure in the shell side is related to the condensing temperature and thus to the heat transfer rate in the exchanger. This last scheme is usually less expensive in implementation; it does not require an orifice with its associated flanges, which can be costly. Both cascade schemes are common in the process industries. Which of the two schemes do you think gives better initial response to disturbances in inlet process temperature $T_i(t)$?

The cascade control systems shown in Figs. 10-4.2a and **b,** temperature controller cascaded to flow or pressure controllers, are very common in industrial practice. A typical application is in distillation columns where temperature is controlled to maintain the desired split. The temperature controller is often cascaded to the steam flow to the reboiler, or the reflux, or distillate flow.

Finally, another very simple example of a cascade control system is a positioner on a control valve. The positioner acts as the innermost controller of the cascade scheme; positioners are discussed in Appendix C.

10-5 FURTHER COMMENTS

So far, we have said nothing about the action of the controllers in a cascade strategy. This is important because, as we saw in Chapter 5, if the actions are not correctly chosen, the controllers will not control. The procedure for choosing the action is the same as that explained in Chapter 5. That is, the action is determined by process requirements and the fail-safe action of the final control element. As previously mentioned, for some of the controllers in the cascade strategy, the final control element is the set point of another controller.

Consider the three-level cascade strategy shown in Fig. 10-3.1. The fail-safe action of the valve is fail-closed, so the action of FC103 is reverse, because if the flow measurement increases above set point, indicating that more flow than required is being delivered by the valve, then the valve opening must be reduced, and for a fail-closed valve this is accomplished by reducing the signal to it. The action of TC102 is also reverse, because if its measurement increases above set point, indicating a higher outlet furnace temperature than required, then the fuel flow must be reduced, and this is accomplished by reducing the set point to FC102. Finally, the action of TClOl is also reverse, because if its measurement increases above set point, indicating a higher reactor temperature than required, then the way to reduce it is by lowering the inlet reactants temperature, which is accomplished by reducing the set point to TIC102. The decision regarding the controller action is simple and easy so long as we understand the significance of what each controller is doing.

Now an interesting question develops. As we have learned, the output from TClOl is a signal, meaning 4 to 20 **mA** or 3 to 15 psig or, in general, 0 to 100%. Then, for a given signal, say 40%, what is the temperature, in degrees, required from TC102? This question is easy to answer if we remember that the job of the controller is to make its measurement equal to set point. Therefore, TC102 will be satisfied when the signal from TT102 is 40%. Thus the required temperature is 40% of the range of TT102.

Considering Fig. 10-1.2, it is important to realize what would happen if TC 102 were

taken off cascade operation while TClOl was left in automatic. If this were done, and if TClOl sensed an error, it would send a new signal (set point) to TC102. However, TC102 would be unable to respond to requests from TClOl. If TClOl had reset mode, it would wind up, because its output (set point changes to TC102) would have no effect on its controlled variable. That is, the effect of taking the secondary controller off cascade is to "open" the feedback loop of the primary controller.

Computers, with their inherit flexibility, offer the capabilities necessary to avoid this wind-up possibility and thus provide for a safer cascade strategy. The computer is programmed, or configured, so that at any time the secondary controller is taken off cascade, the primary controller "automatically" goes into the manual mode if it is in automatic. The primary controller remains in manual as long as the secondary controller remains off cascade. When the secondary controller is returned to cascade, the primary controller immediately returns to automatic. However, if while the secondary controller is off cascade, its set point changes, then at the moment it is returned to cascade mode, its present set point may not be equal to the output of the primary controller. If this occurs, the set point of the secondary controller will immediately jump to equal the output of the primary controller, thus generating a "bump" in the process operation. To obtain a "bumpless" transfer, controllers are programmed so that while the secondary controller is off cascade, the output from the primary controller is forced to equal either the measurement or the set point of the secondary controller. That is, the output from the primary controller "tracks" either variable of the secondary controller. Thus, when the secondary controller is returned to cascade operation, a smooth transfer is obtained.

The tracking option just explained, which is often referred as output tracking, is very important for the smooth and safe operation of cascade control systems; its implementation is very simple. The implementation is similar to the reset feedback technique presented in Chapter 5; Fig. 10-3.1 shows this implementation. While TClOl is in manual, the RFB from TT102 forces the output of TClOl, which is the set point to TC102, to be equal to the measurement received by TC102. Thus the error is zero, and a bumpless transfer occurs. The same procedure takes place between TC102 and FC103.

10-6 SUMMARY

This chapter has presented in detail the fundamentals and benefits of cascade control, a strategy that is simple in concept and implementation and may enhance the control performance provided by feedback control. The reader must remember that in a two-level cascade strategy the secondary loop must be faster than the primary loop. This requirement extends to higher-level cascade strategies. Typical two-level cascaded loops are temperature-to-flow, concentration-to-flow, pressure-to-flow, level-to-flow, and temperature-to-pressure.

REFERENCES

1. Austin, V. D. 1986. "Development of Tuning Relations for Cascade Control Systems." Ph.D. dissertation, Department of Chemical Engineering, University of South Florida, Tampa, Fla.
2. Corripio, A. B. 1990. *Tuning of Industrial Control Systems.* Research Triangle Park, N.C.: Instrument Society of America.
3. Pressler, Gerhard, Regelungs-Technik, Hochschultashenbucher, Band 63, Bibliographischer Institut, Mannheim, West Germany.

PROBLEMS

10-1. For the paper drying process of Fig. P10-1, the following information is available: The flow control loop (FIC47) can be represented by a first-order lag with a gain of 4 gpm/%CO and a time constant of 0.1 min. The transfer function of the air heater outlet temperature to the fuel flow is a second-order lag with time constants of 2 min and 0.8 min. A change in fuel flow of 1 gpm causes a change of 2°F in the outlet air temperature. The drier can be represented by a first-order lag with a time constant of 5 min. A change of 1°F in inlet air temperature causes a change in outlet moisture of 0.5 mass percent. The moisture transmitter (MT47) has a range of 0 to 6 mass percent and a negligible time constant.

(a) For the control scheme shown in Fig. P10-1, draw the block diagram of the moisture control loop showing the transfer functions. Decide on the fail position of the control valve and the controller action, and make sure that the signs in your block diagram correspond to your decisions. Determine the ultimate gain and period of oscillation of the moisture control loop, and then use these values to tune a PID controller for the moisture controller (MC47) for quarter decay ratio response.

(b) Consider a cascade control scheme using an outlet temperature sensor with a range of 250 to 300°F installed on the air line from the heater to the drier and a temperature controller to manipulate the fuel flow set point; the output of the moisture controller (MC47) sets the set point of the outlet air temperature controller. Draw the instrumentation diagram and the block diagram of
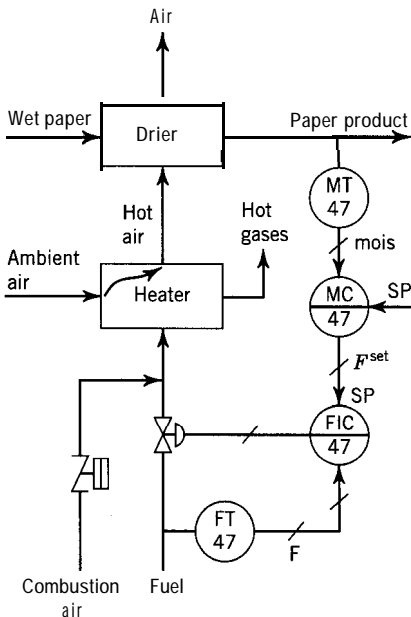
Figure P10-1 Drier for Problem 10-1.

the cascade control scheme. Show the defined transfer functions on the block diagram, and specify the action of each controller.

(c) Determine the ultimate gain and period of oscillation of the slave temperature control loop, and use these values to tune a P controller for quarter decay ratio response; then calculate the ultimate gain and period for the master moisture control loop, and the quarter decay ratio response tuning parameters for a PID moisture controller. Comparing these parameters with those you obtained in part (a), briefly comment on the advantages of cascade control for this application.

10-2. This is a real process with real data. Consider the heater shown in Fig. P10-2. In this process, the heater is actually a reactor where the catalytic reaction of C_3H_8 and steam takes place to form H_2 and CO_2. The temperature of the product gas is controlled by manipulating the flow of fuel. It was suspected that the controllers were not correctly tuned. This suspicion was confirmed by looking at the unstable response, shown in Fig. P10-3, of the reactor after an upset (a change in feed flow) entered the reactor. On the basis of this response, it was decided to re-tune the controllers. Both controllers were set in manual, and after the process reached a steady condition, shown in Fig. P10-3, the output from the flow controller to the valve was changed by 5%. The flow and temperature responses are given in Table P10-1. Interestingly, although the plant personnel were expecting a 20°F change in temperature, the actual change was about 230°F. Figure P10-3 also shows this temperature response. After the temperature reached bottom, the flow controller output was changed by + 5% to bring the temperature back to its
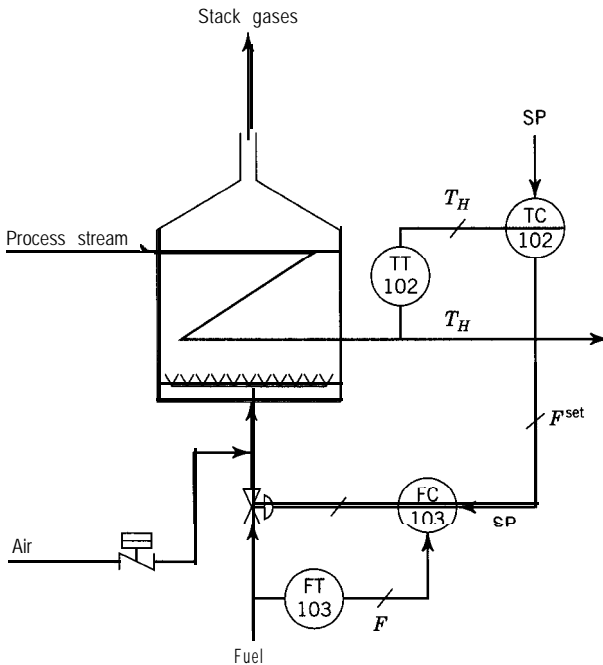
Figure P10-2 Heater for Problem 10-2.
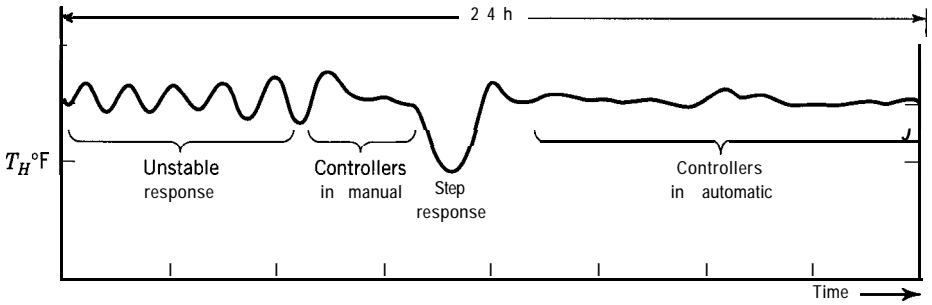
Figure P10-3 Temperature **response for Problem 10-2.**

desired operating condition. On the basis of the process response, both controllers were tuned and set in automatic. Figure P10-3 shows the response under automatic control with the new settings and for the same upset. Obtain the settings for the controllers using Austin's method, which is outlined in this chapter. The temperature transmitter range is 0 to 2000°F, and that of the flow transmitter is 0 to 24 mscfh.

10-3. Figure P10-4 shows the block diagram of a feedback control system. The control engineer in charge of the process decided that a cascade control system could improve the control performance. The proposed cascade scheme consists of mea-

Table P10-1. Test Data for Problem 10-2

Time, s	Flow, mscfh	Time, s	Flow, mscfh
0	19.46	20	17.01
5	17.51	25	16.98
10	17.32	.	
15	17.08	45	16.49

Time, min	T, °F	Time, min	T, °F	Time, mm	T, °F
0	1463	24	1393	46	1293
2	1463	26	1382	48	1287
4	1463	28	1372	50	1281
6	1462	30	1361	52	1275
8	1461	32	1351	54	1275
10	1457	34	1341	56	1263
12	1452	36	1332	58	1258
14	1444	38	1324		
16	1435	40	1316		
18	1426	42	1308		
20	1415	44	1301	94	1235
22	1405				
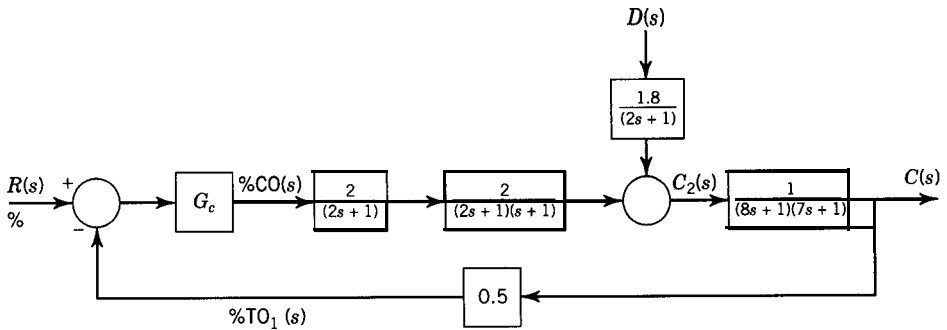
Figure PlO-4 Block diagram for Problem 10-3.

suring C,(s) with a transmitter with a gain of 0.5 and sending the signal to a controller (the slave controller). The controller shown in Fig. P10-4 becomes the master controller. Compare the stability of both systems; that is, obtain the ultimate gain and ultimate period for both systems. Assume the slave controller is proportional-only, and tune it by the Ziegler-Nichols method.

10-4. Consider the jacketed continuous stirred tank reactor (CSTR) sketched in Fig. P10-5. The following information is obtained from testing the reactor and its control system. The transfer function of the reactor temperature to the jacket temperature is a first-order lag with a gain of $0.6°C/°C$ and a time constant of 13 min. The transfer function of the jacket temperature to the coolant flow is a first-order lag with a gain of $-2.0°C/(kg/s)$ and a time constant of 2.5 min. The control valve is linear with constant pressure drop and is sized to pass 12 kg/s when fully opened. Its time constant is negligible. The reactor temperature transmitter is calibrated for a range of 50 to 100°C and has a time constant of 1 min. The jacket temperature transmitter is calibrated for a range of 0 to $100°C$, and its time constant is negligible.
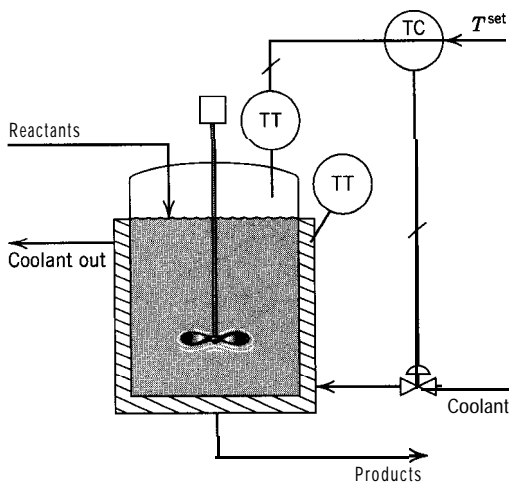
Figure P10-5 Jacketed reactor for Problem 10-4.

(a) Decide on the proper fail position of the control valve and the action of the controller for a simple feedback control loop with the reactor temperature controller manipulating the position of the coolant valve. Draw the block diagram showing all transfer functions, and write the closed-loop transfer function of the reactor temperature to its set point. Pay particular attention to the signs, which must correspond to the fail position of the valve and the controller   action.

(b) Write the characteristic equation for the single feedback loop and calculate its ultimate gain and period by direct substitution.

(c) Design a cascade control system for the reactor temperature with the jacket temperature as the intermediate process variable, specifying the action of both controllers. Draw the complete block diagram for the cascade control system showing all transfer functions and their signs.

(d) Assuming a proportional slave controller with a gain of 2%/%, write the transfer function for the jacket temperature loop and redraw the block diagram with the jacket temperature loop as a single block.

(e) Using the simplified block diagram from part (d), write the characteristic equation of the reactor temperature loop in the cascade control system and calculate the ultimate gain and period of the loop by direct substitution.

10-5. The diagram for a reactor temperature controller cascaded to a coolant flow controller is shown in Fig. P10-6. The control valve is linear with constant pressure drop and is sized for a maximum flow of 500 gpm (gallons per minute); its time constant is 0.2 min. The flow transmitter (FFT) has a range of 0 to 500 gpm and a negligible time constant. The flow controller (FC) is proportional-integral (PI) with a gain of 1.0 %CO/%TO, and the integral time is set equal to the valve time constant. The transfer function of the reactor temperature to the coolant flow is first-order with a gain of -2 °F/gpm and a time constant of 5.0 min. The temperature transmitter (TT) has a range of 160 to 200°F and a time constant of 0.5 min.
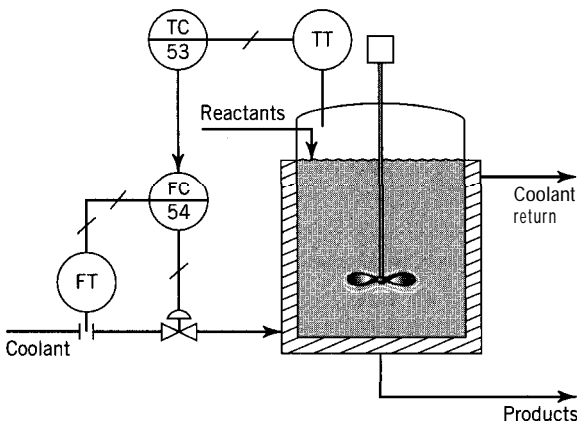
Figure P10-6 Jacketed reactor for Problem 10-5.

(a) Draw the block diagram for the cascade control loop showing the transfer function of each device. Pay particular attention to the valve fail position and the action (direct or reverse) of the flow and temperature controllers, and show the appropriate signs for each of their transfer functions.
(b) Calculate the ultimate gain and period of the temperature control loop.
(c) Repeat parts (a) and (b), assuming that the flow controller is removed and the temperature controller directly manipulates the valve. Caution: The controller action may change.

Chapter **11**

Override and Selective Control

In Chapter 10, we began our discussion of control techniques that enhance simple feedback to provide improved control performance. Specifically, Chapter 10 presented cascade control. Chapter 11 continues this presentation with two other techniques: override and selective control (override control is also sometimes referred to as constraint control). Very often these techniques are implemented for safety and optimization considerations. Also, these techniques often deal with multiple control objectives (controlled variables) and a single manipulated variable. Up to now we have dealt only with processes that have one control objective. In Chapter 13, we will look at processes with multiple objectives and multiple manipulated variables. The chapter begins with a presentation of some computing algorithms needed for these and other control techniques.

11-1 COMPUTING ALGORITHMS

Many of the control techniques presented in this and subsequent chapters require some amount of computing power. That is, many of these techniques require the multiplication, division, addition, subtraction, etc., of different signals. Several years ago, and even in some places today, all these calculations were implemented with pneumatic or electronic analog instrumentation. The devices used to implement these calculations are often referred to as computing relays. Computers, as expected, allow for a simpler, more flexible, more accurate, more reliable, and less expensive implementation of these functions.

There are several ways to implement these mathematical manipulations with computer control systems. Some manufacturers allow the control system to be programmed using a higher-level language in a fashion similar to other computers. Several other manufacturers offer software in a "subroutine-type" form referred to as computing algorithms or computing blocks. Each computing block performs a specified mathematical manipulation. To develop a control strategy, these computing blocks are linked together, the output of one block being the input to another. This "linking" procedure is referred to as "configuring" the control system, as opposed to "programming" it.

Most control system manufacturers and vendors offer both means of implementing the control strategies: programming and configuring.

Some typical calculations performed by computing relays, or computing blocks, are as follows:

1. **Addition/subtraction.** The output signal is obtained by adding and/or subtracting the input signals.
2. **Multiplication/division.** The output signal is obtained by multiplying and/or dividing the input signals.
3. **Square root.** The output signal is obtained by extracting the square root of the input signal.
4. **High/low selector.** The output signal is the highest or lowest of two or more input signals.
5. **High/low limiter.** The output signal is the input signal limited to some preset high or low limit value.
6. Function **generator.** The output signal is a function of the input signal. This function is usually approximated by a series of straight lines.
7. **Integrator.** The output signal is the time integral of the input signal. The industrial term for integrator is *totalizer*.
8. **Lead/lag.** The output signal is the response of the transfer function

$$\text{Output} = \left(\frac{\tau_{ld}s + 1}{\tau_{lg}s + 1} \right) \text{Input}$$

This calculation is often used in control schemes, such as feedforward, where dynamic compensation is required.
9. **Dead time.** The output signal is equal to a delayed input signal. This calculation is very easily done with computers, but it is extremely difficult to do with analog instrumentation.

Table 11-1.1 shows an example of computing blocks used in Honeywell's TDC Extended Controller, and Table 11-1.2 shows an example of computing blocks used in Bailey Control's controllers. These are by no means the only manufacturers; vendors of high-quality blocks abound.

There are two different ways in which field signals are handled once they enter the computer control system. The first way is to convert the number received by the computer to engineering units. That is, if a signal is read from a temperature transmitter, the number kept in memory by the computer is the temperature in degrees. The computer is given the zero and the span of the transmitter, and with this information it converts the raw signal from the field into a number indicating temperature in degrees. The second way of handling signals is not to convert them to engineering units but rather to keep them as a percentage, or fraction, of the span. That is, assume a temperature is 160°C and the sensor/transmitter measuring it is calibrated between 100°C and 200°C. The output of the transmitter in this case would be 13.6 mA, which would then be read by the control system as 60%, or 0.6. That is, 60% or 0.6 is the number kept in memory for the signal. Table 11-1.1 is an example of a system that handles signals in percentages of span. Table 1 l-l.2 is an example of a system that converts the signals to engineering units.

Table 11-1.1 Comnutinn Blocks"

The signals in this table are given in percentage of span. Here OUT = output signal and X, Y, Z = input signals.

Summer

$$OUT = K_X X + K_Y Y + K_Z Z + B_0$$

where

$$K_X, K_Y, K_Z = -9.999 \text{ to} + 9.999$$
$$B_0 = -100\% \text{ to} + 100\%$$

Multiplier/Divider

$$OUT = \frac{K_A(K_X X + B_X)(K_Y Y + B_Y)}{(K_Z Z + B_Z)} + B_0$$

where

$$K_A = 0 \text{ to} 2$$
$$K_X, K_Y, K_Z = 0.1 \text{ to } 9.999$$
$$B_X, B_Y, B_Z, B_0 = -100\% \text{ to } 100\%$$

Sum of Square Roots

$$OUT = 10K_X \sqrt{X} + 10K_Y \sqrt{Y} + 10K_Z \sqrt{Z} + B_0$$

Square Root of Product

$$OUT = (\text{Factor})K_A \sqrt{X \cdot Y \cdot Z} + B_0$$

where (for both square roots)

$$K_A, K_X, K_Y, K_Z = -9.999 \text{ to} + 9.999$$
$$B_0 = -100\% \text{ to} + 100\%$$

Factor: 0.1 when all three input signals are used; 1.0 when two input signals are used; 10.0 when only one input signal is used.

Mass Flow

$$OUT = K_A \cdot \text{ XSQRT} \cdot \sqrt{\frac{K_Y Y + B_Y}{K_Z Z + B_Z}}$$

where
XSQRT = square root of the differential pressure
(must be supplied to the algorithm)

$$K_A = 0 \text{ to} 2$$
$$K_Y, K_Z = 0.1 \text{ to } 1.00$$
$$B_Y, B_Z = 0.0\% \text{ to } 100\%$$

Lead/Lag Summer

$$OUT = \frac{K_A(sT_2 + 1)}{(sT_1 + 1)(sT_3 + 1)}(X - Y) + K_Z Z + B_0$$

Table 11-1.1 (Continued)

Lead/Lag with Multiplier

$$OUT = \frac{0.01K_A(sT_2 + 1)}{(ST_1 + 1)(sT_3 + 1)}(X - Y)(K_ZZ) + B_0$$

where (for both lead/lags)

K_A = 0.1 to 99.99
K_Z = -9.999 to +9.999
B_0 = - 100% to + 100%
T_2 = lead time constant, min
   = 0.02 to 99.99; 0 = off
T_1 = first lag time constant, min
   = 0 to 91.4; 0 = off
T_3 = second lag time constant, min
   = 0 to 91.02; 0 = off

Y can be set to zero if it is not needed.

External Ratio and Bias

$$OUT = Y(\text{effective ratio}) + (\text{effective bias})$$

where

effective ratio = $K_XX + B_X$ when there is a configured input X
   = RATIO when there is no configured input X
effective bias = $K_ZZ + B_Z$ when there is a configured input Z
   = BIAS when there is no configured input Z

where

K_X, K_Z = -9.999 to +9.999
B_X, B_Z = - 100% to + 100%
RATIO = - 9.999 to + 9.999
BIAS = - 100% to + 100%

Selector

$$OUT = \text{maximum of used inputs } X, Y, Z, M, A, C$$

$$OUT = \text{minimum of used inputs } X, Y, Z, M, A, C$$

where
X, Y, Z, M, A, C = input signals, % of range

Dead Time

$$\text{Output} = \text{Input delayed by } t_0$$

or

$$OUT = X(t - t_0)$$

[a] These are examples of the computing blocks of the Honeywell TDC microprocessor-based control system (extended controller).

Table 11-1.2 Computing Blocks"

The signals in this table are in engineering units. (〉
denotes input signals, and ( ) denotes constant val-
ues set by the user.

4-Input Summer

$$\text{OUTPUT} = \langle S1 \rangle + (S2) + (S3) + (S4)$$

2-Input Summer

$$\text{OUTPUT} = \langle S1 \rangle (S3) + \langle S2 \rangle (S4)$$

Multiply

$$\text{OUTPUT} = (S3) * [\langle S1 \rangle * \langle S2 \rangle]$$

Divide

$$\text{OUTPUT} = (S3) * [\langle S1 \rangle / \langle S2 \rangle]$$

Square Root

$$\text{OUTPUT} = (S2) * \sqrt{\langle S1 \rangle}$$

Lead/Lag

$$\text{OUTPUT} = \langle S1 \rangle * \left[\frac{T_1 s + 1}{T_2 s + 1} \right]$$

High   Select

$$\text{OUTPUT} = \text{highest of } \langle S1 \rangle, \langle S2 \rangle, \langle S3 \rangle, \langle S4 \rangle$$

Low Select

$$\text{OUTPUT} = \text{lowest of } \langle S1 \rangle, \langle S2 \rangle, \langle S3 \rangle, \langle S4 \rangle$$

Dead Time

$$\text{OUTPUT delay of } \langle S1 \rangle$$

Dead time can be either fixed or variable.

[a] These are examples of the computing algorithms provided by Bail-
ey's controllers.

11-1.1 Scaling Computing Algorithms

The second way to handle field signals necessitates an additional calculation before the
required mathematical manipulations can be performed. We will first explain the ne-
cessity and meaning of these additional calculations.

Consider a tank, shown in Fig. 1 1- 1.1, where temperature transmitters with different
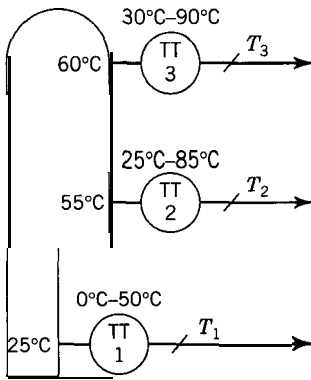
Figure 11-1.1 Tank with three temperature transmitters.

ranges measure temperatures at three different locations in the tank. The figure shows the transmitter range and the steady-state values of each temperature, which are at the mid-value of each range. It is desired to compute the average temperature in the tank. This computation is straightforward for the control system that reads each signal and converts it to engineering units: the three values are added together and divided by three. However, for the control systems that treat each signal as a percentage of the span, this simple computation would result in an answer without much significance. That is, in the example at hand, because each signal is 50% of its range, the computation result is also 50%. However, 50% of what range? How do we translate this answer into a temperature? Furthermore, note that even though every input signal is 50%, the measured temperatures are different because the ranges are different. All of this indicates that for the computation to "make sense," the range of each input signal and a chosen range for the output variable must be considered. Considering each range ensures compatibility between input and output signals.

The procedure used to consider the range of each signal is referred to as *scaling.* The scaling method presented here is very simple and applies to computers as well as to analog instrumentation. The following three steps constitute the method.

1. Write the equation, in engineering units, to be solved along with the range of each variable. Assign each variable a signal name.
2. Relate each variable in engineering units to its signal name by a scaled equation. Often, the signals are referred to as *scaled signals.*
3. Substitute the set of scaled equations into the original equation and solve for the output signal.

Let's illustrate the application of this method by using a typical example.

EXAMPLE 11-1.1

Assume it is necessary to calculate the mass flow rate of a certain gas as it flows through a process pipe, as shown in Fig 11-1.2. This figure shows an orifice installed in the
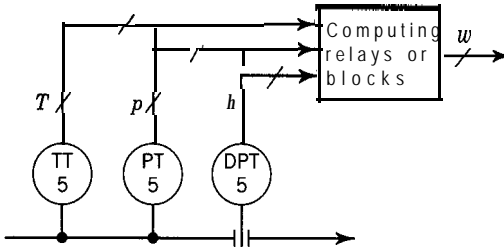
Figure 11-1.2 Gas flow through process pipe.

pipe. As presented in Appendix C, a simple equation for the calculation of mass flow through an orifice is

$$w = K[h\rho]^{1/2} \tag{11-1.1}$$

where

w = mass flow, lb/h
h = differential pressure across orifice, in. H_2O
ρ = density of gas, lb/ft^3
K = orifice coefficient, $\dfrac{\text{lb/h}}{\text{(in. } H_2O\text{-lb/ft}^3)^{1/2}}$

The density of the gas around the operating conditions is given by the following linearized equation:

$$\rho = 0.13 + 0.003(p - 30) - 0.00013(T - 500) \tag{11-1.2}$$

Thus the equation that gives the mass flow is

$$w = K\{ \ h[0.13 + 0.003(p - 30) - 0.00013(T - 500)]\}^{1/2} \tag{11-1.3}$$

For this process, the ranges of the variables are as follows:

Signal	Variable	Range	Steady State
S1	h	0- 100 in. H_2O	50 in. H_2O
S2	T	300-700°F	500°F
S3	p	0-50 psig	30 psig
S4	w	0-700 lb/h	500 lb/h

The orifice coefficient K = 196.1 $\dfrac{\text{lb/h}}{\text{(in. } H_2O\text{-lb/ft}^3)^{1/2}}$.

Equation (1 1-1 .3) and the table given-constitute step 1 of the scaling method; all of this information is known by the process engineer. Note that a signal name is assigned to each process variable.

Step 2 calls for relating each process variable to its signal name by a scaled equation. This means that as the variable varies between the low and high values of the range, the signal varies between the values of 0% and 100%. A simple equation to accomplish this scaling is

$$\text{Variable} = \frac{\text{Span}}{100} \cdot \text{Signal} + \text{Low value of range} \tag{11-1.4}$$

Applying this equation to each variable yields

$$h = \frac{100 - 0}{100} S1 + 0 = S1 \tag{11-1.5}$$

$$T = \frac{700 - 300}{100} S2 + 300 = 4\ S2 + 300 \tag{11-1.6}$$

$$p = \frac{50 - 0}{100} s3 + 0 = 0.5\ s3 \tag{11-1.7}$$

and

$$w = \frac{700 - 0}{100} S4 = 7\ S4 \tag{11-1.8}$$

Step 3 calls for substituting the scaled equations, Eqs. 1 l-l.5 through 11-1.8, into Eq. 11-1.3.

$$7\ S4 = 196.1\{S1[0.130 + 0.0030(0.5\ S3 - 30) - 0.00013(4\ S2 + 300 - 500)]\}^{1/2}$$

This equation is now solved for the output signal, $S4$, and simplified to

$$S4 = 1.085[S1(S3 - 0.35\ S2 + 44)]^{1/2} \tag{11-1.9}$$

This is the equation to be implemented with the computing algorithms, and it is referred to as the **final scaled equation.** All signals are between 0% and 100%.

The computing blocks of Table 1 1- 1.1 are used to implement Eq. 1 1- 1.9. Because there is no one single block to implement this equation, the implementation will be done by parts; the first part is to calculate the term in parentheses. We first assign this term a signal name, $S5$; that is, $S5 = S3 - 0.35\ S2 + 44$; an add/subtract block can be used to implement this summation. Note that $S5$ will be the output signal from the computing block, so it must follow the same requirement that any other signal must follow: it must fit in the range from 0% to 100%. An analysis of the equation reveals that, depending on the operating conditions, this is not the case. For example, if $p = 40$ psig (S3 = 80%) and $T = 500°F$ (S2 = 50%), then $S5$ is equal to 107.5%, which cannot be. This condition is easily avoided by noting that the worst condition is when $p = 50$ psig (S3 = 100%) and $T = 300°F$ (S2 = 0%), which will result in $S5 =$

144%. Thus, what can be done to avoid a number greater than 100% is to divide the $S5$ equation by 1.44 or a greater number. That is,

$$S4 = 1.085\left[1.44\ S1\left(\frac{S3}{1.44} - \frac{0.35}{1.44}S2 + \frac{44}{1.44}\right)\right]^{1/2}$$

or

$$S4 = 1.302[S1(0.694\ S3 - 0.243\ S2 + 30.55)]^{1/2} \qquad \textbf{(11-1.10)}$$

Then

$$S5 = 0.694\ S3 - 0.243\ S2 + 30.55$$

The add/subtract block shown in Table 1 1- 1.1 is

$$\text{OUT} = K_X X + K_Y Y + K_Z Z + B_0$$

Let X = S3, Y = S2, and OUT = $S5$; the Z input is not used. That is, signal S3, the signal from the pressure transmitter, is connected to the X input. Signal S2, the signal from the temperature transmitter, is connected to the Y input. Matching the last two equations, we get

$$K_X = 0.694 \qquad K_Y = -0.243 \qquad B_0 = 30.55\%$$

The equation now is

$$S4 = 1.302[S1*S5]^{1/2}$$

This calculation is then implemented with a "square root of product" block.

$$\text{OUT} = (\text{Factor})K_A \sqrt{X \cdot Y \cdot Z} + B_0$$

Let X = $S1$, Y = $S5$, and OUT = S4; the Z input is not used. That is, signal $S1$, the signal from the differential pressure transmitter, is connected to the X input, and signal S5, the output signal from the previous block, is connected to the Y input. Because only two input signals are used, the term Factor becomes 1 automatically. Then, matching the last two equations, we get

$$K_A = 1.302 \qquad B_0 = 0\%$$

Figure 1 1-l.3 shows the diagram of the required blocks. The output signal from $FY5B$, the square root extractor, is linearly related to the gas mass flow. As the mass flow varies between 0 and 700 lb/h, the signal varies between 0% and 100%. This signal can now be used to perform any control or recording function.

The application of the method shown is regardless of the type of signal. It is not necessary to specify an electrical, pneumatic, or digital signal when scaling. The method is applied in the same fashion for all three of these signals and for any other, which makes it very powerful.
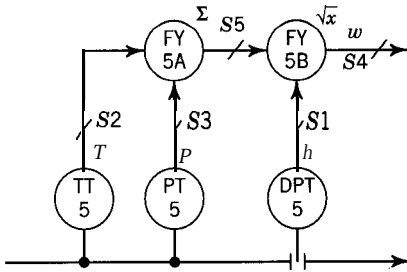
Figure 11-1.3 Gas mass flow calculation.

As previously mentioned, some control systems treat the signals as fraction of span, instead of percentage of span. That is, the signals are between 0 and 1, instead of between 0% and 100%. In these cases the scaling procedure is exactly the same as before, except that the scaling equation is given by

$$\text{Variable} = \text{Span} \cdot \text{Signal} + \text{Low value of range}$$

Example 1 1-l. 1 showed a simple case, but sometimes a significant amount of calculation is involved before we get to the final scaled equation. Thus the possibility of miscalculations is realistic. To check for these miscalculations, it is important to verify the scaled equation, Eq. 1 1- 1.10, before implementing it. This is easily done using the steady-state values shown in the table developed in step 1. The steady-state values for the signals, on the basis of the scaled equations, are

$$\overline{S1} = 50\% \qquad \overline{S2} = 50\% \qquad \overline{S3} = 60\% \qquad \overline{S4} = 71.4\%$$

Substituting $\overline{S1}, \overline{S2},$ and $\overline{S3}$ into Eq. 1 1-l. 10 yields

$$\overline{S4} = 1.302\{50[0.694(60) - 0.243(50) + 30.5]\}^{1/2} = 71.3\%$$

The difference between the two values of $\overline{S4}$, the expected and the calculated, is small (0.14%) and is due mainly to truncation errors. Certainly, in large-volume processes even this error, integrated over time, may become significant.

This section has presented an introduction to different computing blocks available for implementing the control strategies discussed in this and **the** following chapters. We have paid particular attention to those systems that require scaling of these algorithms. The scaling procedure shown is simple to apply. For systems that convert the signals to values in engineering units, scaling is not necessary.

11-1.2 Physical Significance of Signals

During the presentation of the types of field signals in Chapter 1, and in the discussion in Section 10-1, we noted that signals are used by the different instruments to convey information and that, therefore, every signal has "physical significance." That is, every signal used in the control scheme has some "meaning." This fact was also mentioned

in Chapter 10 on cascade control. It is now important to stress again this fact as we embark on the design of complex strategies to improve the control performance.

As already mentioned in this chapter, the new strategies frequently require the manipulation of signals in order to calculate control variables or set points or to decide on control actions. To perform these calculations correctly, it is most important to understand the significance of the signals.

Very often the first step in the design of a control strategy is to give a signal, which is sometimes referred to as the "master" signal, a physical significance. Then, on the basis of the given significance the strategy is designed. This presentation may seem somewhat abstract right now, but as we continue in the next chapters with the study of different control strategies, the presentation will become clear and realistic.

To help keep all of the information in order, and to understand what the calculations are doing, we will indicate next to each signal its significance and direction of information flow, just as we did in Figs. 11-1.1, 11-1.2, and 11-1.3. This practice is not common in industry, but we believe it facilitates understanding the subject.

11-2   OVERRIDE, OR CONSTRAINT, CONTROL

Override control, or *constraint control,* is a powerful yet simple control strategy generally used as (1) a protective strategy to maintain process variables within limits that must be enforced to ensure the safety of personnel and equipment, and product quality and (2) an optimization strategy that permits smooth transition between controllers to obtain maximum benefit. As a protective strategy, override control is not so drastic as interlock control. Interlock controls are used primarily to protect against equipment malfunction. When a malfunction is detected, the interlock system usually shuts down the process. Interlock systems are not presented, but see Becker and Hill (1979) and Becker (1979). Two examples of constraint control are now presented to demonstrate the concept and the implementation of the strategy.

EXAMPLE 11-2.1

Consider the process shown in Fig. 11-2.1. A hot saturated liquid enters a tank and from there is pumped under flow control back to the process. Under normal operation, the level in the tank is at height h_1. If under any circumstances the liquid level drops to height h_2, the liquid will not have enough net positive suction head (NPSH), and cavitation at the pump will result. It is therefore necessary to design a control scheme that avoids this condition. This new control scheme is shown in Fig. 11-2.2.

The level in the tank is now measured and controlled. It is important to note the action of the controllers and final control element. The variable-speed pump is such that as the energy (current in this case) input to it increases, it pumps more liquid. Therefore, the FC50 is a reverse-acting controller, whereas the LC50 is a direct-acting controller. The output of each controller is connected to a low selector, LS50, and the signal from this selector goes to the pump.

Under normal operating conditions the level is at h_1, which is above the set point to the level controller. Consequently, the controller will try to speed up the pump as much as possible, increasing its output to 100%. The output of the flow controller, under
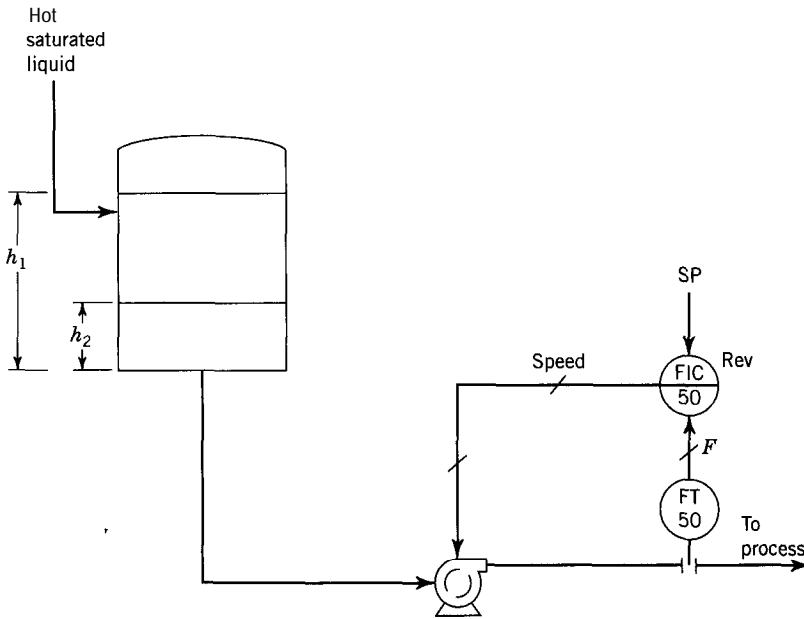
Figure 11-2.1 Tank and flow control loop.

normal conditions, may be 75%, and thus, the low selector switch selects this signal to manipulate the pump speed. This is the desired operating condition.

Let us now suppose that the flow of hot saturated liquid slows down and the level in the tank starts to drop. As soon as the level drops below the set point on the level controller, this controller will try to slow down the pump by reducing its output. When the level controller's output drops below the output of the flow controller, the low selector selects the output of the level controller to manipulate the pump. It can be said that the level controller "overrides" the flow controller.

When the flow of hot liquid returns to its normal value, and the level increases above the set point, the level controller increases its output to speed up the pump. Once the output from the level controller increases above the output from the flow controller, the low selector selects the flow controller, and the operation returns to its normal condition.

An important consideration in designing an override control system is that of reset windup protection on any controller that has integration. The output of the controller not selected must stop at 100%, not at a higher value, or at 0%, not at a lower value (actually, most computer control systems' output can go between 105% and − 5%, or as high as 110% and as low as − 10%). An even more desirable operation is the one that, if the selected controller output is 75%, forces the nonselected output to be close to 75%, not even 100%. That is, it forces the nonselected controller output to be close to the selected output. This desirable operation is easily accomplished using the *reset feedback* technique presented in Chapter 5. Figure 11-2.2 shows the reset feedback (RFB) connections (dotted lines) to both controllers. In this case the reset feedback signal to the controller(s) comes from an external computation, a low selector, not from the controller itself as shown in Fig. 5-3.17; sometimes we refer to this signal as external reset feedback. The output from the low selector is the one that goes to the pump; it is used as the reset feedback signal to the controllers.
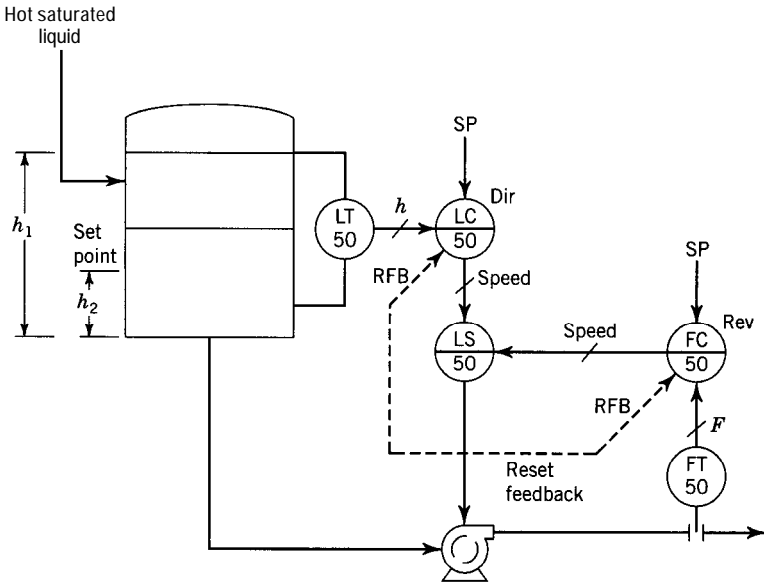
Figure **11-2.2** Override control scheme.

Figure 11-2.3 shows an schematic of the two controllers with the external reset feedback. To further examine how this system works, consider that at steady state the flow controller outputs a 75% signal to maintain its set point, and the level in the tank is above its set point. In this case the level controller increases its output, to speed up the pump, and control at its set point. Thus the low selector selects the 75% signal from the flow controller; this is the normal operating condition. The output signal from the selector is the RFB signal to the controllers, and the corresponding M_I signal will also be 75%. At this steady state, the error in the flow controller is zero, and the proportional calculation of this controller is also zero. Because the level in the tank is above the set point, the error in the level controller is positive (direct-acting) and the proportional calculation will have a certain output, depending on the error and controller gain, say 10%. The M_I signal is 75%, so the output from the level controller to the low selector is 85%. Now suppose that the input flow to the tank decreases and the level in the tank
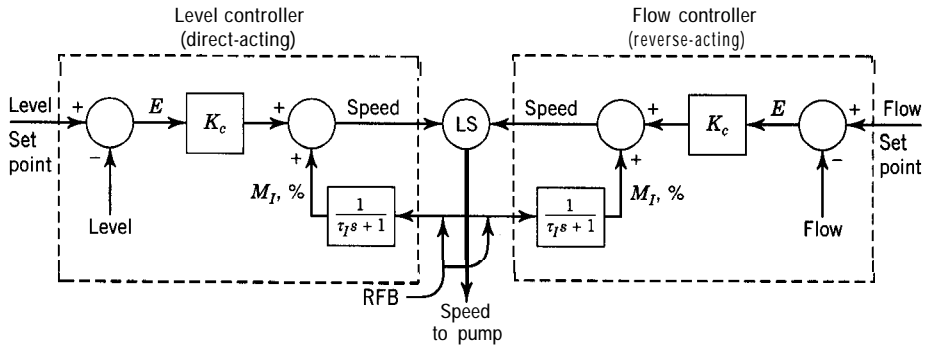
Figure **11-2.3** Controllers with external reset feedback (RFB).

starts to drop. As this happens, the proportional calculation in the level controller also starts to decrease from 10% down, and the controller output from 85% down. As the level in the tank drops below the set point, the error in the controller, and the proportional calculation, become negative, which results in an output less than 75% — say 74%. At this moment the selector selects this signal, and the level controller overrides the flow controller and sends it to the pump to slow it down; this new value is also the RFB signal to the controllers. As the pump slows down to avoid low level in the tank, the error in the flow controller becomes positive (reverse-acting), and the proportional calculation increases to increase the flow and correct for the error. However, the low selector will not allow this particular corrective action; it is more important to avoid a low level in the tank. Note that the output from the flow controller will then be equal to the output from the low selector plus its own proportional calculation.

Most controllers offer this external reset feedback capability. To summarize, this capability, which is also sometimes referred to as output tracking, allows the controller not selected to override the controller selected as soon as its error changes sign. More than two controllers can provide signals to a selector and have RFB signals; this is shown in the following example.

EXAMPLE 11-2.2

A fired heater, or furnace, is another common process that requires the implementation of constraint control. Figure 11-2.4 shows a heater with temperature control manipu-
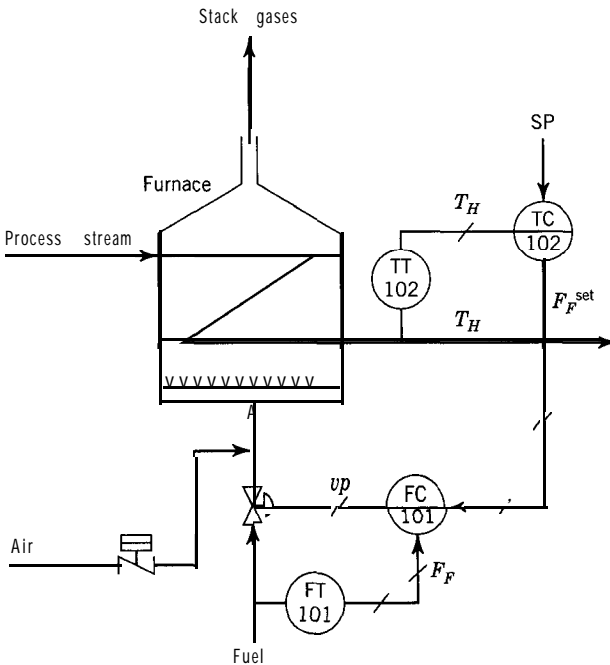
Figure 11-2.4 Heater temperature control.

lating the gas fuel flow. The manipulation of the combustion air has been omitted to simplify the diagram, but it will be discussed in detail in Chapter 12. There are several conditions in this heater that can prove quite hazardous. These conditions include (1) higher fuel pressure than can sustain a stable flame and (2) higher stack, or tube, temperature than the equipment can safely handle. If either of these two conditions exists, the gas fuel flow must decrease to avoid the unsafe condition; at this moment, the temperature control is certainly not so important as the safety of the operation. Only when such conditions disappear is it permissible to return to straight temperature control.

Figure 11-2.5 shows a constraint control strategy to guard against the unsafe conditions we have described. The gas fuel pressure is usually below the set point to $PC103$, so the controller will try to raise the set point to the fuel flow controller. Usually, the stack temperature is also below the set point to $TC101$, so the controller will also try to raise the set point to the fuel flow controller. Thus, under normal conditions, the exit heater temperature controller is the controller selected by the low selector, because its output will be the lowest of the three controllers. Only when one of the unsafe conditions exists is TC102 "overridden" by one of the other controllers.
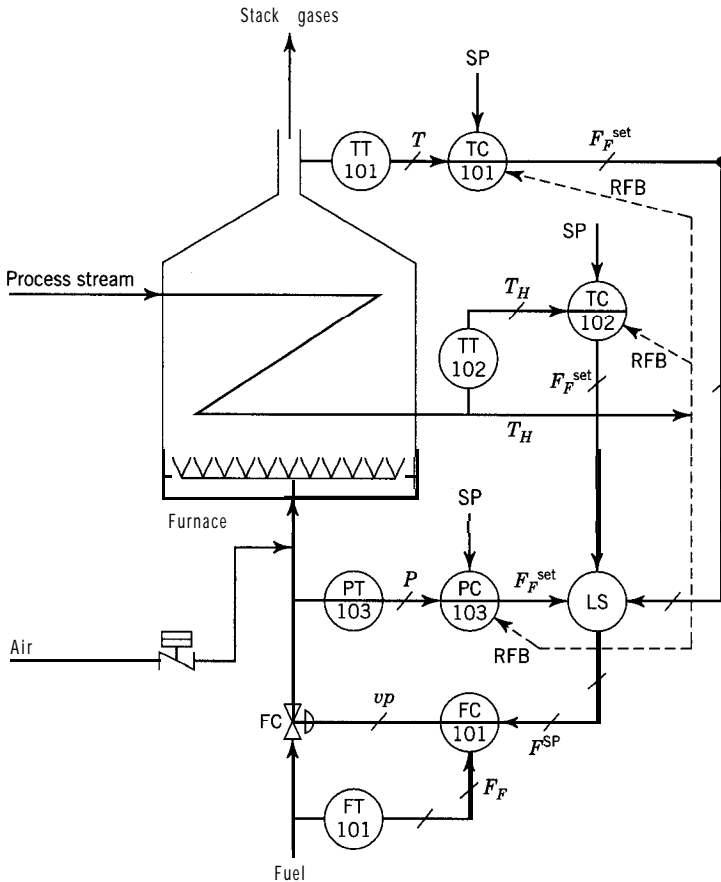
Figure   11-2.5   Heater   temperature   control-constraint   control.

As explained in Example 11-2.1, it is important to prevent windup of the controllers that are not selected. Thus the control system must be configured to provide external reset feedback. This is shown by the dotted lines in the figure.

The constraint control scheme shown in Fig. 11-2.5 contains a possible safety difficulty. If at any time the operating personnel were to set the flow controller FC101 in the automatic or the manual mode (off cascade mode), then the safety provided by TC101 and PC103 would not be in effect. This would result in an unsafe and unacceptable operating condition. You may want to think how to design a new constraint control to permit the operating personnel to set the flow controller in automatic or manual and still have the safety provided by TC102 and PC103 in effect. This is a problem at the end of this chapter.

The introduction to this section mentioned that override control is commonly used as a protective scheme, and Examples 11-2.1 and 11-2.2 presented two such applications. As soon as the process returns to normal operating conditions, the override scheme returns automatically to its normal operating status. The two examples presented show multiple control objectives (controlled variables) with a single manipulated variable; however, only one objective is enforced at a time.

11-3   SELECTIVE   CONTROL

Selective control is another interesting control scheme used for safety considerations and process optimization. Two examples are presented to show its principles and implementation.

EXAMPLE 11-3.1

Consider the exothermic catalytic plug flow reactor shown in Fig. 11-3.1. The figure shows the reactor temperature control and the temperature profile along the reactor with its typical "hot spot." The sensor providing the temperature measurement should be located at the hot spot. As the catalyst in the reactor ages, or as conditions change, the hot spot moves. It is desired to design a control scheme so that its measured variable "moves" as the hot spot moves. A control strategy that accomplishes the desired specifications is shown in Fig. 11-3.2. The high selector in this scheme selects the transmitter with the highest output, and thus the controlled variable is always the highest, or closest to the highest, temperature.

In implementing this control strategy, an important consideration is that all temperature transmitters must have the same range so that their output signals can be compared on the same basis. Another consideration that may be important is installing some kind of indication of which transmitter is giving the highest signal. If the hot spot moves past the last transmitter, TT11D, this may be an indication that it is time either to regenerate or to change the catalyst. The length of reactor left for the reaction is probably not enough to obtain the desired conversion.
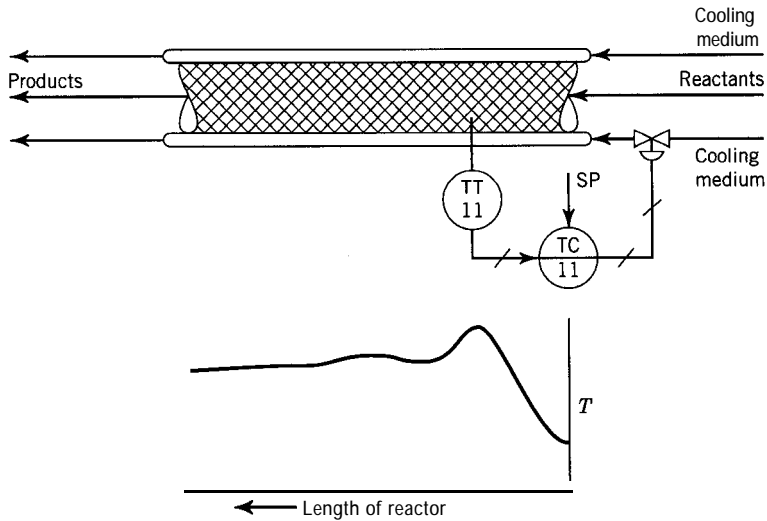
Figure   11-3.1   Temperature   control   of   a   plug   flow   reactor.

EXAMPLE 11-3.2

An instructive and realistic process wherein selective control can improve the operation
is shown in Fig. 11-3.3. A furnace heats a heat transfer oil to provide an energy source
to several process units. Each individual unit manipulates the flow of oil required to
maintain its controlled variable at set point. In addition, the outlet oil temperature from
the furnace is controlled by manipulating the fuel flow. A bypass control loop, DPC101,
is provided.
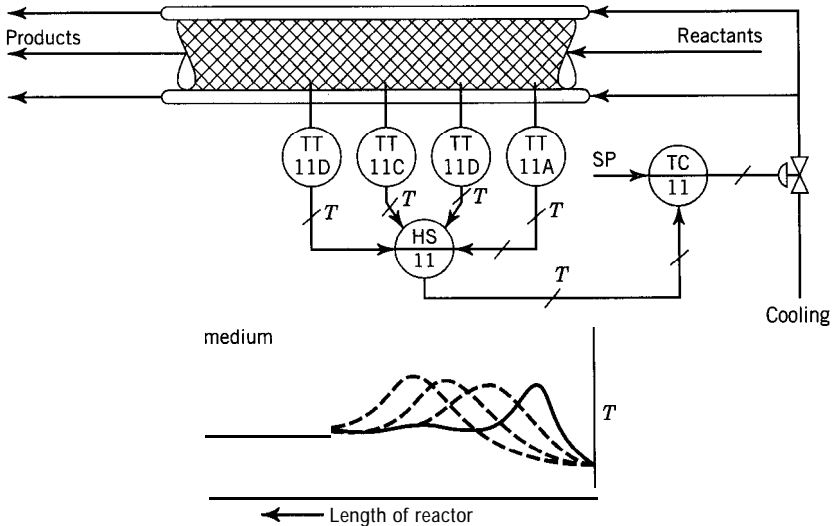
Figure   11-3.2   Selective   control   for   a   plug   flow   reactor.
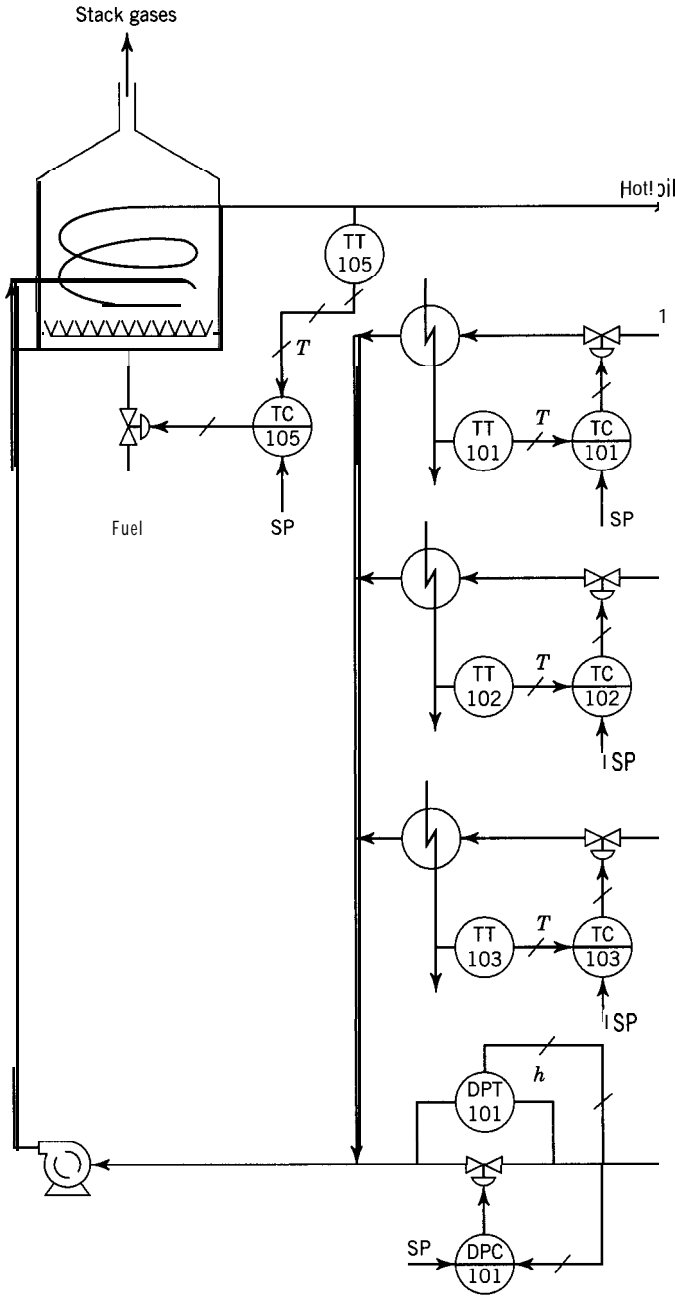
Figure 11-3.3 Hot oil system.
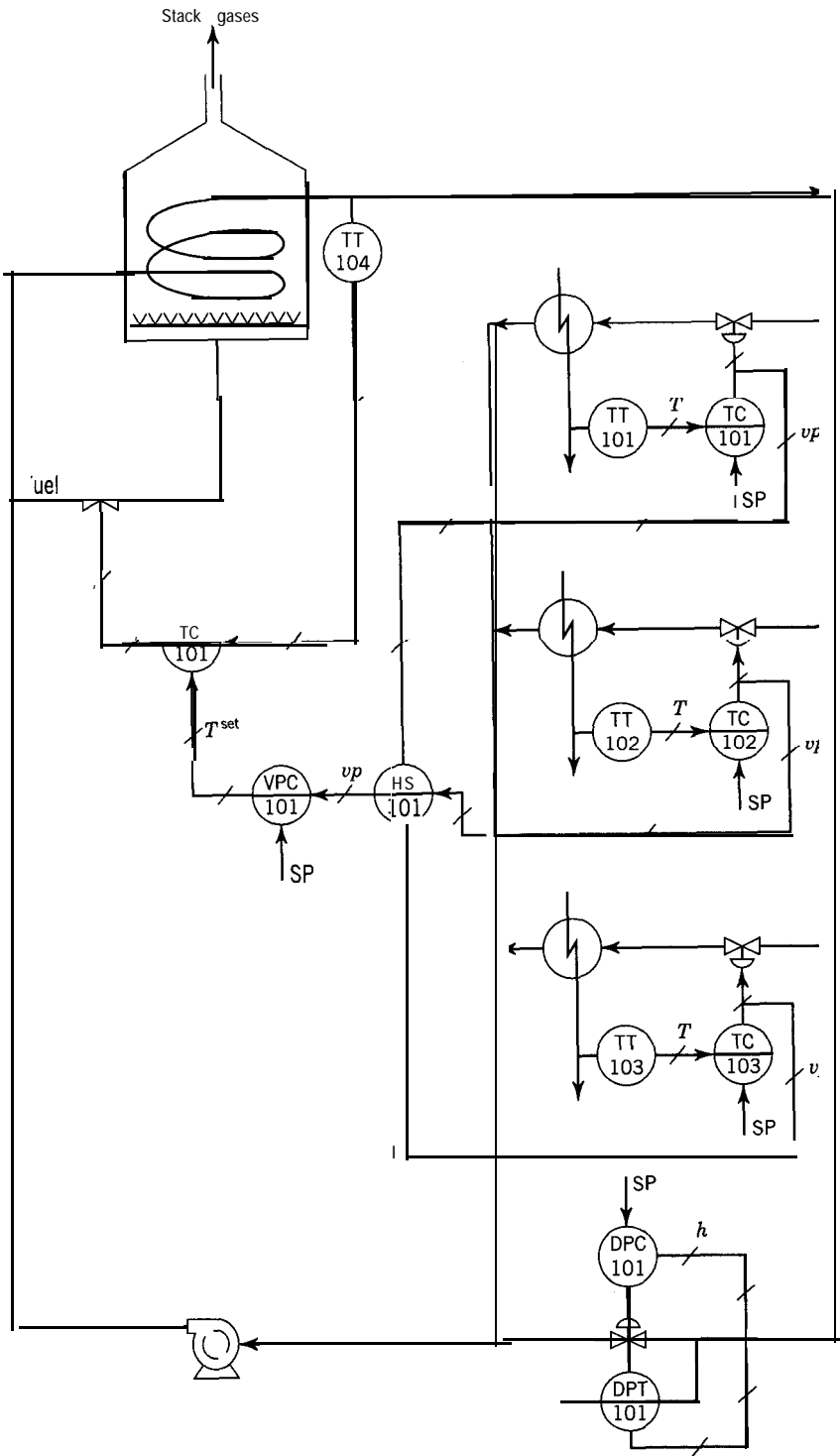
Figure 11-3.4 Selective control for hot oil system.

Suppose it is noticed that the control valve in each unit is not open very much— that is, TV101 is 20% open, TV102 is 15% open, and TV103 is 30% open. This indicates that the hot oil temperature provided by the furnace may be higher than required by the users. Consequently, not much oil flow is necessary, and much of it will bypass the users. This situation is energy-inefficient because a large quantity of fuel must be burned to obtain a high oil temperature. Also, a significant amount of the energy provided by the fuel is lost to the surroundings in the piping system and through the stack gases.

A more efficient operation is the one that maintains the oil leaving the furnace at a temperature just hot enough to provide the necessary energy to the users with hardly any flow through the bypass valve. In this case most of the temperature control valves would be open most of the time. Figure 11-3.4 shows a selective control strategy that provides this type of operation. The strategy first selects the most open valve using a high selector, TY101. The valve position controller, VPC101, controls the selected valve position-at, say, at 90% open-by manipulating the set point of the furnace temperature controller. Thus this strategy ensures that the oil temperature from the furnace is just "hot enough."

Note that because the most open valve is selected by comparing the signals to the valves, they should all have the same characteristics.

The selective control strategy shows again that with a bit of logic, a process operation can be significantly improved.

11-4 SUMMARY

This chapter introduced some computing algorithms provided by manufacturers. An explanation of the need for scaling was given, and the method for scaling was presented in detail. A brief discussion of the significance of field signals was included.

The chapter also presented the concepts of override and selective control and their applications. These techniques provide a realistic and simple way to improve process safety, product quality, and process operation.

REFERENCES

1. Becker, J. V. 1979. "Designing Safe Interlock Systems." *Chemical Engineering Magazine,* October 15.
2. Becker, J. V., and R. Hill. 1979. "Fundamentals of Interlock Systems." *Chemical Engineering Magazine,* October 15.

PROBLEMS

11-1. The tank shown in Fig. 11-1.1 contains temperature transmitters that measure temperatures at three different locations in the tank. Specify the instrumentation required to calculate the average temperature in the tank. Use Table 11-1.1 and specify the required scale factors.

11-2. *Mass Flow Computer.* A standard mass flow computer calculates the mass flow of a gas from the orifice equation

$$w(t) = K_0\sqrt{Mp(t)h(t)/RT(t)}$$

where $w(t)$ is the mass flow, K_0 is the orifice coefficient, M is the molecular weight of the gas, $p(t)$ is the absolute pressure, R is the ideal gas constant, $T(t)$ is the absolute temperature of the gas, and $h(t)$ is the differential pressure across the orifice. For this application,

$$K_0 = 200 \ (\text{lb/h})/(\text{in. } H_2O\text{-lb/ft}^3)^{1/2}$$

$$M = 40 \ \text{lb/lbmol}$$

$$R = 10.73 \ \text{psia-ft}^3/\text{lbmol-}°R$$

with the following transmitter ranges:

$$p(t) = 0 \text{ to } 50 \text{ psig} \qquad T(t) = 100°F \text{ to } 200°F$$

$$h(t) = 0 \text{ to } 100 \text{ in. } H_2O \qquad w(t) = 0 \text{ to } 1500 \text{ lb/h}$$

(a) Convert the equation to an equation in the scaled signals, assuming that all signals are in percentage of range.
(b) Draw the instrumentation diagram using the computing blocks from Table 1 1- 1.1. Specify the scale factors.

11-3. Consider the piping system shown in Fig. Pll-1, in which natural gas flows into a process. It is necessary to meter the total flow of gas that enters the process
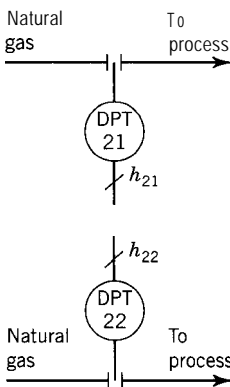
Figure Pll-1 Natural gas pipes for Problem 11-3.

and totalize it so that every 24 h the total amount is known. The flow rate through each meter, as differential pressure, is given by

$$\text{DPT21:}\ \ Q_{21} = 44.736,\ \text{mscfh}$$

$$\text{DPT22:}\ \ Q_{22} = 48.106,\ \text{mscfh}$$

where h is the differential pressure in inches of H_2O. Both transmitters have a range of 0 to 100 in. H_2O. Specify the instrumentation required to calculate the instantaneous total flow rate into the process, using Table 11-1.1. Determine the scale factors. (mscfh = 1000 ft^3/h at standard temperature and pressure.)

11-4. *Heat Rate Computer. The* heat exchanger shown in Fig. 10-4.1 heats a process fluid by condensing steam. A control scheme calls for controlling the heat transferred to the fluid. This heat transfer is calculated using the equation

$$Q = F\rho c_p(T - T_1)$$

The following information is known:

Variable	Range	Steady State
F	0-50 gpm	30 gpm
T	50°F– 120°F	80°F
T_1	25°F–60°F	50°F
Q	o-?	?

The density (ρ = 5.62 lb/gal) and heat capacity (c_p = 0.60 Btu/lb) are assumed to be constant. Using Table 1 1-1. 1, specify the instrumentation required to calculate the energy transferred. Determine the scale factors.

11-5. Figure Pl 1-2 shows the reflux to the top of a distillation column. The "internal reflux computer" computes the set point, L_E^{set}, of the external reflux flow controller so as to maintain the internal reflux L_I, at some desired value, L_I^{set}. The internal reflux is greater than the external reflux because of the condensation of
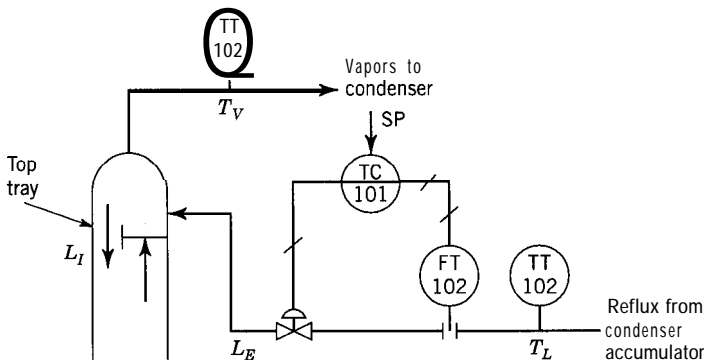
Figure Pll-2 Distillation column for Problem 11-5.

vapors on the top tray, which is required to bring the subcooled reflux at T_L up to its bubble point, T_V. An energy balance on the top tray yields the working equation

$$(L_I - L_E)\lambda = L_E C_{PL}(T_V - T_L)$$

Show all of the instrumentation required for the internal reflux computer and compute the coefficients, using Table 11-1.2. For this process, the heatcapacity of the liquid and the latent heat can be assumed constants at values of $C_{PL} = 0.76$ Btu/lb-°F and $\lambda = 285$ Btu/lb. Other design specifications are as follows:

Transmitter	Range	Normal Value
FT102 (L_E)	0-5000 lb/h	3000 lb/h
TT102 (T_L)	100°F–300°F	195°F
TT101 (T,)	150°F–250°F	205°F

11-6. Figure Pll-3 shows a system designed to filter an oil before processing. The oil enters a header in which the pressure is controlled, for safe operation, by manipulating the inlet valve. From the header, the oil is distributed to four filters. The filters consist of a shell with tubes inside similar to heat exchangers. The tube wall is the filter medium through which the oil must flow to be filtered. The oil enters the shell and flows through the medium into the tubes. As time passes, the filter starts to build up a cake, and consequently, the oil pressure required for flow increases. If the pressure increases too much, then the walls may collapse. Thus at some point, the filter must be taken out and cleaned. Under normal conditions, the total oil flow can be handled by three filters.

(a) Design a flow control system to set the total oil flow through the system.
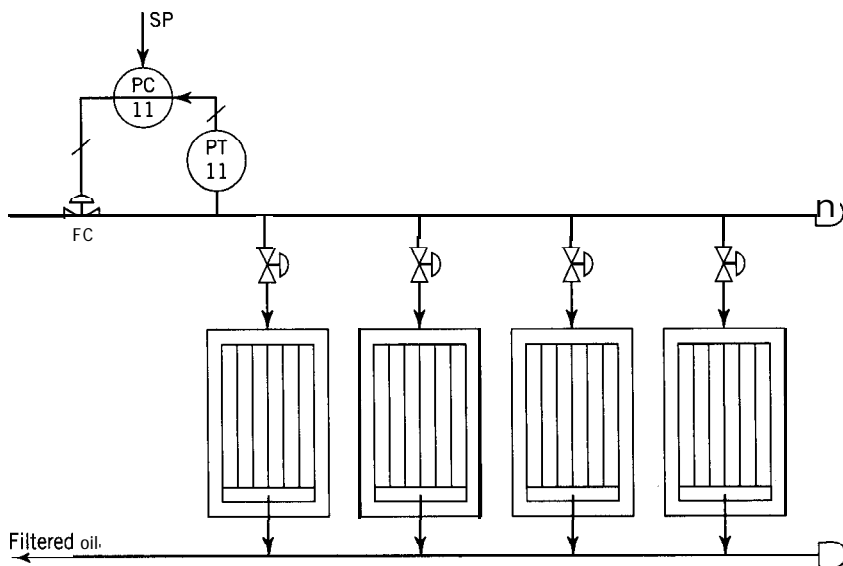
Figure Pll-3 Filters for Problem 11-6.

(b) Design a control system so that as the oil pressure drop in each filter increases above some predetermined value, the oil flow to that filter starts to decrease. Once the feed valve is 10% open, it is time to close down the filter for cleaning. The total oil flow through the system must still be maintained.

11-7. Figure Pl l-4 shows a process often found in chemical plants. R-101 is a reactor where high-pressure gas is generated. It is necessary to transfer the gas to a low-pressure vessel, V-102, at about 50 psig. As an energy saving measure, the gas pressure is dropped across a power recovery turbine, T-102. The work produced in T-102 can be used to drive a compressor, C-102. However, the work produced in T-102 is usually not enough to run C-102, so a steam turbine, T-103, is connected in series with T-102 to provide the necessary work. The figure shows the control systems to control the pressure in R-101 and the pressure of the gas leaving the compressor. ST-16 is a speed transmitter connected to the turbine's shaft; SC-16 is a controller controlling the shaft rotational speed. Out of R-101 there is a line that by-passes T-102 and goes directly to V-102. This line is used in case T-102 is down or in case any emergency develops and it is necessary to relieve the pressure in R-101 quickly. The set point to PC-14 is 500 psig, whereas that to PC-15 is 510 psig. A condition occurs when the
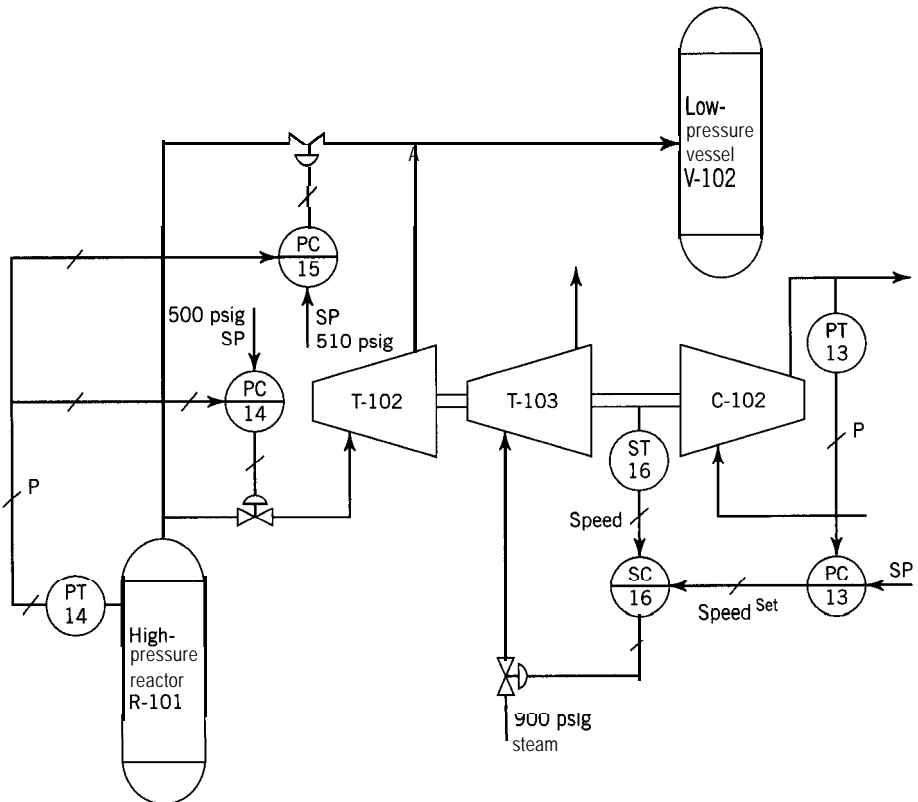
Figure Pll-4 Process for Problem 11-7.

gas produced in R-101 increases significantly, thus increasing the pressure. In this case PC-14 opens the valve to T-102 wider to relieve the pressure. When this occurs, the steam valve to T-103 eventually backs off to control the pressure of the compressed gas. It has been determined that if the steam valve is less than 10% open, then severe mechanical damage can occur in T-103. Design a control scheme to avoid this condition, and specify the fail-safe action of all the valves.

11-8. In Example 11-2.2 the constraint control scheme shown in Fig. 11-2.4 was presented and discussed. In the discussion of the scheme, it was mentioned that "if at any time the operating personnel were to set the flow controller FClOl in the automatic or the manual mode (off cascade mode), then the safety provided by TClOl and PC103 would not be in effect. This would result in an unsafe and unacceptable operating condition." Modify the control scheme shown such that even when FClOl is off cascade, controllers TClOl and PC103 still provide the necessary safety override.

11-9. Consider the turbine/compressor process shown in Fig. Pl l-5. The motive force for the turbine, T-201, is a high-pressure gas, and the compressor, C-103, compresses a refrigerant gas. The operator sets the valve position of the high-pressure gas valve, which in turn results in a certain compressor speed. A lag unit is used to avoid sudden changes in the high-pressure valve position. Under normal operating conditions, the valve should respond to the operator's set value. However, there are some special conditions that the control system must guard against.

- Under normal conditions, the pressure in the refrigerant gas line is about 15 psig. However, during start-up and under other circumstances, the pressure in the line tends to drop below 8 psig, which condition is dangerous to the compressor. In this case the compressor velocity must be reduced to pull in less gas, thus increasing the pressure in the line. The lowest safe pressure in the refrigerant line is 8 psig.
- Because of mechanical difficulties, the compressor velocity must not under operation exceed 95% of its maximum rated velocity. Also, it must not drop below 50% of its maximum velocity.
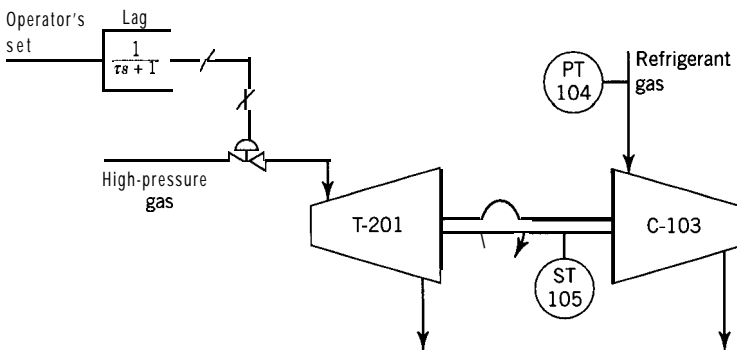
Figure   Pll-5   Turbine/compressor   process   for   Problem   11-9.
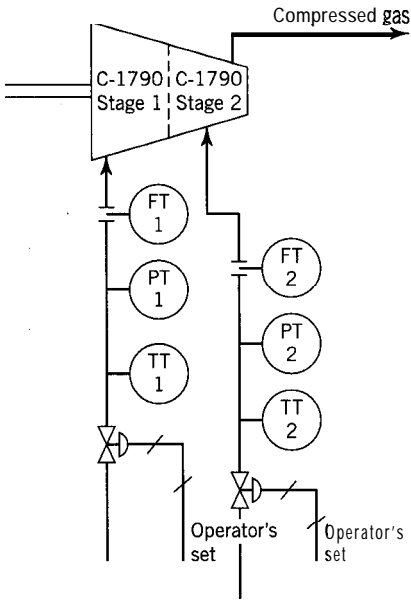
Figure Pll-6 Compressor for Problem
11-10.

Design, to manipulate the high-pressure gas valve, a control strategy that does
not violate the foregoing constraints.

11-10. Consider the compressor shown in Fig. Pll-6. This two-stage compressor has
two different suction points. In each suction line there is a volumetric flow meter
calibrated at 0°C and 1 atm, a pressure transmitter, and a temperature transmit-
ter. An important consideration in the control of the compressor is to avoid the
surge condition, Figure Pl l-7 shows a curve indicating the minimum inlet flow,
in ACFM (actual ft^3/min) required, for a given inlet pressure, to avoid surge.
Each stage can go into surge independently. Under normal operating conditions,
the operator sets the position to each suction valve. However, for safety reasons,
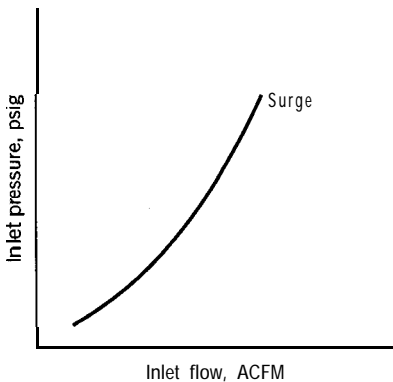
Figure Pll-7 Surge curve for Problem
11-10.

the operator must not close the valves below the surge limit. Design a control strategy to avoid closing the valves below the surge limit. Also as a result of safety considerations, although it is permissible to open the valves very fast, closing the valves must be done slowly. Design this constraint into your previous design.

Chapter 12

Ratio and Feedforward Control

This chapter continues the study of advanced control techniques that enhance process control. Specifically, the chapter presents the principles and practice of ratio and feed-forward control. These control strategies usually involve multiple measurements but a single control objective and a single manipulated variable. Without any doubt, the use of computers has been responsible for an increase in the use of both techniques.

12-1 RATIO CONTROL

A commonly used process control technique is ratio control. *Ratio control* is a strategy wherein one variable is manipulated to keep it as a ratio or proportion of another. This section presents two industrial examples to illustrate its meaning and implementation. The first example is a simple one, but it clearly demonstrates the need for ratio control.

Assume that it is required to blend two liquid streams, A and B, in some proportion, or ratio, *R;* the process is shown in Fig. 12-1.1. That is,

$$R = \frac{F_B}{F_A}$$

where F_A and F_B are the flow rates of streams A and B, respectively.

   An easy way of accomplishing this task is shown in the figure. Each stream is controlled by a flow loop in which the set points to the controllers are set such that the
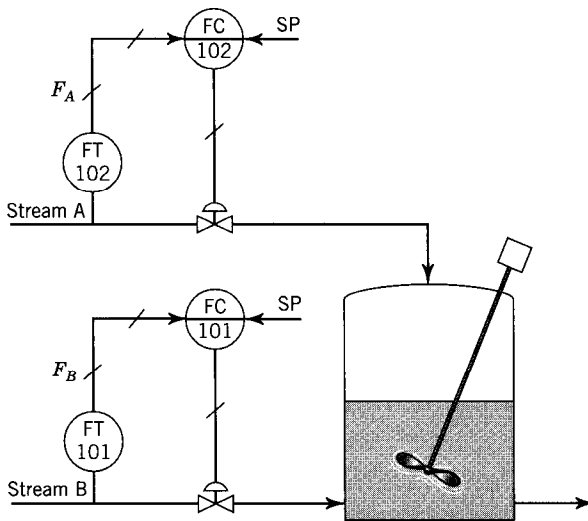
Figure 12-1.1 Control of blending of two liquid streams.

liquids are blended in the correct ratio. However, suppose now that stream A cannot be controlled but only measured. The flow rate of this stream, often referred to as "wild flow," is usually manipulated to control something else, such as level or temperature, upstream. The controlling task is now more difficult. Somehow the flow rate of stream B must vary, as the flow rate of stream A varies, to maintain the blending in the correct ratio. Two possible ratio control schemes are shown in Fig. 12-1.2.
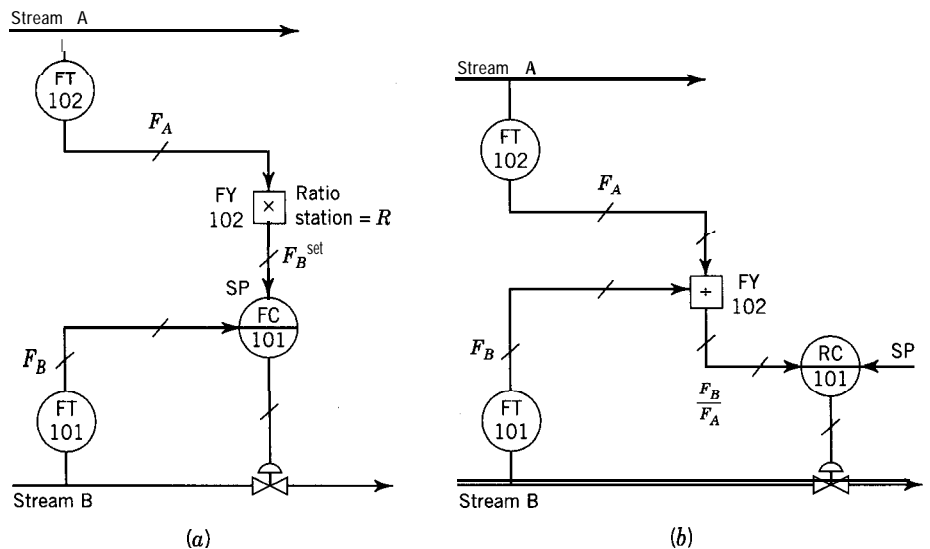
Figure 12-1.2 Ratio control of blending system.

The first scheme, shown in Fig. 12-1.2a, consists of measuring the wild flow and multiplying it by the desired ratio, in FY102, to obtain the required flow rate of stream B; that is, $F_B = RF$. The output of the multiplier, or ratio station, FY 102, is the required flow rate of stream B, and it is used as the set point to the flow controller of stream B, FC101. (Recall that depending on the control system being used, the calculation done in the ratio station may have to be scaled as shown in Chapter 11.) Thus as the flow rate of stream A varies, the set point to the flow controller of stream B will vary accordingly to maintain both streams at the required ratio. If a new ratio between the two streams is required, the new ratio is set in the multiplier. The set point to the flow controller of stream B is set from a computation, not locally.

The second ratio control scheme, shown in Fig. 12-1.2b, consists of measuring both streams and dividing them, in FY102, to obtain the actual ratio flowing through the system. The calculated ratio is then sent to a controller, RC101, which manipulates the flow of stream B to maintain set point. The set point to this controller is the required ratio and is set locally.

Both control schemes shown in Fig. 12-1.2 are used, but the scheme shown in Fig. 12-1.2a is preferred because it results in a more linear system than the one shown in Fig. 12-1.2b. This is demonstrated by analyzing the mathematical manipulations in both schemes. In the first scheme, FY 102 solves the equation

$$F_B = RF,$$

The gain of this device-that is, how much its output changes per change in flow rate of stream A-is given by

$$\frac{\partial F_B}{\partial F_A} = R$$

which, as long as the required ratio is constant, is a constant value. In the second scheme, FY 102 solves the equation

$$R = \frac{F_B}{F_A}$$

Its gain is given by

$$\frac{\partial R}{\partial F_A} = -\frac{F_B}{F_A^2} = -\frac{R}{F_A}$$

so as the flow rate of stream A changes, this gain also changes, yielding a nonlinearity.

From a practical point of view, even if both streams can be controlled, the implementation of ratio control may still be more convenient than the control system shown in Fig. 12- 1.1. Figure 12- 1.3 shows a ratio control scheme for this case. If the total flow
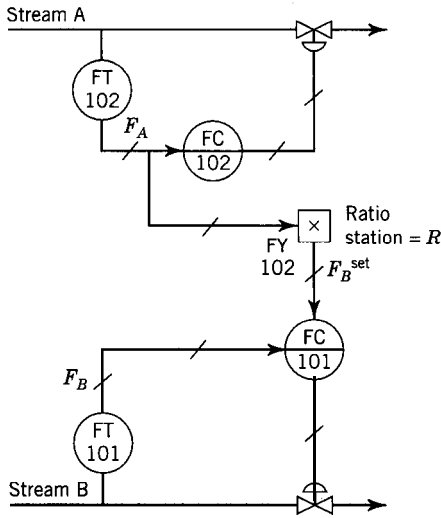
Figure   12-1.3   Ratio   control   of   blending   system.

must be changed, the operator needs to change only one flow, the set point to FC102; then the set point to FC101 changes automatically once the flow rate of stream A changes. In the control system of Fig. 12-1.1, the operator needs to change two flows, the set points to both FC101 and FC102.

The schemes shown in Figs. 12-1.2a and 12-1.3 are quite common in the process industries. Recalling what we learned in Chapter 11 about computing blocks, we realize that implementation of the ratio stations can be accomplished simply with the use of a ratio unit such as the one shown in Table 11-1.2. Most computer control systems offer a controller, referred to as PID-RATIO, that accepts a signal, applies the same algorithm as the ratio unit, FY102, in Fig. 12-1.2a, and uses the internal result as its set point. That is, if the PID-RATIO is used, then the calculations done by FY102 and FC102 are performed in only one block, and in this case the usual notation for the block is FFC, as opposed to just FC.

As we noted in previous chapters, it is helpful in developing control schemes to remember that every signal must have a physical significance. In Figs. 12-1.2 and 12-1.3, we have labeled each signal with its significance. For example, in Fig. 12-1.2a the output signal from FT102 is related to the flow rate of stream A and has the label F_A. If this signal is then multiplied by the ratio F_B/F_A, then the output signal from FY102 is the required flow rate of stream B, F_B^{set}. Even though this use of labels is not standard practice, we will continue to label signals with their significance throughout the chapter for pedagogical reasons. We recommend that the reader do the same.

EXAMPLE 12-2

Another common example of ratio control used in the process industries is control of the air/fuel ratio to a boiler or furnace. Air is introduced in a set excess of that stoichi-ometrically required for combustion of the fuel; this is done to ensure complete com-

bustion. Incomplete combustion not only results in the inefficient use of the fuel but may also result in smoke and the production of other pollutants. The amount of excess air introduced depends on the type of fuel, the fuel composition, and the equipment used. However, the greater the amount of excess air introduced, the greater the energy losses through the stack gases. Therefore, the control of the air and fuel flows is most important for proper safe and economical operation.

The flow of combustibles is generally used as the manipulated variable to maintain the pressure of the steam produced in the boiler at some desired value. Figure 12-1.4 shows one way to control the steam pressure as well as the fuel/air ratio control scheme. This scheme is called **parallel positioning control** (O'Meara, 1979; Scheib and Russell, 1981; Congdon, 1981) with manually adjusted fuel/air ratio. The steam pressure is transmitted by PT101 to the pressure controller PC101, and this controller manipulates a signal, referred to as the **boiler master signal,** to the fuel valve. Simultaneously, the controller also manipulates the air damper through the ratio station FY 101 C. This ratio station sets the fuel/air ratio required.

The control scheme shown in Fig. 12-1.4 does not actually maintain a ratio of fuel flow to air flow; rather, it maintains a ratio of the signals to the final control elements. The flow through these elements depends on these signals and on the pressure drop across them. Consequently, any pressure fluctuation across the valve or air damper changes the flow, even though the opening has not changed, and this in turn affects the combustion process and steam pressure. A better control scheme to avoid these types of disturbances is **full metering control** (O'Meara, 1979). It is shown in Fig. 12-1.5. (Such "top-down" instrumentation diagrams are commonly used to present control schemes.) The fuel/air ratio is still manually adjusted. In this scheme the pressure controller sets the flow of fuel, and the air flow is ratioed to the fuel flow. The flow loops correct for any flow disturbances.
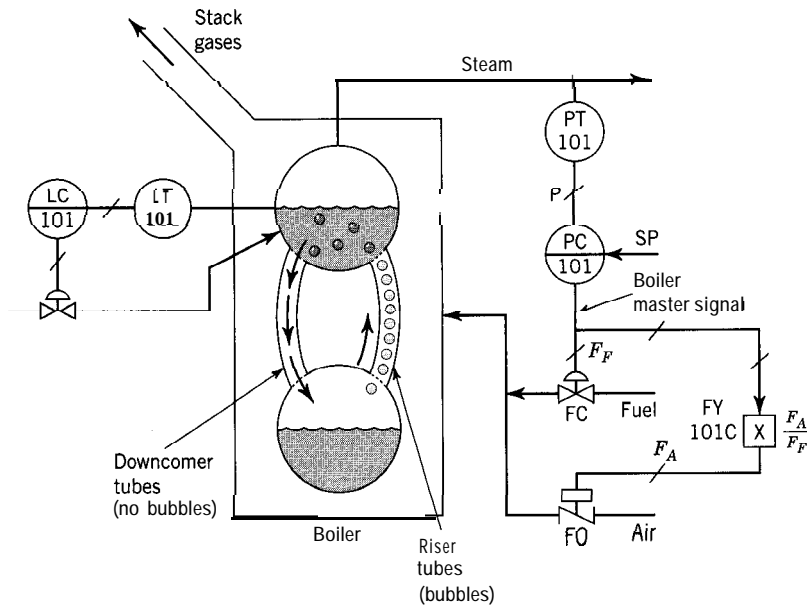
Figure 12-1.4 Parallel positioning control with manually adjusted fuel/air ratio.
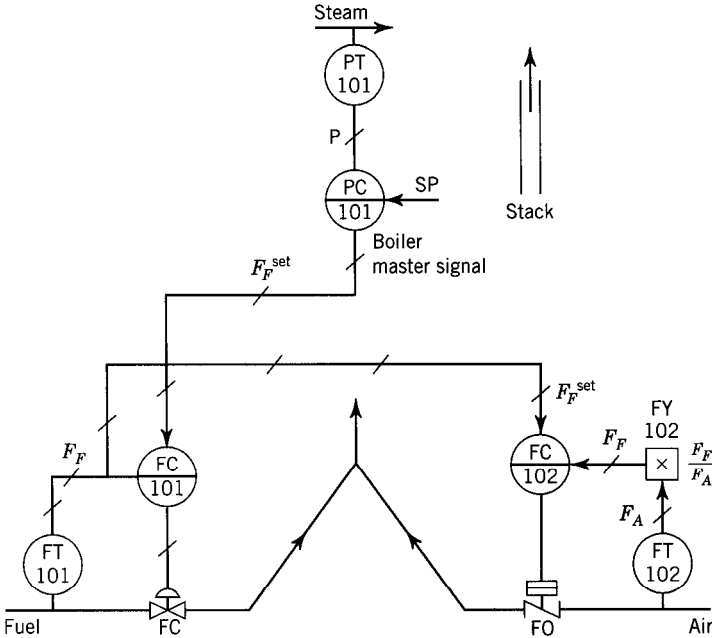
Figure 12-1.5 Full metering control with manually adjusted fuel/air ratio.

Let us analyze the control scheme shown in Fig. 12-1.5 in more detail. When the steam header pressure increases, probably because of a decrease in steam demand, the pressure controller reduces the demand for fuel. As the set point to the fuel flow controller is reduced, the controller closes the valve to satisfy the set point. As the fuel flow decreases, the set point to the air flow controller is also reduced. Thus the air flow follows the fuel flow, and during a transient period, the entering combustible mixture is richer than usual in air. Note that in the figure we have indicated the significance of each signal. Also note the flow loop in the air stream.

Now consider the case wherein the header pressure decreases, probably because of an increase in steam demand, and the pressure controller increases the demand for fuel. As the set point to the fuel flow controller increases, the controller opens the valve to satisfy the set point. As the fuel flow increases, the set point to the air flow is increased; the air flow again follows the fuel flow. In this last case, the entering combustible mixture is not richer in air during a transient period, and if we are not careful it may be lean in air. This situation is certainly not desirable for two important reasons. First, a lean air mixture may result in pockets of excess fuel in the combustion chamber: not a very safe (in fact, an explosive!) condition. Second, a lean air mixture results in unburned fuel in the stack gases, which constitutes an environmental hazard and a waste of energy and money. Therefore, a control scheme must be designed to avoid these situations. The control scheme must be such that when more combustibles are required to maintain pressure, it increases the air first and then the fuel. When less combustibles are required, it decreases the fuel first and then the air. This pattern ensures that the combustible mixture is air-rich during transient periods. Figure 12-1.6 shows a scheme,
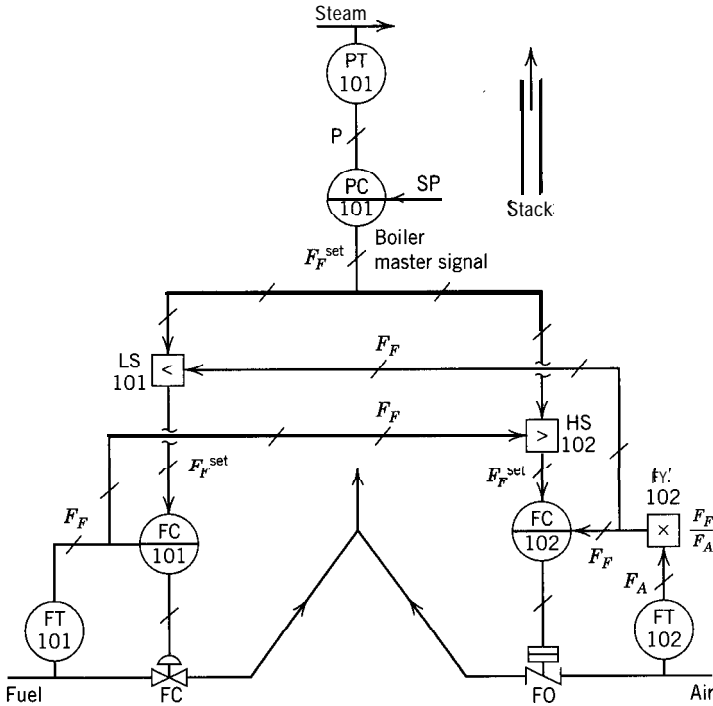
Figure 12-1.6 Cross-limiting control.

referred to as *cross-limiting control,* that provides the required control. Only two selectors, LS-101 and HS-102, are added to the previous control scheme. The selectors provide a way to decide which device sets the set point to the controller; their action is similar to that of an override scheme as presented in Chapter 11. The reader is encouraged to "go through" the scheme to understand how it works. As a way to do so, assume that the required air/fuel ratio is 2 and that at steady-state the required fuel is 10 units of flow. Consider next what happens if the header pressure increases and the pressure controller asks for 8 units of fuel flow. Finally, consider what happens if the header pressure decreases and the pressure controller asks for 12 units of fuel flow.

Because the amount of excess air is so important to the economical and environmentally responsible operation of boilers, it has been proposed that some feedback signal based on an analysis of the stack gases be provided; the analysis is usually percent 0, or CO. It is proposed that the fuel/air ratio then be adjusted on the basis of this analysis. This new scheme, which is shown in Fig. 12-1.7, consists of an analyzer transmitter, AT103, and a controller, AC103. The controller maintains the required percent O_2, for example, in the stack gases by setting the required fuel/air ratio. The figure shows the use of high and low limiters, HL103 and LL103. These two units are used mainly for safety reasons. They ensure that the air/fuel ratio will always be between some preset high and low values.

This section has shown two applications of ratio control. As we noted at the beginning of the section, ratio control is widely used in the process industries; it is simple and easy to use.
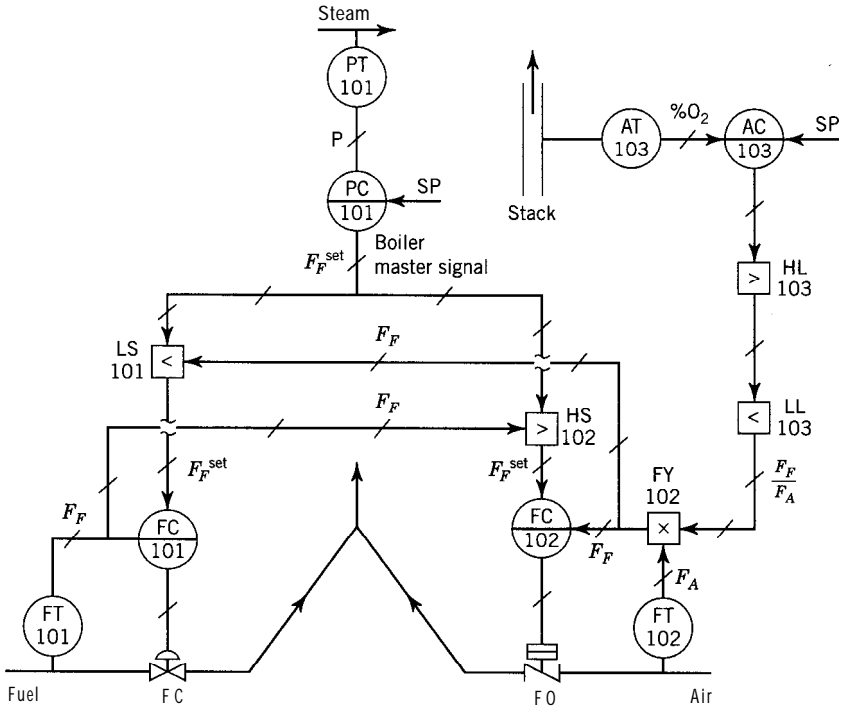
Figure **12-1.7** Cross-limiting with 0, trim control.

12-2 FEEDFORWARD CONTROL

This section presents the principles and application of feedforward control, quite often a most profitable control strategy. Feedforward is not a new strategy; the first reports date back to the early 1960s (Dobson, 1960; Shinskey, 1963). However, the use of computers has significantly simplified and expanded its implementation. Feedforward requires a thorough knowledge of the steady-state and dynamics characteristics of the process. Thus good process engineering knowledge is basic to its application.

12-2.1 The Feedforward Concept

To help us understand the concept of feedforward control, let's briefly review feedback control; Fig. 12-2.1 depicts the feedback concept. As different disturbances, d,(t), $d_2(t)$, . . . , d,(t), enter the process, the controlled variable, $c(t)$, deviates from set point, and feedback compensates by manipulating another input to the process, the manipulated variable, $m(t)$, The advantage of feedback control is its simplicity. Its disadvantage is that in order to compensate for disturbances, the controlled variable must first deviate from set point. Feedback acts on an error between the set point and the controlled variable. It may be thought of as a *reactive* control strategy: it waits until the process has been upset before it begins to take corrective action.

By its very nature, feedback control results in a temporary deviation in the controlled variable. Many processes can permit some amount of deviation. In many other pro-
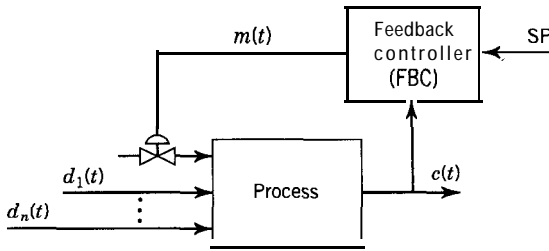
Figure   12-2.1   Feedback   concept.

cesses, however, this deviation must be minimized to such an extent that feedback control may not provide the required performance. For these cases, feedforward control may prove helpful.

The idea of feedforward control is to compensate for disturbances before they affect the controlled variable. Specifically, feedforward calls for measuring the disturbances before they enter the process and, on the basis of these measurements, calculating the required manipulated variable to maintain the controlled variable at set point. If the calculation and action are done correctly, the controlled variable should remain undisturbed. Thus feedforward control may be thought of as a proactive control strategy; Fig. 12-2.2 depicts this concept.

To further explain, consider a disturbance d(t), as shown in Fig. 12-2.3, entering the process. As soon as the feedforward controller (FFC) realizes that a change has occurred, it calculates a new value of $m_{FF}(t)$ and sends it to the process (valve). This is done such that path G_M negates the effect of path G,. G_M is the transfer function that describes how the manipulated variable, $m_{FF}(t)$, affects the controlled variable, and G_D is the transfer function that describes how the disturbance, d(t), affects the controlled variable.

To attain perfect negation, the feedforward controller must be designed by taking into account the steady-state characteristics of the process. That is, assume that a change of plus one unit in d(t) affects c(t) by plus ten units and that a change of plus one unit in m,(t) affects c(t) by plus five units. Thus if d(t) changes by plus one unit, affecting c(t) by plus ten units, then the feedforward controller must change $m_{FF}(t)$ by minus two units, affecting c(t) by minus ten units and, therefore, negating the effect of d(t).
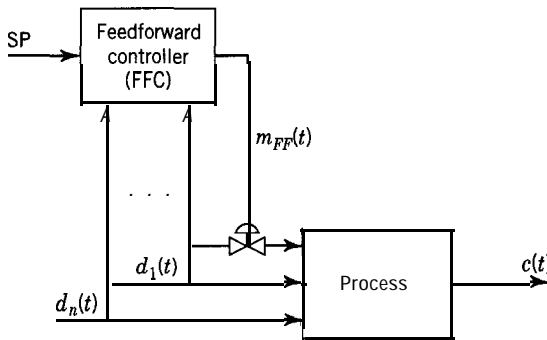
Figure 12-2.2 Feedforward concept.
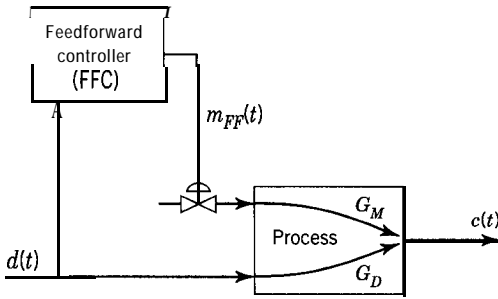
Figure 12-2.3 Feedforward control.

The preceding paragraph explains how the feedforward control strategy compensates through consideration of the steady-state characteristics of the process. However, to avoid any change in the controlled variable, the dynamic characteristics of the process must also be considered. It is desired that the effects of m,(t) and d(t) affect c(t) at the same time. That is, consider that when d(t) changes, the feedforward controller changes m,(t) almost at the same time. If as a result of process dynamics, the effect of $m_{FF}(t)$ on c(t) is faster than the effect of d(t) on c(t), then c(t) will deviate from its desired value because of m„(t), not because of d(t)! In this case, perfect compensation requires "slowing down" the feedforward controller. That is, the feedforward controller should not take immediate corrective action; rather, it should wait a certain time before taking action so that the negating effects reach c(t) at the same time. In other processes, the effect of d(t) on c(t) may be faster than the effect of $m_{FF}(t)$ on c(t). In these cases, perfect compensation requires "speeding up" the feedforward controller. Thus *the feed-forward controller must be designed to provide both the required steady-state and dynamic compensations.*

Figure 12-2.2 shows feedforward compensation for all the disturbances entering the process. However, very often it may be difficult, if not impossible, to measure some disturbances. In addition, some of the measurable disturbances may occur so infre-quently that the need for compensation by feedforward is questionable. Therefore, **feed-**forward control is used to compensate for the major measurable disturbances. It is up to the operating personnel to define "major" disturbances (those that occur often and cause significant deviations in the controlled variable). Feedback control is then used to compensate for those disturbances that are not compensated for by feedforward. Fig. 12-2.4 shows a possible implementation of this feedforward/feedback control.

The foregoing paragraphs have explained the objective, and some design consider-ations, of feedforward control. The things to keep in mind are that the feedforward controller must usually contain steady-state and dynamic compensations and that feed-back compensation must always be present.

12-2.2 Block Diagram Design of Linear Feedforward Controllers

This section and the next three show the design of feedforward controllers. The mixing system shown in Fig. 12-2.5 is used to illustrate this design; Table 12-2.1 gives the steady-state conditions and other process information. In this process, three different streams are mixed and diluted with water to a final desired composition of component
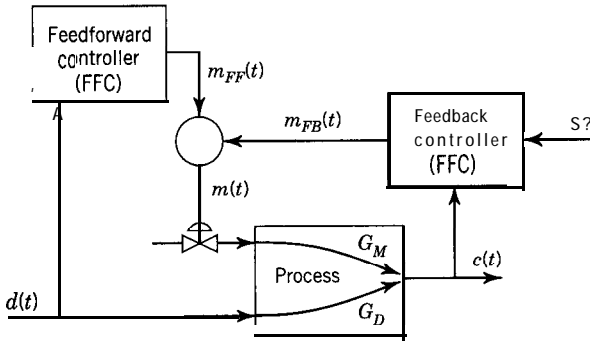
Figure   12-2.4 Feedforward/feedback   control.

A, $x_6(t)$. Process considerations dictate that the mixing be done in three constant-volume tanks, as shown in the figure. All of the input streams represent possible disturbances to the process; that is, the flows and compositions of streams 5, 2, and 7 may vary. However, the major disturbances usually come from stream 2. Commonly, the stream flow, $f_2(t)$, may double, whereas the mass fraction, $x_2(t)$, may decrease as much as 20% of its steady-state value. Figure 12-2.6 shows the control, provided by feedback control, when $f_2(t)$ changes from 1000 gpm to 2000 gpm. The composition changes from its steady-state value of 0.472 mass fraction (mf) to about 0.510 mf, a 8.05% change from set point. An index often used to evaluate control performances is the integral of the absolute value of the error (IAE), which is the total area under the curve, or error. The IAE value for the response shown in Fig. 12-2.6 is 33.34 mf-min. Fig.
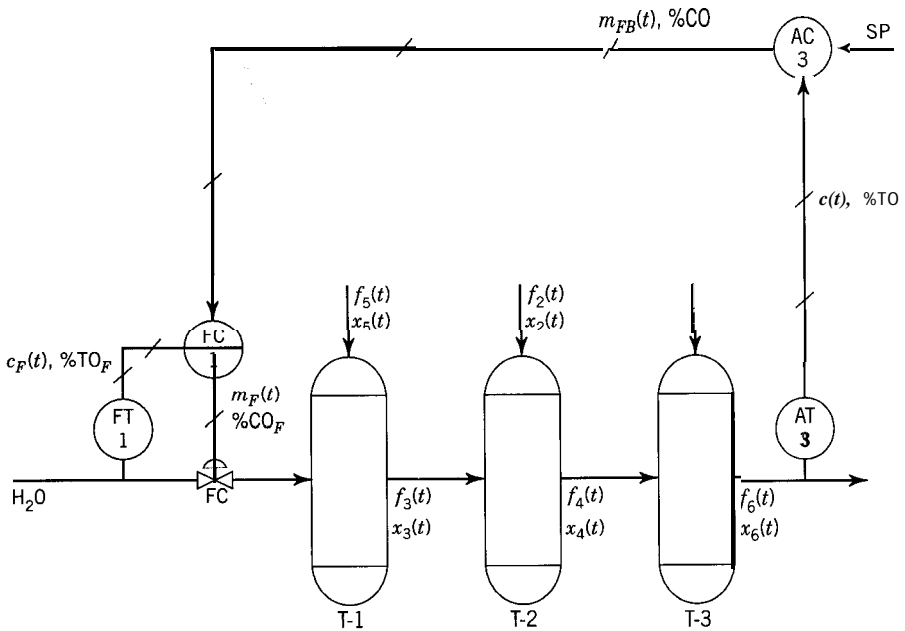
Figure   12-2.5 Mixing   process.

Table 12-2.1 Process Information and Steady-State Values for Mixing Process

Information

Concentration transmitter range: 0.3-0.7 mass fraction. Its dynamic can be described by a time constant of 0.1 min.

The pressure drop across the valve can be considered constant, and the maximum flow provided by the valve is 3800 gpm. The valve dynamics can be described by a time constant of 0.1 min.

The densities of all streams are also considered similar and constant.

Steady-State Values

Stream	Flow, gpm	Mass Fraction	Stream	Flow, gpm	Mass Fraction
1	1900	0.000	5	500	0.800
2	1000	0.990	6	3900	0.472
3	2400	0.167	7	500	0.900
4	3400	0.409			

12-2.6 shows the performance provided by feedback control, but the maximum deviation permitted for this process is $\pm$ 1.5% from set point. That is, the maximum value of composition permitted is 0.479 mf, and the minimum value permitted is 0.465 mf. Thus it does not appear that simple feedback can provide the required performance; feedforward control may be justified.

Assuming for the moment that $f_2(t)$ is the major disturbance, the application of feedforward to this process calls for measuring this flow and, upon a change, taking corrective action. Let's examine the design of this feedforward controller.

The block diagram for this process is shown in Fig. 12-2.7a. The diagram shows that $f_2(t)$ is the disturbance of concern. We next note the significance of each transfer function.
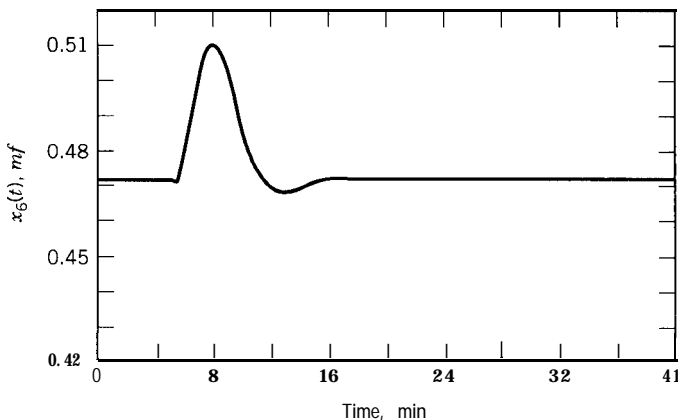
Figure 12-2.6 Feedback control--f,(t) changed from 1000 gpm to 2000 gpm.
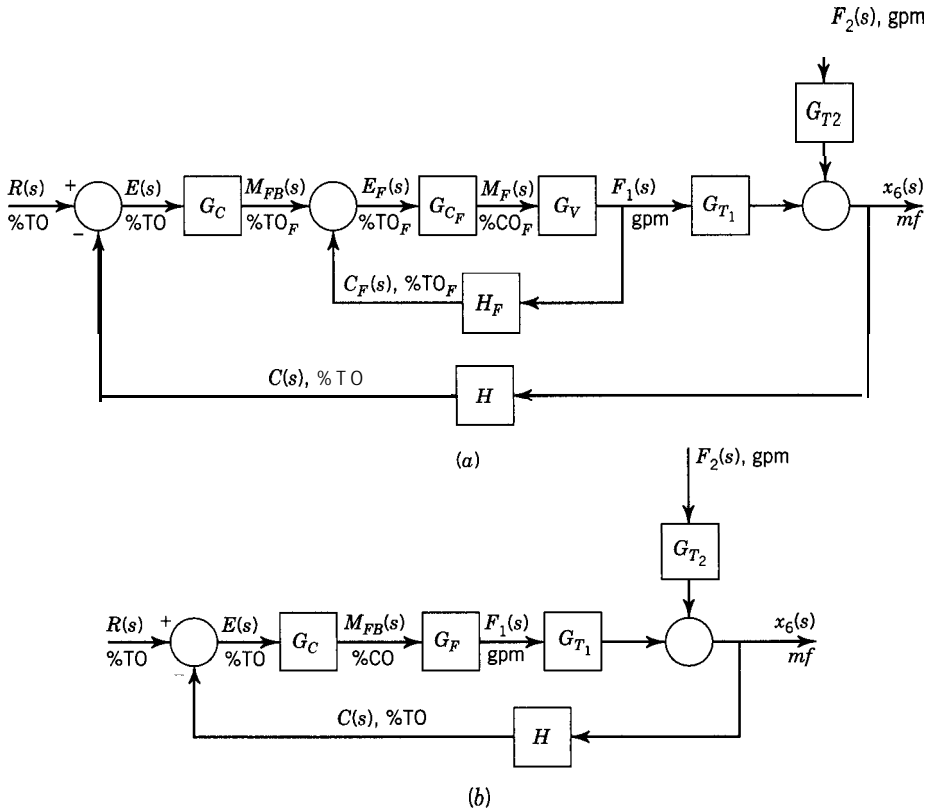
Figure 12-2.7 Block diagram of mixing process.

G_c = transfer function of the composition controller.

G_{C_F} = transfer function of the flow controller.

G_v = transfer function of the valve. It describes how the water flow is affected by the flow controller output.

G_{T_1} = transfer function of the mixing process. It describes how $x_6(t)$ is affected by the water flow.

G_{T_2} = transfer function of the mixing process. It describes how $x_6(t)$ is affected by $f_2(t)$.

H_F = transfer function of the flow sensor and transmitter.

H = transfer function of the concentration sensor and transmitter.

Because the flow loop, once tuned, is fast and stable, Fig. 12-2.7a can be simplified as shown in Fig. 12-2.7b. The new transfer function is

G_F = transfer function that describes how the water flow is affected by the composition controller, $G_F = \dfrac{G_{C_F}G_v}{1 + G_{C_F}G_vH_F}$.

A more condensed block diagram, shown in Fig. 12-2.8, can be drawn and compared to that of Fig. 12-2.7. The significance of each transfer function follows.
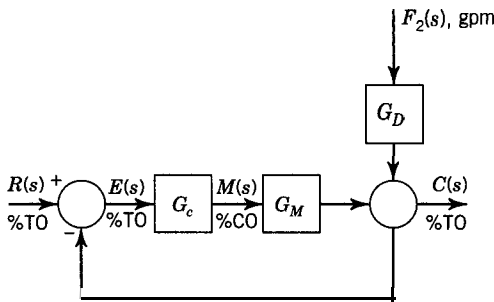
Figure 12-2.8 Block diagram of mixing process.

G_M = transfer function that describes how the manipulated variable, $M_{FB}(s)$, affects the controlled variable, $C(s)$. In this case, $G_M = G_F G_{T_1} H$.

G_D = transfer function that describes how the disturbance, $F_2(s)$, affects the controlled variable. In this case, $G_D = G_{T_2} H$.

To review, the objective of feedforward control is to measure the input(s) and, if a disturbance is detected, to adjust the manipulated variable to maintain the controlled variable at set point. This control operation is shown in Fig. 12-2.9. The significance of each new transfer function follows.

H_D = transfer function that describes the sensor/transmitter that measures the disturbance.

FFC = transfer function of feedforward controller.

Note that in Fig 12-2.9 the feedback controller has been "disconnected." This controller will be "connected" again later.

Figure 12-2.9 shows that the way the disturbance, $F_2(s)$, affects the controlled variable, $C(s)$, is given by

$$C(s) = G_D F_2(s) + H_D(FFC)(G_M)F_2(s)$$
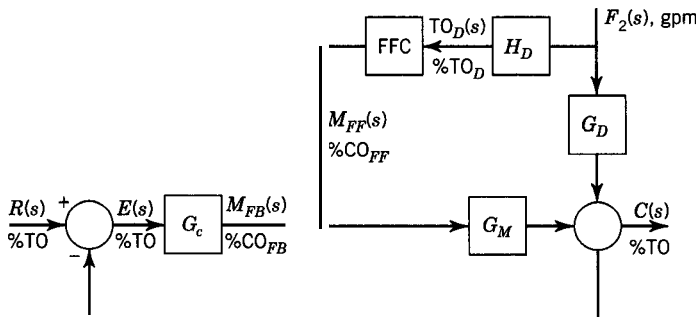
Figure 12-2.9 Block diagram of feedforward control for mixing process.

The objective is to design FFC such that a change in F,(s) does not affect C(s)-that is, such that C(s) = 0. Thus

$$0 = G_D F_2(s) + H_D(\text{FFC})(G_M)F_2(s)$$

Dividing both sides by $F_2(s)$ and solving for FFC, we get

$$\text{FFC} = -\frac{G_D}{H_D G_M} \tag{12-2.1}$$

Equation 12-2.1 is the design formula of the feedforward controller.
   As we noted in previous chapters, first-order-plus-dead-time transfer functions are commonly used as an approximation to describe processes; Chapter 7 showed how to evaluate this transfer function from step inputs. Using this type of approximation for this process, we have

$$G_D = \frac{K_D e^{-t_{0_D}s}}{\tau_D s + 1}, \frac{\% \text{ T O}}{\text{gpm}} \tag{12-2.2}$$

$$G_M = \frac{K_M e^{-t_{0_M}s}}{\tau_M s + 1}, \frac{\% \text{ T O}}{\% \text{CO}} \tag{12-2.3}$$

And assuming that H_D is only a gain,

$$H_D = K_{T_D}, \frac{\% \text{TO}_D}{\text{gpm}} \tag{12-2.4}$$

Substituting Eqs. 12-2.2, 12-2.3, and 12-2.4 into Eq. 12-2.1 yields

$$\text{FFC} = \frac{M_{FF}(s)}{\text{TO},(s)} = -\frac{K_D}{K_{T_D} K_M} \left(\frac{\tau_M s + 1}{\tau_D s + 1}\right) e^{-(t_{0_D} - t_{0_M})s} \tag{12-2.5}$$

Equation 12-2.5 contains several terms that will be explained in detail next; the implementation of this equation is shown in Fig. 12-2.10.
   The first element of the feedforward controller, $-\dfrac{K_D}{K_{T_D} K_M}$, contains only gain terms. This term is the part of the feedforward controller that compensates for the steady-state differences between the "G, and G_D paths." The units of this term also help in understanding its significance:

$$\frac{K_D}{K_{T_D} K_M} = \frac{\dfrac{\% \text{ T O}}{\text{gpm}}}{\dfrac{\% \text{TO}_D}{\text{gpm}} \dfrac{\% \text{ T O}}{\% \text{CO}_{FF}}} = \frac{\% \text{CO}_{FF}}{\% \text{TO}_D}$$
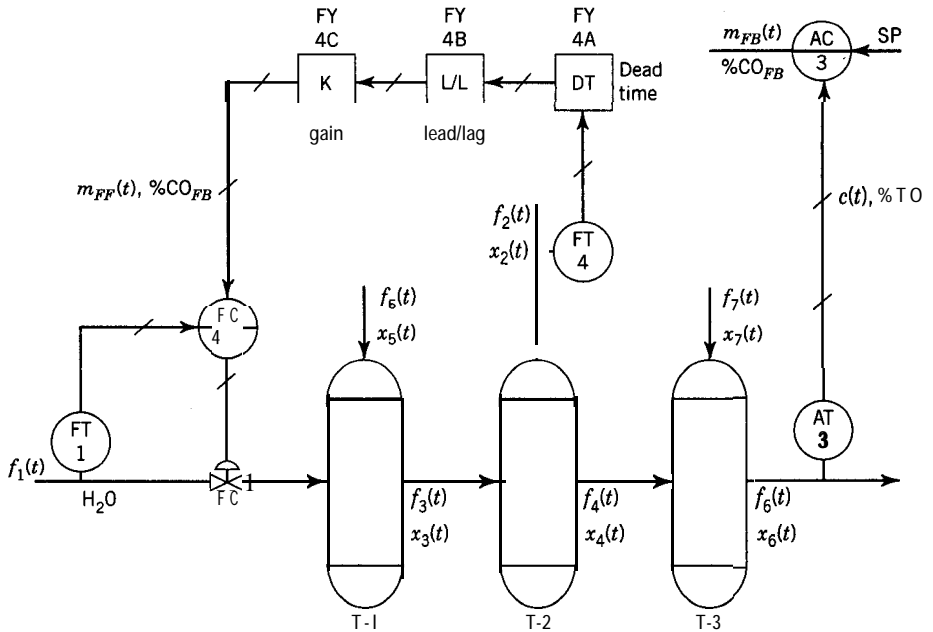
Figure 12-2.10 Implementation of feedforward control.

Thus the units show that the term indicates how much the feedforward controller output, $M_{FF}(s)$, changes per unit change in the transmitter's output, TO,(s).

Note the minus sign in front of the gain term in Eq. 12-2.1 and Eq. 12-2.5. This sign helps to decide the "action" of the controller. In the process at hand, K_D is positive; as $f_2(t)$ increases, the outlet concentration, $x_6(t)$, also increases because stream 2 is more concentrated than the outlet stream. K_M is negative; as the signal to the flow controller increases, the valve opens and the outlet concentration decreases. Finally, K_{T_D} is positive: as $f_2(t)$ increases, the signal from the transmitter also increases. Thus the sign of the gain term is negative:

$$\frac{K_D}{K_{T_D}K_M} \rightarrow \frac{(+)}{(+)(-)} = -$$

A negative sign means that as $TO_D(t)$ increases, indicating an increase in $f_2(t)$, the feedforward controller output, $m_{FF}(t)$, should decrease, reducing the water flow. This action does not make sense. As $f_2(t)$ increases, tending to increase the concentration of the output stream, the water flow should also increase to dilute the outlet concentration, thus negating the effect of $f_2(t)$. Therefore, the sign of the gain term should be positive. Note that multiplying the negative sign in front of the gain term by the sign of this term results in the correct feedforward action.

The second term of the feedforward controller includes only the time constants of the "G_D and G_M paths." This term, referred to as a lead/lag (L/L), compensates for the differences in time constants between the two paths. Section 12-2.3 discusses this term in detail.

The last term of the feedforward controller contains only the dead-time terms of the "G_D and G_M paths." This term, referred to as a dead-time compensator, compensates for the differences in dead time between the two paths. Sometimes the term $(t_{0_D} - t_{0_M})$ may be negative, yielding a positive exponent. As we noted in Chapter 2, the Laplace representation of dead time includes a negative sign in the exponent. When the sign is positive, it is definitely not a dead time and cannot be implemented. A negative sign in the exponent is interpreted as "delaying" an input; a positive sign may indicate a "forecasting." That is, the controller requires taking action before the disturbance happens. This is not possible! When this occurs, there is quite often a physical explanation, as the present example will show.

Thus it can be said that the first term of the feedforward controller is a steady-state compensator, whereas the last two terms are dynamic compensators. All these terms are easily implemented using computer control software. Years ago, however, when only analog instrumentation was used, the dead-time compensator was either extremely difficult or impossible to implement. At that time, the better approach was to implement only the steady-state and lead-lag compensators. Figure 12-2.10 shows a component for each calculation needed for the feedforward controller-that is, one component for the dead time, one for the lead/lag, and one for the gain. Very often, however, lead/lags have adjustable gains (see Tables 11-1.1 and 1 1-1.2), and in this case we can combine the components **FY4B** and **FY4C** into one component. This is more efficient because it uses fewer components. From now on we will show only one component, the lead/lag, and the gain will be included.

Let's return to the mixing system. Under open-loop conditions, a step change of 5% in the signal to the valve provides a process response from which the following first-order-plus-dead-time approximation is obtained:

$$G_M = \frac{-1,095e^{-0.93s}}{3.82s + 1} \frac{\%\,TO}{\%\,CO} \tag{12-2.6}$$

Also, under open-loop conditions, $f_2(t)$ was allowed to change by 10 gpm in a step fashion, and from the process response the following approximation is obtained:

$$G_D = \frac{0.0325e^{-0.75s}}{2.75s + 1} \frac{\%\,TO}{gpm} \tag{12-2.7}$$

Finally, assuming that the flow transmitter in stream 2 is calibrated from 0 to 3000 gpm, its transfer function is given by

$$H_D = K_{T_D} = \frac{100 \,\%TO_D}{3000 \ gpm} = 0.0333 \ \frac{\%TO_D}{gpm}$$

Substituting these three transfer functions into Eq. 12-2.1 yields

$$FFC = \frac{M_{FF}(s)}{TO_D(s)} = 0.891 \left(\frac{3.82s + 1}{2.75s + 1} \right) e^{-(0.75-0.93)s} \tag{12-2.8}$$

The dead time indicated, 0.75-0.93, is negative, so the dead-time compensator cannot be implemented. Thus the implementable, or realizable, feedforward controller is

$$\text{FFC} = \frac{M_{FF}(s)}{TO_D(s)} = \mathbf{0.891} \left(\frac{3.82s+1}{2.75s+1} \right) \tag{12-2.9}$$

Figure 12-2.11 shows the implementation of this controller. The figure shows that the feedback compensation has also been implemented. This implementation has been accomplished by adding the output of both feedforward and feedback controllers in FY3. Section 12-2.4 discusses how to implement this addition. Figure 12-2.12 is the block diagram for this combined control scheme.

Figure 12-2.13 shows the response of the composition when $f_2(t)$ doubles from 1000 gpm to 2000 gpm. The figure compares the control provided by feedback control, steady-state feedforward control, and dynamic feedforward control. In steady-state feedforward control, no dynamic compensation is implemented; that is, in this case, FFC = 0.891. Dynamic feedforward control includes the complete controller, Eq. 12-2.9. Under steady-state feedforward, the mass fraction increased to 0.477 mf, a 1.05% change from set point. Under dynamic feedforward, the mass fraction increased to 0.473 mf, a 0.21% change. The improvement provided by feedforward control is quite impressive. Figure 12-2.13 also shows that the process response tends to decrease first and then to increase; we will discuss this response later.

The previous paragraphs and figures have shown the development of a linear feedforward controller and the responses obtained. At this stage, because we have not of-
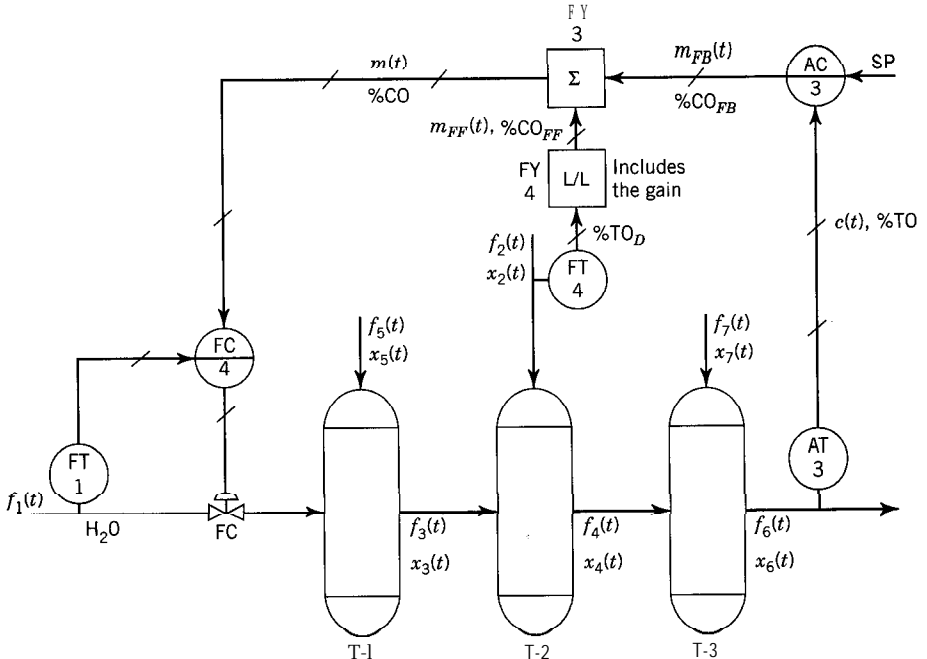
Figure 12-2.11 Implementation of feedforward/feedback control.
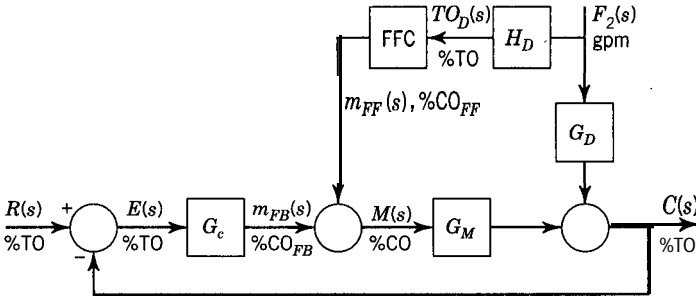
Figure 12-2.12 Block diagram of feedforward/feedback control.

fered an explanation of the lead/lag unit, the reader may be wondering about it. Thus we will explain this term before further discussing feedforward control.

12-2.3 Lead/Lag Term

As indicated in Eqs. 12-2.5 and 12-2.9, the lead/lag term is composed of a ratio of two $(\tau s + 1)$ terms. More specifically, its transfer function is

$$\frac{O(s)}{I(s)} = \frac{\tau_{ld}s + 1}{\tau_{lg}s + 1} \tag{12-2.10}$$

where

$O(s) = $ **Laplace** transform of output variable
$I(s) = $ **Laplace** transform of input variable
$\tau_{ld} = $ lead time constant
$\tau_{lg} = $ lag time constant
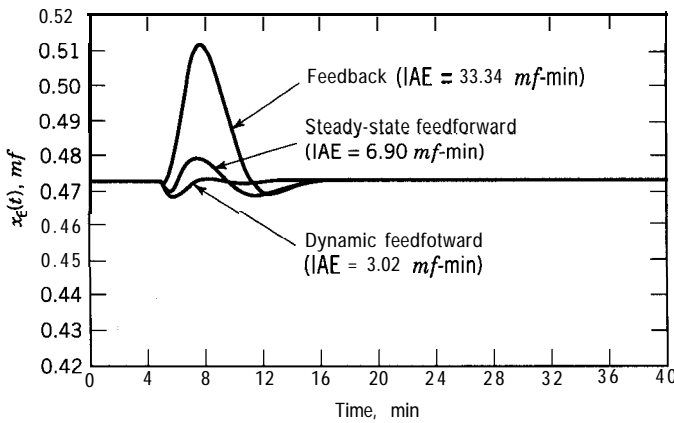
Figure 12-2.13 Comparison of feedback and feedforward control to a change of 1000 gpm in $f_2(t)$.

This lead/lag term was discussed in Chapter 2 (see Section 2-4), but a brief review is warranted here. Let us suppose that the input changes, in a step fashion, A units of amplitude; that is,

$$I(s) = \frac{A}{s}$$

Substituting this expression for $I(s)$ in Eq. 12-2.10 and inverting the equation back to the time domain, we get

$$O(t) = A\left[1 + \frac{\tau_{ld} - \tau_{lg}}{\tau_{lg}} e^{-t/\tau_{lg}} \right] \tag{12-2.11}$$

Figure 12-2.14a shows the response for different values of the ratio τ_{ld}/τ_{lg} while τ_{lg} remains 1. The figure shows that as the ratio τ_{ld}/τ_{lg} increases, the initial response also increases; as time increases, the response approaches exponentially its final value of A. For values of $\tau_{ld}/\tau_{lg} > 1$ the initial response is greater than its final value, whereas for values of $\tau_{ld}/\tau_{lg} < 1$ the initial response is less than its final value. Therefore, the initial response depends on the ratio of the lead time constant to the lag time constant, τ_{ld}/τ_{lg}. The approach to the final value depends only on the lag time constant. Thus, in tuning a lead/lag, we must provide both τ_{ld} and τ_{lg}.
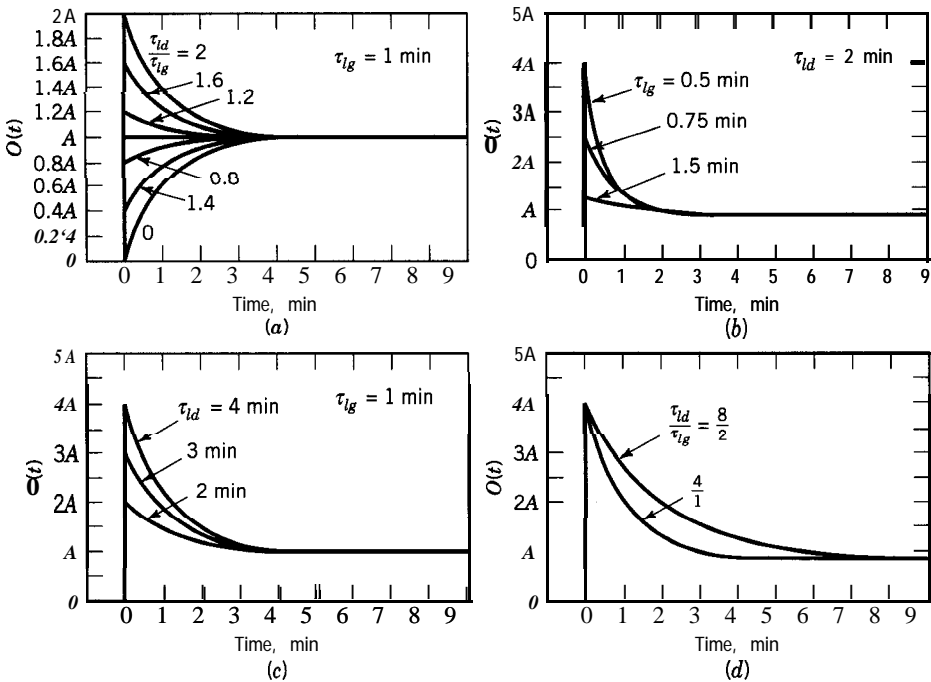
Figure 12-2.14 Response of lead/lag to an input of A units of magnitude.

Figures 12-2.14*b* through 12-2.14*d* help further illustrate the effects of τ_{ld} and τ_{lg} on the response of the lead/lag to a step change of *A* units of magnitude. Figure 12-2.14*b* shows how τ_{lg} affects the response while τ_{ld} is kept constant. The figure shows that as τ_{lg} decreases, the ratio τ_{ld}/τ_{lg} increases, the magnitude of the initial output response increases, and the rate at which the response approaches its final value increases. Figure 12-2.14*c* shows how τ_{ld} affects the response while τ_{lg} is kept constant. The figure shows that as τ_{ld} increases, the ratio τ_{ld}/τ_{lg} also increases, and the magnitude of the initial output response increases. The figure also shows that all curves reach the final value at the same time, because τ_{lg} is the same in all cases. Finally, Fig. 12-2.14*d* shows two response curves with identical values of the ratio τ_{ld}/τ_{lg} but different individual values of τ_{ld} and τ_{lg}. The figure shows that the magnitude of the initial output response is the same, because the ratio is the same, but the response with the larger τ_{lg} takes longer to reach the final value.

Equation 12-2.5 reflects the use of a lead/lag term in the feedforward controller. The equation indicates that τ_{ld} should be set equal to τ_M and that τ_{lg} should be set equal to τ_D.

12-2.4   Back to the Previous Example

With an understanding of the lead/lag term we can now return to the example of Section 12-2.2 and specifically to a discussion of the dynamic compensation of the feedforward controller.

Comparing the transfer functions given by Eqs. 12-2.6 and 12-2.7, it is easy to see that the controlled variable, c(t), responds slower to a change in the manipulated variable, m(t), than to a change in the disturbance, $f_2(t)$. Recall that one design consideration for a feedforward controller is to compensate for the dynamic differences between the manipulated and the disturbance paths, the G_M and G_D paths. The feedforward controller for this process should be designed to "speed up" the response of the controlled variable to a change in the manipulated variable. That is, the feedforward controller should "speed up" the G_M path; the resulting controller, Eq. 12-2.9, does exactly this. First, realize that the resulting lead/lag term has a τ_{ld}/τ_{lg} ratio greater than 1; $\dfrac{\tau_{ld}}{\tau_{lg}} = \dfrac{3.82}{2.75} =$ 1.39. This means that at the moment the signal from the flow transmitter changes by 1%, indicating a certain change in $f_2(t)$, the lead/lag output changes by 1.39%, resulting in an initial output change from the feedforward controller of $(0.891)(1.39) =$ 1.238%. Eventually, the lead/lag output approaches 1%, and the feedforward controller output approaches 0.891%. This type of action results in an initial increase in $f_1(t)$ greater than the one really needed for the specific increase in $f_2(t)$. This initial greater increase provides a "kick" to the G_M path to move faster, resulting in a tighter control than that provided by steady-state feedforward control; this is shown in Fig. 12-2.12. Second, note that the feedforward controller equation does not contain a dead-time term. There is no need to delay the feedforward action. On the contrary, the present process makes it necessary to "speed up" the feedforward action. Thus the absence of a dead-time term makes sense.

It is important to realize that this feedforward controller, Eq. 12-2.9, compensates only for changes in $f_2(t)$. Any other disturbance will not be compensated for by the controller, and in the absence of a feedback controller, it would result in a deviation of

the controlled variable. The implementation of feedforward control requires the presence of feedback control. Feedforward control compensates for the major measurable disturbances, while feedback control takes care of all other disturbances. In addition, the feedback controller compensates for any inexactness in the feedforward controller. Thus *feedforward control must, whenever possible, be implemented with feedback compensation!* Feedback from the controlled variable must, whenever possible, be present.

Figure 12-2.11 shows a summer wherein the signals from the feedforward controller, $m_{FF}(t)$, and from the feedback controller, $m,(t)$, are combined. A way to think about the significance of the signals entering the summer is to note that the signal from the feedforward calculation, $m,(t)$, is related to the flow $f_1(t)$ required to compensate for the major disturbances. The signal from the feedback controller, $m_{FB}(t)$, is $\Delta f_1(t)$. That is, this feedback signal biases the feedforward calculation to correct for unmeasured disturbances or for errors in the feedforward calculation. The summer solves the equation

$$\text{OUT} = \text{Feedback signal} + \text{Feedforward signal} + \text{Bias}$$

To be more specific, consider the use of the summer shown in Table 1 1- 1.1,

$$\text{OUT} = K_X X + K_Y Y + K_Z Z + B_0$$

Let the feedback signal be the X input, let the feedforward signal be the Y input, and the Z input is not used. Therefore,

$$\text{OUT} = K_X[m_{FB}(t)] + K_Y[m_{FF}(t)] + B_0$$

As previously discussed, the sign of the steady-state part of the feedforward controller, $\dfrac{-K_D}{K_{TD}K_m}$, is positive for this process, so the value of K_Y is set to $+1$ (if the sign had been negative, then K_Y would have been set to -1). The value of K_X is also set to $+1$. Note that setting K_Y to 0 or to 1 provides an easy way to "turn off" or "turn on" the feedforward controller.

Let us suppose that the process is at steady-state under feedback control only ($K_Y = 0$, $B_0 = 0$) and that it is now desired to "turn on" the feedforward controller. Furthermore, it is desired to turn on the feedforward controller without upsetting the signal to the flow controller, because the process is at steady state. That is, a "bumpless-transfer" from simple feedback control to feedforward/feedback control is desired. To accomplish this transfer, the summer is first set to "manual," which freezes its output; K_Y is set to $+1$, the output of the feedforward controller, $m,(t)$, is read from FY4; the bias term, B_0, is set equal to the negative of the value read in FY4; and finally, the summer is set to automatic again. This procedure results in the bias term canceling the feedforward controller output. To be a bit more specific, suppose the process is running under feedback control only, with a signal to the flow controller equal to $m,(t)$. It is then desired to "turn on" the feedforward controller, and at this time the process is at steady state with $f_2(t) = 1500$ gpm. Under this condition, the output of the flow trans-

mitter is at 50%, and m,(t) is at $(50\%)(0.891)$, or 44.55%. Then the procedure just explained is followed, yielding

$$\text{OUT} = (1)m_{FB}(t) + (1)(44.55) - 44.55 = m,(t)$$

Now suppose $f_2(t)$ changes from 1500 gpm to 1800 gpm, making the output from the flow transmitter equal to 60%. After the transients through the lead/lag have died out, the output from the feedforward controller becomes equal to $(60\%)(0.891) = 53.46\%$. Thus the feedforward controller asks for 8.91% more signal to the flow controller to compensate for the disturbance. At this moment, the summer output signal becomes

$$\text{OUT} = (1)m_{FB}(t) + (1)(53.46) - 44.55 = m_{FB}(t) + 8.91\%$$

which changes the signal to the flow controller by the required amount.

The procedure just described to implement the summer is easy, but it requires the manual intervention of the operating personnel. Most control systems can be easily configured to perform the procedure automatically. For instance, consider the use of an on-off switch. Whenever the switch is off (indicating that no feedforward is desired), the bias value is set equal to the negative of the output of FY4 before the summation is performed. Thus they cancel out during the summation, and the result is equal to the output of the feedback controller only, $m_{FB}(t)$, which is the required result. Whenever the switch is set to on (feedforward control is desired), the bias is kept constant at the last value set. Therefore, whenever the feedforward controller changes its output, the summer changes by the same amount.

The previous paragraphs have explained just one way to implement the summer where the feedback and feedforward signals are combined. The importance of the bumpless transfer was stressed. The way the summer is implemented depends on the algorithms provided by the control system used. For example, there is one control system that provides a lead/lag and a summer in only one algorithm, called a lead/lag summer (see Table 11-1.1). In this case, the feedback signal can be brought directly into the lead/lag, and the summation is done in the same unit; the summer unit is not needed. There are other control systems that provide what is called a PID-FEEDFORWARD. In this case, the feedforward signal is brought into the feedback controller (the PID-FEEDFORWARD controller) and added to the feedback signal calculated by the controller. How the bumpless transfer is accomplished depends on the control system.

So far in our example, feedforward control has been implemented to compensate for $f_2(t)$ only. But what if it is necessary to compensate for another disturbance such as $x_2(t)$? The technique for designing this new feedforward controller is the same as before; Fig. 12-2.15 shows a block diagram, including the new disturbance with the new feedforward controller FFC,. The new controller equation is

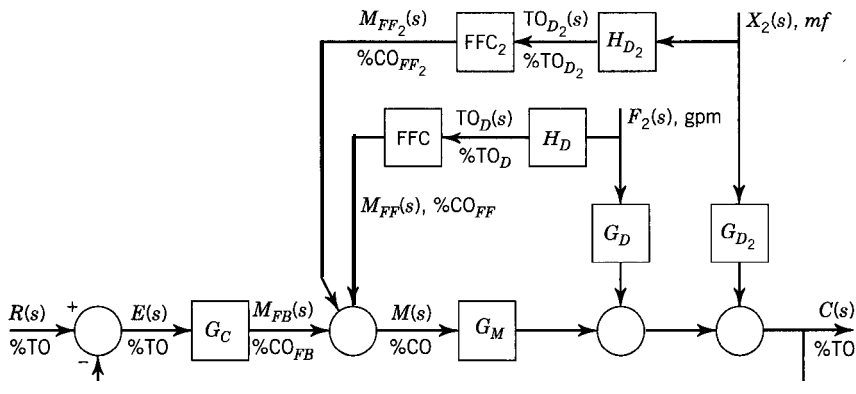$$\text{FFC}_2 = - \frac{\sigma_{D_2}}{H_{D_2}G_M} \tag{12-2.12}$$

Figure 12-2.15 Block diagram of feedforward control for two disturbances.

Step testing yields the following transfer function:

$$G_{D_2} = \frac{64.1e^{-0.85s}}{3.15s + 1} \frac{\% \, TO}{mf} \tag{12-2.13}$$

Assuming that the concentration transmitter in stream 2 has a negligible lag and that it has been calibrated from 0.5 — 1.0 mf, its transfer function is given by

$$H_{D_2} = \frac{100 \, \%TO_{D_2}}{0.5 \, mf} = 200 \, \frac{\%TO_{D_2}}{mf} \tag{12-2.14}$$

Finally, substituting Eqs. 12-2.6, 12-2.13, and 12-2.14 into Eq. 12-2.12 yields

$$FFC_2 = 0.293 \left(-\frac{3.82s + 1}{3.15s + 1} \right) e^{-(0.85 - 0.95)s} \tag{12-2.15}$$

Because the dead time is again negative,

$$FFC, = 0.293 \left(\frac{3.82s + 1}{3.15s + 1} \right) \tag{12-2.16}$$

Figure 12-2.16 shows the implementation of this new feedforward controller added to the previous one and to the feedback controller. Fig. 12-2.17 shows the response $x_6(t)$ to a change of — 0.2 mf in $x_2(t)$ under feedback, steady-state feedforward, and dynamic feedforward control. The improvement provided by feedforward control is certainly significant. Most of the improvement in this case is provided by the steady-state term; arguably, the addition of the lead/lag also provides an improvement. It is a judgment call in this case whether to implement the lead/lag. Note that the ratio of the lead time constant to the lag time constant is 1.25, which is close to 1.0. On the basis of our discussion on the lead/lag term, the closer the ratio is to 1.0, the less the need for lead/lag compensation. Here is a *rule of thumb* that can be used to decide whether to use the lead/lag: If $1.0 < \tau_{ld}/\tau_{lg} < 1.5$ or $0.67 < \tau_{ld}/\tau_{lg} < 1.0$, do not use lead/lag.
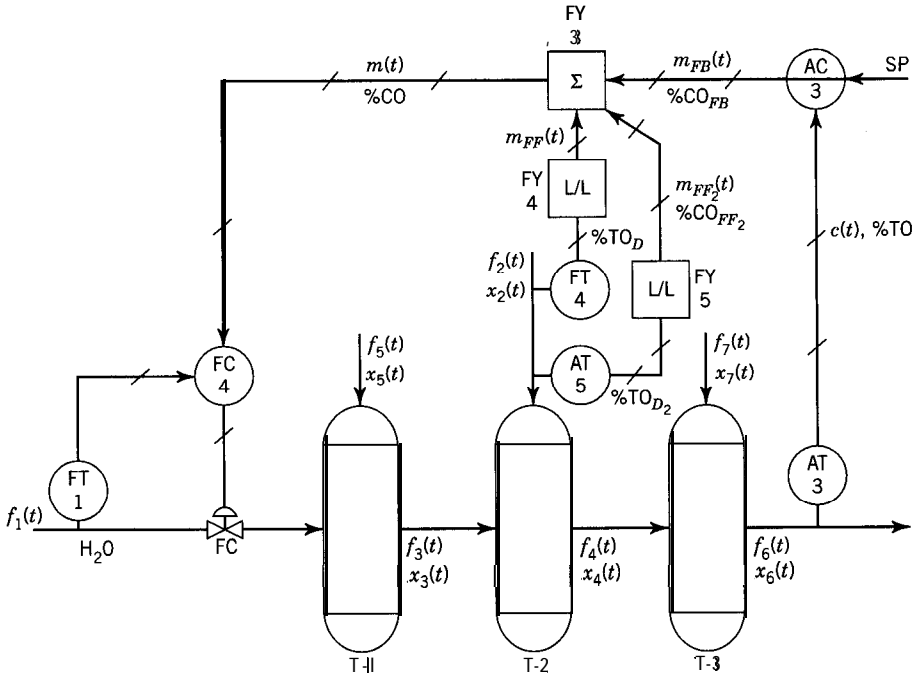
Figure 12-2.16 Implementation of feedforward control for two disturbances.

Outside these limits, the use of lead/lag may significantly improve the control performance.

12-2.5 Design of Nonlinear Feedforward Controllers from Basic Process Principles

The feedforward controllers developed thus far, Eqs. 12-2.9 and 12-2.16, are linear controllers. They were developed from linear models of the process that are valid only
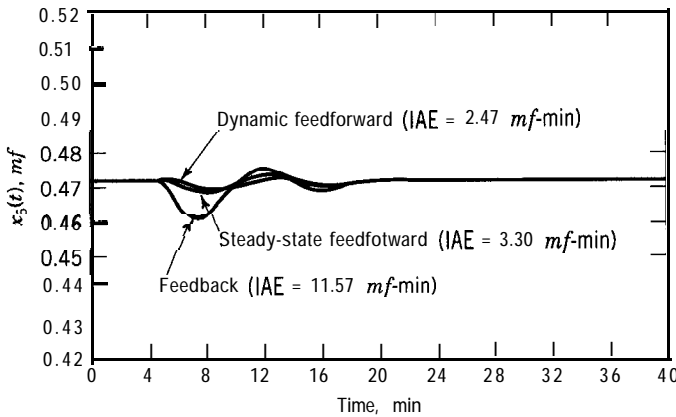

Figure 12-2.17 Comparison of feedback and feedforward control to a change of -0.2 mf in $x_2(t)$.

for small deviations around the operating point where the step tests were performed. These controllers are then used with the same constant parameters without consideration of the operating conditions. As we noted in Chapters 3 and 4, processes most often have nonlinear characteristics, so as operating conditions change, the control performance provided by linear controllers degrades.

As shown in Section 12-2.2, the feedforward controllers are composed of **steady-state** and dynamic compensators. Very often the steady-state compensator, represented by the $-\dfrac{K_D}{K_{T_D}K_m}$ term, can be obtained by other means that yield a nonlinear **compensator,** thus providing an improved control performance over a wide range of operating conditions.

A method for obtaining the nonlinear steady-state compensator consists of starting from first principles, usually mass or energy balances. Using first principles, it is desired to develop an equation that provides the manipulated variable as a function of the disturbances and of the set point of the controlled variable. That is,

$$m(r) = f[d_1(t), d_2(t), \ldots, d_n(t), \text{set point}]$$

For the process at hand,

$$f_1(t) = f[f_5(t), x_5(t), f_2(t), x_2(t), f_7(t), x_7(t), x_6^{\text{set}}(t)]$$

where $x_6^{\text{set}}(t)$ is the desired value of $x_6(t)$.

In the previous section we decided that for this process, the major disturbances are $f_2(t)$ and $x_2(t)$ and that the other inlet flows and compositions are minor disturbances. Thus we need to develop an equation, the steady-state feedforward controller, that expresses the manipulated **variable** $f_1(t)$ in terms of the **disturbances** $f_2(t)$ and $x_2(t)$. In this equation, we consider all other inlet flows and compositions at their steady-state values. That is,

$$f_1(t) = f[\bar{f}_5, \bar{x}_5, f_2(t), x_2(t), \bar{f}_7, \bar{x}_7, x_6^{\text{set}}(t)]$$

where the bar indicates the steady-state values of the variables.

Because we are dealing with compositions and flows, mass balances are the appropriate first principle to use. There are two components, A and water, so we can write two independent mass balances. We start with a total mass balance around the three tanks:

$$\rho\bar{f}_5 + \rho f_1(t) + \rho f_2(t) + \rho\bar{f}_7 - \rho f_6(t) = 0 \qquad \text{(12-2.17)}$$

$$\text{1 eq., 2 unk.}\,[f_1(t), f_6(t)]$$

Note that $f_2(t)$ is not considered an unknown; it will be measured, and thus its value will be known. A mass balance on component A provides the other equation

$$\rho\bar{f}_5\bar{x}_5 + \rho f_2(t)x_2(t) + \rho\bar{f}_7\bar{x}_7 - \rho f_6(t)x_6^{\text{set}}(t) = 0: \qquad \text{(12-2.18)}$$

$$\textbf{2}\ \text{eq., 2 unk.}$$

Since $x_2(t)$ will also be measured, it is not considered an unknown. Solving for $f_6(t)$ from Eq. 12-2.17, substituting into Eq. 12-2.18, and rearranging yield

$$f_1(t) = \frac{1}{x_6^{set}(t)} [\bar{x}_5 \bar{f}_5 + \bar{x}_7 \bar{f}_7] - \bar{f}_5 - \bar{f}_7 + \frac{1}{x_6^{set}} [x_2(t) - x_6^{set}(t)] f_2(t) \quad \textbf{(12-2.19)}$$

Substituting the steady-state values into Eq. 12-2.19 yields

$$f_1(t) = \frac{1}{x_6^{set}(t)} [850 + f_2(t)x_2(t)] - f_2(t) - 1000 \quad \textbf{(12-2.20)}$$

Equation 12-2.20 is the desired steady-state feedforward controller.

The implementation of Eq. 12-2.20 depends on how the feedback correction, the output of the feedback controller, is implemented. This implementation depends on the physical significance given to the feedback signal; there are several ways to decide this. One way is to decide that the significance of the feedback signal is $\Delta f_1(t)$ and to use a summer similar to that shown in Fig. 12-2.16. In this case, we first substitute $x_6^{set} = 0.472$ into Eq. 12-2.20 and obtain

$$f_1(t) = 800.85 + f_2(t) \left[\frac{x_2(t)}{0.472} - 1 \right]$$

This equation is written in engineering units. Depending on the control system being used, the equation may have to be scaled before it is implemented. Assuming that this is done, if needed, Fig. 12-2.18 shows the implementation of this controller; a multiplier only is needed, with no dynamic compensation.

The second way to implement the feedback compensation is by deciding that the significance of the feedback signals is $1/x_6^{set}$. This signal is then input into Eq. 12-2.20 to calculate $f_1(t)$. Thus in this case, the feedback signal is used directly in the feedforward calculation, not to bias it; Fig, 12-2.19 shows the implementation of this controller. The figure shows only one block, referred to as CALC. The actual number of computing blocks required to implement Eq. 12-2.20 depends on the control system used. As an exercise, the reader may determine how many blocks are required when Table 11-1.1 is used and how many are required with Table 1 1-1 .2.

Figure 12-2.20 shows the response of the process under feedback controller, linear steady-state feedforward, and the two nonlinear steady-state feedforward controllers to disturbances of a 1000-gpm increase in $f_2(t)$ and a 0.2-mf change in $x_2(t)$. The improvement in control performance obtained with the nonlinear controllers is obvious. The performance obtained with the second nonlinear controller is quite impressive. This controller better maintains the nonlinearity characteristics of the process and can provide better control.

Instead of calling the output of the feedback controller $1/x_6^{set}$, we could call it x_6^{set}; the control performance would be the same. But what about the action of the feedback controller in both cases? Think about it.

The previous paragraphs have shown two different ways to implement the nonlinear steady-state feedforward controller depending on the significance given to the feedback signal, $\Delta f_1(t)$ or $1/x_6^{set}$. The designer has complete freedom to make this decision. In the
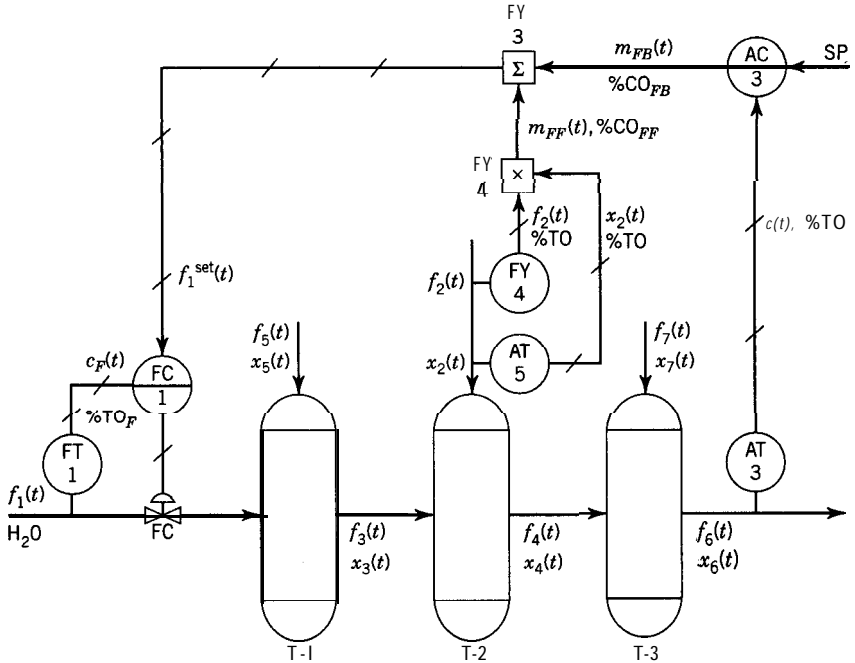
Figure 12-2.18 Implementation of nonlinear steady-state feedforward control.
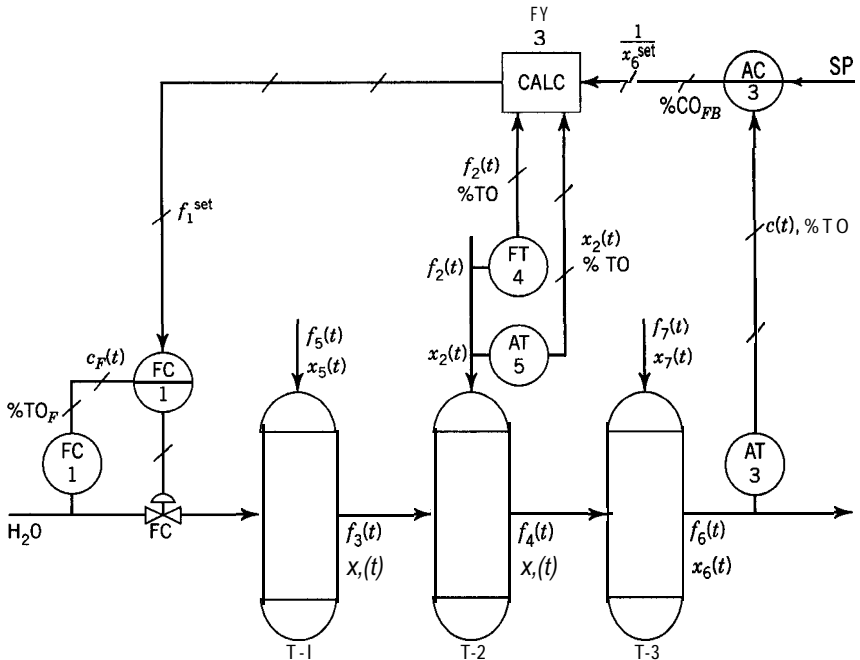

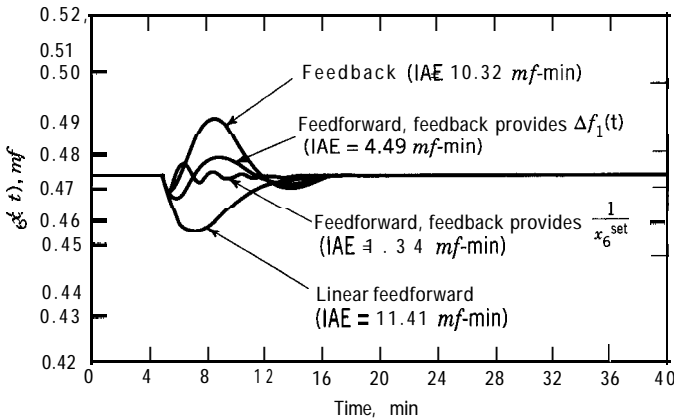**Figure 12-2.19** Implementation of nonlinear steady-state feedforward control.

Figure 12-2.20 Response of mixing process under feedback, linear steady-state, and nonlinear steady-state feedforwrd control to a disturbance of a 1000 gpm change in $f_2(t)$ and a -0.2 mf change in $x_2(t)$.

first case, the feedback controller biased the feedforward calculation. This is a simple and valid choice, and it is the one generally used when the steady-state part of the controller is obtained as in Section 12-2.2. The second choice, $1/x_6^{set}$, or x_6^{set}, is also a simple choice that helps maintain the nonlinearity of the controller. Please note that the actual desired value of x,(t) is the set point to the feedback controller. The controller changes the term $1/x_6^{set}$, or x_6^{set}, in the feedforward equation to keep its own set point. Whenever possible, the feedback signal should adjust the set point of the feedforward controller.

Sometimes, developing a nonlinear steady-state compensator from first principles is just too difficult. Fortunately, process engineering tools provide yet another way to develop the controller. Processes are usually designed by either steady-state flowsheet simulators or any other steady-state simulation. These simulators, along with regression analysis tools, provide another means of designing the steady-state compensator. The simulation can be run at different conditions-that is, different $f_2(t)$, $x_2(t)$, and x_6^{set} — and the required manipulated variable, $f_1(t)$, can be calculated to keep the controlled variable at set point. This information can then be fed to a multiple regression program to develop an equation relating the manipulated variables to the disturbances and set point.

12-2.6   Some Closing Comments and Outline of Feedforward Controller Design

Before proceeding with more examples, we want to make the following comments about the process and example presented in this section, and about feedforward control in general.

1. The first comment refers to the process itself. Figure 12-2.13 shows the response of the control system when $f_2(t)$ changes from 1000 gpm to 2000 gpm. The composition of this stream is quite high (0.99 mf), so this change tends to increase the composition of outlet stream 6. However, the response shown in Fig. 12-2.13 shows that the composition $x_6(t)$ first tends to decrease and then increases. This behavior is an inverse response of the type presented in Section 4-4, and, of course, there is an explanation

for this behavior. Because the tanks are constant-volume tanks, an increase in $f_2(t)$ results in an immediate increase in $f_4(t)$. The composition of stream 4, which enters the third tank, is less than the composition of stream 6, which exits the third tank. Thus this increase in $f_4(t)$ tends initially to dilute-decrease-the composition $x_6(t)$. Eventually, the increase in $f_2(t)$ results in an increase in the composition entering the third tank and a corresponding increase in $x_6(t)$. The transfer function relating $f_2(t)$ and $x_6(t)$ should show a negative zero; see Problem 4-4. Figure 12-2.13 shows that the response under feedforward control exhibits a more pronounced inverse response. What happens is that when $f_2(t)$ increases, f,(t) is also increased by the feedforward controller. Thus the total flow to the third tank increases more, and the dilution effect in that tank is more pronounced. Can you explain why the inverse response is more pronounced under dynamic feedforward than under steady-state feedforward?

2. The second comment refers to the lead/lag term. The lead/lag unit is a simple unit used to implement the dynamic compensation for the linear and nonlinear feedforward controllers. We showed how to tune the lead/lag, obtaining τ_{ld} and τ_{lg}, on the basis of step testing the process. This method gives an initial tuning for the unit. But what if the step testing cannot be done? How do we go about tuning the unit? Obviously, a good dynamic simulation can provide the required tunings. However, when this simulation is not available, what we learned in Section 12-2.3 and Section 2-4.5 suggests that we can provide some guidelines to answer these questions. Figure 2-4.6 shows the response of a lead/lag unit to a ramp input. Note in Fig. 2-4.6 that the amount of time the output lags the input depends on the *net lag,* defined as $\tau_{lg} - \tau_{ld}$. The amount of time the output leads the input depends on the *net lead,* defined as $\tau_{ld} - \tau_{lg}$. The response of the lead/lag unit to a ramp input is important, because disturbances usually look more like ramps than like steps. Some tuning guidelines follow.

- If you need to lag the input signal, set the lead to zero and select a lag. The lead won't make much difference; it is the net lag that matters.
- If you need to lead the input signal, concentrate on the net lead term. However, you must also choose a lag.
- From the response of the lead/lag unit to a step change in input, it is clear that if $\tau_{ld} > \tau_{lg}$, it amplifies the input signal. For noisy signals, such as flow, do not use ratios greater than 2.
- Because dead time just adds to the lag, a negative dead time would effectively decrease the net lag if it could be implemented. Thus we could decrease the lag in the lead/lag unit by the positive dead time. That is,

$$\tau_{lg} \text{ to be used } = \ \tau_{lg} \text{ calculated } + (t_{0_D} - t_{0_M})$$

Alternatively, we could increase the lead in the lead/lag unit by the negative of the dead time. That is,

$$\tau_{ld} \text{ to be used } = \ \tau_{ld} \text{ calculated } - (t_{0_D} - t_{0_M})$$

- If significant dead time is needed, then use a lag, with no lead, and a dead time. It would not make sense to delay the signal and then lead it, even if the transfer functions called for it.

3. The third comment also refers to the lead/lag unit-specifically, to the location of the unit when multiple disturbances are measured and used in the feedforward controller. If linear compensators are implemented, all that is needed is a single lead/lag unit with adjustable gain for each input. The outputs from the units are then added in the summer, as shown in Fig. 12-2.16. When dynamic compensation is required with nonlinear steady-state compensators, the individual lead/lag units should be installed just after each transmitter-that is, on the inputs to the steady-state compensator. This permits the dynamic compensation for each disturbance be implemented individually. It would be impossible to provide individual dynamic compensations after the measurements were combined in the steady-state compensator.

4. The fourth comment refers to the steady-state portion of the feedforward controller. This section demonstrated the development of a linear compensator and of a nonlinear compensator. The nonlinear compensator has shown better performance. Often it is easy to develop this nonlinear compensator by using first principles, or a steady-state simulation. If the development of a nonlinear compensator is possible, this is the preferred method. However, if this development is not possible, a linear compensator can be set up with a lead/lag unit, with adjustable gain, for each input, and a summer. The adjustable gain can be tuned either by the step test method presented in Section 12-2.2 or by trial and error if needed. Which method to use depends on the process.

5. The fifth comment refers to the comparison of feedforward control to cascade and ratio control; all three of these techniques take corrective action before the controlled variable deviates from set point. Feedforward control takes corrective action *before, or* at the *same time* as, the disturbance enters **the** process. Cascade control takes corrective action before the primary controlled variable is affected but after the disturbance has entered the process. The block diagrams of Fig. 12-2.12 and 10-2.2 graphically illustrate these differences. Figure 12-2.10 shows the implementation of feedforward control only-that is, with no feedback compensation. Interestingly, this scheme is similar to the ratio control scheme shown in Fig. 12-1.2*a*. The ratio control scheme does not have dynamic compensation, but the ratio station in Fig. 12-1.2*a* serves the same function as the gain unit shown in Fig. 12-2.10. Thus ratio control is the simplest form of feedforward   control.

6. Finally, an outline of the different steps in designing a feedforward control strategy should be useful. The following eight steps can serve as a design procedure (Corripio, 1990).

Step 1. State the control objective; that is, define which variable needs to be controlled and what its set point is. The set point should be adjustable by the operator, not a constant.

Step 2. Enumerate the possible measured disturbances. Which disturbances can be easily measured? How much and how fast should each be expected to vary? Have an idea of the cost and maintenance of each sensor. Knowing the answers to these questions should help designers decide which disturbance(s) will be considered major and will thus be compensated for by the feedforward controller.

Step 3. State which variable is going to be manipulated by the feedforward controller. In a cascade arrangement, wherein the feedforward controller is cascaded to a slave controller, the manipulated variable is the set point to the slave controller. This makes sense, because the feedforward controller manipulates **the** set point to the slave controller. Such a case was presented in this section.

Step 4. Now you are ready to design the feedforward controller. The feedforward controller consists of two parts: steady-state and dynamic compensators. Develop the steady-state compensator first: specifically, a nonlinear compensator using first principles, or an existing steady-state simulation. The compensator should be an equation such that the manipulated variable, identified in step 3, can be calculated from the measured disturbances, identified in step 2, and the control objective (set point), identified in step 1. Keep the equation as simple as possible. If the steady-state nonlinear compensator cannot be developed by any of the methods mentioned, then use the procedure developed in Section 12-2.2, step testing.

Step 5. At this point, reevaluate the list of disturbances. If a nonlinear compensator has been developed, it may help in the reevaluation. The effect of a disturbance on the controlled variable can be calculated from the equation. A disturbance that was not in the original list may appear in the equation and may be important. The final decision as to which disturbance to compensate for using feedforward depends on its effect on the controlled variable, on the frequency and magnitude of variation, and on the capital cost and maintenance of the sensor. Unmeasured disturbances can be treated as constants at their steady-state or expected values.

Step 6. Introduce the feedback compensation. The way feedback is introduced depends on the physical significance assigned to the feedback signal.

Step 7. Decide whether dynamic compensation, lead/lag, and/or dead time is required, and decide how to introduce it into the design.

Step 8. Draw the instrumentation diagram for the feedforward strategy. The details of the diagram depend largely on the control system being used. A good design should be able to continue to operate safely when some of its input measurement fails. This characteristic of the design is known as "graceful degradation."

12-2.7 Three Other Examples

EXAMPLE 12-2.1

This example presents the control of the process shown in Fig. 12-2.21. The process is similar to the one presented in Section 12-2.2. However, it is dissimilar enough to require a different feedforward controller, principally in its dynamic compensation; Table 12-2.2 presents the steady-state values and other information. It is desired to maintain the outlet composition $x_6(t)$ at 0.472 mass fraction (**mf**) of component A. Any flow or composition entering the process is a possible disturbance. However, operating experience has shown that the flow of stream $2, f_2(t)$, is the major upset; the stream can double its flow from 1000 gpm to 2000 gpm almost instantaneously. This upset occurs when another process upstream comes on line. The flow of water, $f_1(t)$, is the manipulated variable.

Figure 12-2.22 shows the process response to this upset when feedback control is used. This response shows that $x_6(t)$ deviates from its set point of 0.472 mf to 0.487 mf, a 3.18% change. However, this process requires a tighter quality control. Feedforward can be used to minimize the effect of the disturbance.

We have already completed the first three steps of the design procedure outlined in the previous section. That is, we have stated the control objective (step 1), we have
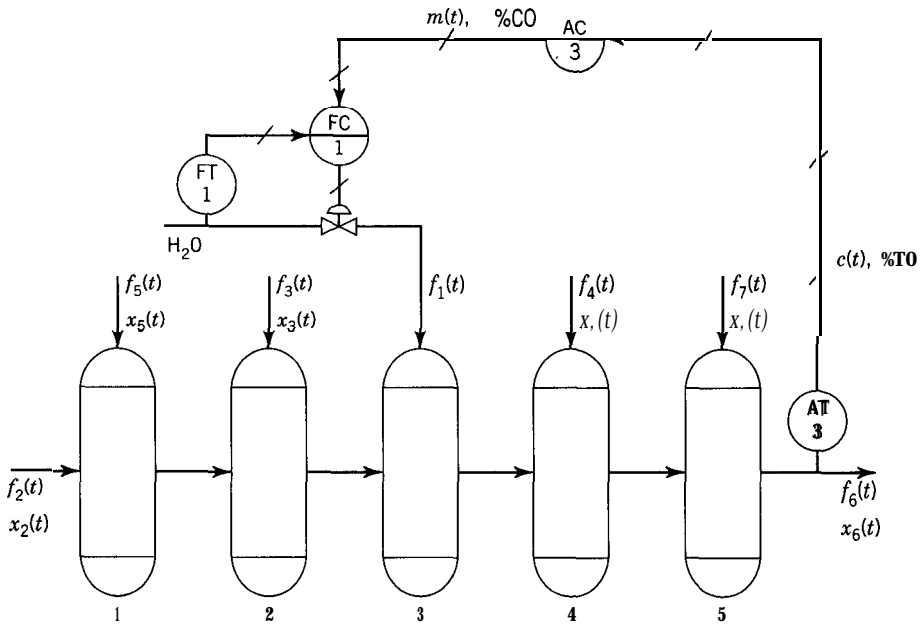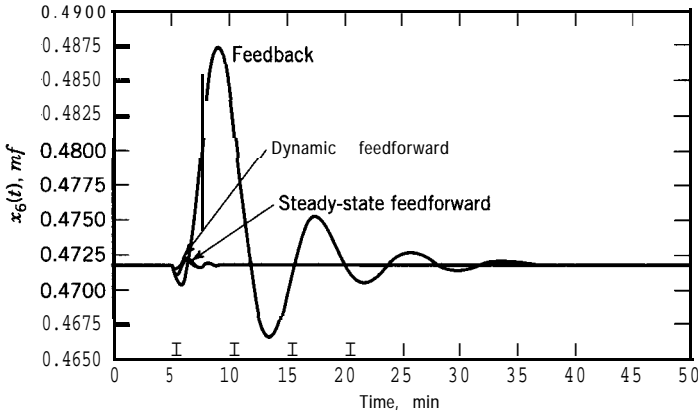
Figure 12-2.21 Mixing process for Example 12-2.1.

enumerated the disturbances and chosen the major one (step 2), and we have stated the manipulated variable (step 3). We now proceed to design the feedforward controller (step 4).

The block diagram for this process is identical to the one shown in Fig. 12-2.8. As we saw in previous sections, it is necessary to determine the transfer functions G_D and

Table 12-2.2 Process Information and Steady-State Values for Mixing Process of Example 12-2.1

Information
Concentration transmitter range: 0.3-0.7 mass fraction. Its dynamic can be described by a time constant of 0.1 min.
The pressure drop across the valve can be considered constant, and the maximum flow provided by the valve is 8000 gpm. The valve dynamics can be described by a time constant of 0.1 min.
The densities of all streams are also considered similar and constant.

Steady-State Values

Stream	Flow, gpm	Mass Fraction	Stream	Flow, gpm	Mass Fraction
1	3983	0.000	5	500	0.800
2	1000	0.990	6	8483	0.472
3	1100	0.850	7	500	0.900
4	1500	0.875			

Figure 12-2.22 Process response under feedback, nonlinear steady-state feedforward, and dynamic (dead time only) feedforward control--f,(t) changed from 1000 gpm to 2000 gpm.

G_M to design the feedforward controller. G_D is determined by step changing $f_2(t)$ and recording $x_6(t)$. Using this procedure, we obtain the transfer function

$$G_D = \frac{0.014e^{-2.2s}}{3.6s + 1}, \frac{\% \text{TO}}{\text{gpm}} \tag{12-2.21}$$

G_M is determined by step changing the controller's output and recording $x_6(t)$. The following transfer function is obtained.

$$G_M = \frac{-1.065e^{-1.2s}}{3.15s + 1}, \frac{\% \text{TO}}{\% \text{CO}} \tag{12-2.22}$$

Because the flow transmitter for $f_2(t)$ has a range of 0 to 3000 gpm and has negligible dynamics, its transfer function is given by

$$H_D = 0.0333 \frac{\% \text{TO}_D}{\text{gpm}} \tag{12-2.23}$$

Using Eqs. 12-2.22, 12-2.23, and 12-2.24, we design the following feedforward controller:

$$\text{FFC} = 0.3947 \left(\frac{3.15s + 1}{3.60s + 1} \right) e^{-1.00s} \tag{12-2.24}$$

Even though a steady-state compensator is available (0.3947) from the process testing, let us develop a nonlinear steady-state compensator starting from first principles. This compensator, in engineering units, is

$$f_1(t) = \frac{1}{x_6^{\text{set}}(t)} [3012.5 + 0.99f_2(t)] - 3500 - f_2(t) \tag{12-2.25}$$
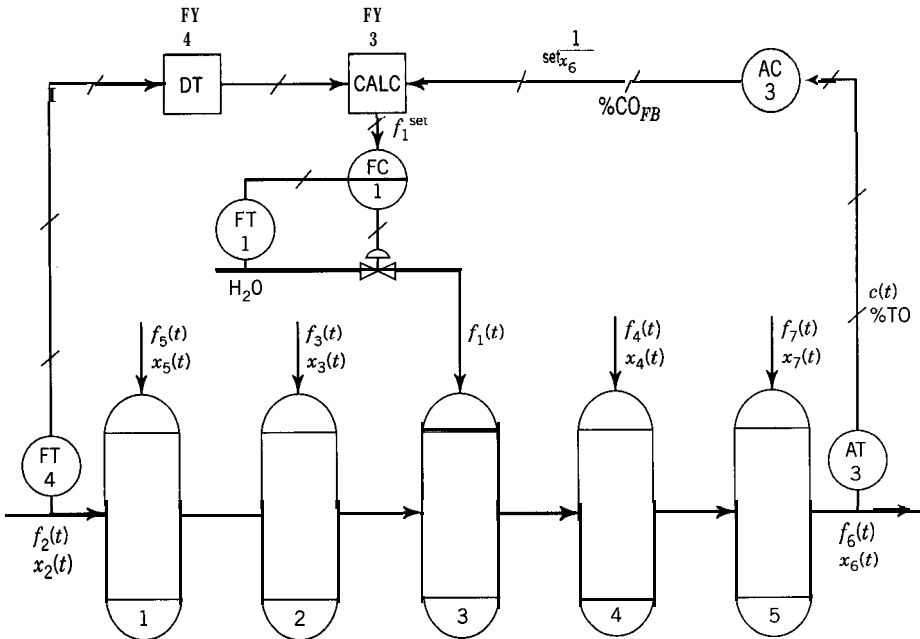
Figure 12-2.23 Implementation of feedforward control.

The feedback compensation will be used to adjust the $1/x_6^{set}(t)$ in Eq. 12-2.25. Choosing how the feedback is to be introduced is step 6 in the procedure.

Step 7 is deciding on the dynamic compensation. In this controller, the dynamic compensation terms, lead/lag and dead time, indicate that the action of the manipulated variable should be delayed. The ratio of the lead time constant to the lag time constant is less than 1, and the dead time is positive. However, the lead/lag term in Eq. 12-2.24 indicates that $\tau_{ld}/\tau_{lg} = 3.15/3.6 = 0.875$. The ratio is relatively close to 1, and applying the rule of thumb presented at the end of Section 12-2.4, we can neglect it. Furthermore, the dead time itself does not seem very large when compared with the time constants. This may be an indication that we can also neglect it. As an exercise, we will leave it in the controller to test its contribution. Figure 12-2.23 shows the implementation of this feedforward controller (step 8). The block referred to as CALC implements Eq. 12-2.25.

Figure 12-2.22 compares the control performance provided by feedback control, by nonlinear steady-state feedforward control, and by dynamic (nonlinear steady-state and dead time) feedforward control. Obviously, the addition of a dead time to the feedforward control does not improve the control performance in this case, so it can be neglected.

EXAMPLE 12-2.2

An interesting and challenging process is control of the liquid level in a boiler drum. Figure 12-2.24 shows a schematic of a boiler drum. Control of the level in the drum is very important. A high level may result in carrying over of water, and perhaps impu-
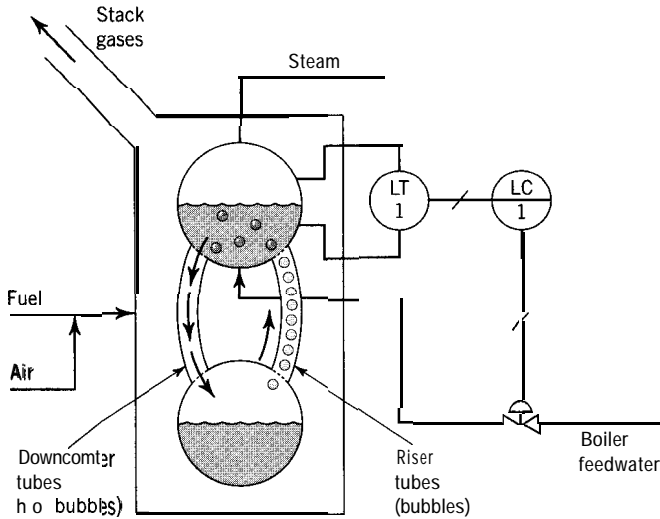
Figure 12-2.24 "Single-element" control in a boiler drum.

rities, into the steam system; a low level may result in tube failure due to overheating as a result of lack of water in the boiling surfaces.

Figure 12-2.24 shows steam bubbles flowing upward through riser tubes into the liquid water; this is an important phenomenon. The specific volume (volume/mass) of the bubbles is very large, so these bubbles displace the water. This results in a higher apparent level than the level due to water only. The presence of these bubbles also presents a problem under transient conditions. Consider the situation in which the pressure in the steam header drops because of an increased demand for steam by the users. This drop in pressure results in a certain quantity of water flashed into steam bubbles. These new bubbles tend to increase the apparent level in the drum. The drop in pressure also causes the volume of the existing bubbles to expand, further increasing the apparent level. Such a surge in level resulting from a decrease in pressure is called swell. An increase in steam header pressure, brought about by a decreased demand for steam by the users, has the opposite effect on the apparent level and is called shrink.

The swell/shrink phenomena, combined with the importance of maintaining a good level, makes the level control even more critical. The following paragraphs explain some of the level control schemes presently used in industry.

The drum level control is accomplished by manipulating the flow of feedwater. Figure 12-2.24 shows the simplest type of level control, which is referred to as *single-element control.* A standard differential pressure sensor/transmitter is usually used. Because this control scheme relies only on the drum level measurement, that measurement must be reliable. Under frequent transients the swell/shrink phenomena do not render a reliable measurement, so a control scheme that compensates for these phenomena is required. Single-element control is good for boilers that operate at a constant load.

Two-element control, shown in Fig. 12-2.25, is essentially a **feedforward/feedback** control system. The idea behind this scheme is that the major reason for level changes is changes in steam demand and that for every pound of steam produced, a pound of feedwater should enter the drum; that is, there should be a mass balance. The signal output from **FT5** provides the feedforward part of the scheme, and **LC1** provides the
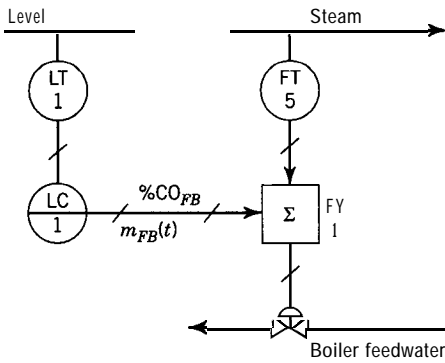
Figure   12-2.25   "Two-element"   control.

feedback compensation for any unmeasured flows such as blowdown. The feedback controller also helps to compensate for errors in the flowmeters.

The two-element control scheme works quite well in many industrial boiler drums. However, there are some systems that exhibit variable pressure drop across the feed-water valve. The two-element control scheme does not directly compensate for this disturbance, and consequently, it upsets drum level control by momentarily upsetting the mass balance. The *three-element control* scheme, shown in Fig. 12-2.26, incorporates the required compensation. This scheme provides a tight mass balance control during transients. It is interesting to note that all that has been added to the two-element control scheme is a cascade control system.

The boiler drum level provides a realistic example wherein the cascade and feedforward control schemes are used to improve the performance provided by feedback control. In this particular example, the use of these schemes is almost mandatory to avoid mechanical and process failures. Every step taken to improve the control was justified.
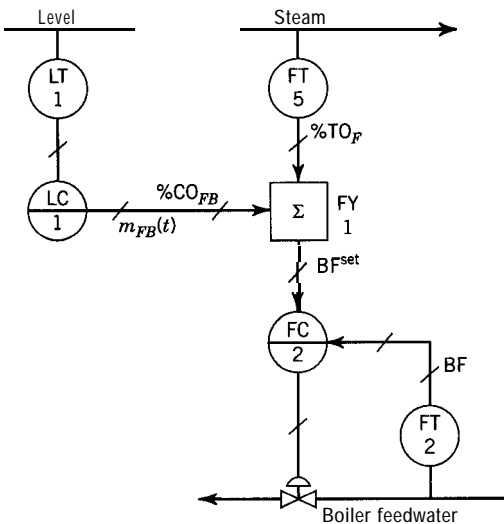
Figure   12-2.26   "Three-element"   control.

Otherwise, there is no need to complicate matters. See **O'Meara** (1979) and Scheib and Russell (198 1) for another complete discussion of this subject.

EXAMPLE 12-2.3

We now present another industrial example that has proved to be a successful application of feedforward control. The example is concerned with temperature control in the rectifying section of a distillation column. Figure 12-2.27 shows the bottom of the column and the control scheme originally proposed and implemented. This column uses two reboilers. One of the reboilers, **R-10B**, uses a condensing process stream as a heating medium, and the other reboiler, **R-10A**, uses condensing steam. For energy-efficient operation, the operating procedure calls for using as much of the process stream as possible. This stream must be condensed anyway, and thus it serves as a "free" energy source. The steam flow is used to control the temperature in the column.

After start-up of this column, it was noticed that the process stream serving as heating medium experienced changes in flow and in pressure. These changes acted as disturbances to the column, and consequently, the temperature controller needed to compensate for these disturbances continually. The time constants and dead time in the column and reboilers complicated the temperature control. After the problem was studied, de-
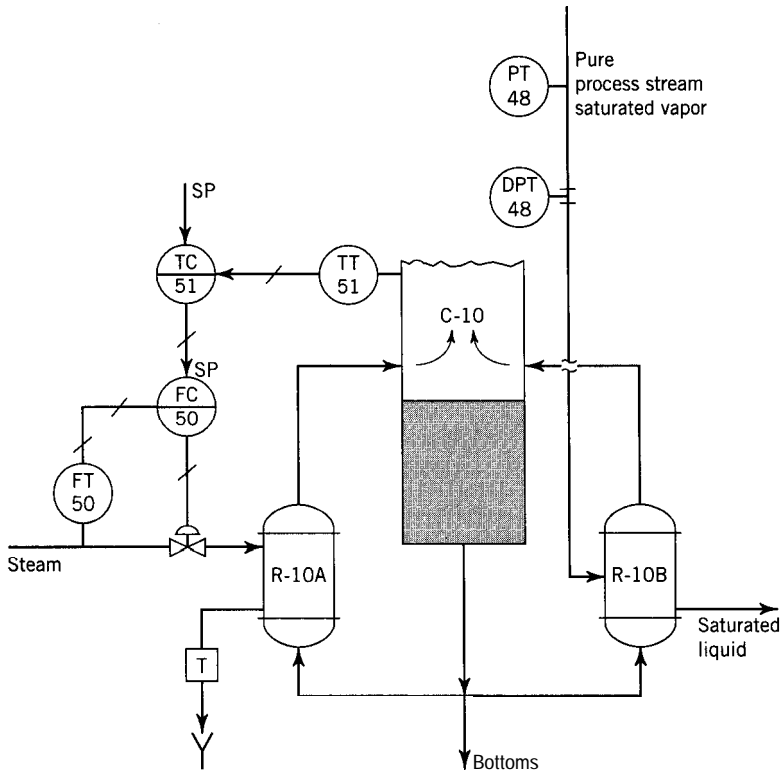
Figure 12-2.27 Distillation column control.

signers decided to use feedforward control. A pressure transmitter and a differential pressure transmitter had been installed in the process stream; from them, the amount of energy given off by the stream in condensing could be calculated. This information made it possible to calculate the amount of steam required to maintain the temperature at set point, and thus corrective action could be taken before the temperature deviated from set point. This is a perfect application of feedforward control.

Specifically, the procedure implemented was as follows: Because the process stream is pure and saturated, the density, ρ, is a function of pressure only. Therefore, the density of the stream can be obtained by using a thermodynamic correlation.

$$P = f_1(P) \tag{12-2.27}$$

From this density and the differential pressure, h, obtained from the transmitter DPT48, the mass flow of the stream can be calculated from the orifice equation

$$w = K_0 \sqrt{h\rho}, \text{ lbm/hr} \tag{12-2.28}$$

where K_0 is the orifice coefficient.

Knowing the stream pressure and using another thermodynamic relation, it is possible to obtain the latent heat of condensation, λ.

$$\lambda = f_2(P) \tag{12-2.29}$$

Finally, multiplying the mass flow rate times the latent heat yields the energy, q_1, given off by the process stream in condensing.

$$q_1 = w\lambda \tag{12-2.30}$$

Figure 12-2.28 shows the implementation of Eqs. 12-2.27 through 12-2.30 and the rest of the feedforward scheme. Block PY48A performs Eq. 12-2.27, block PY48B performs Eq. 12-2.28, block PY48C performs Eq. 12-2.29, and block PY48D performs Eq. 12-2.30. Therefore, the output of relay PY48D is q_1, the energy given off by the condensing process stream.

To complete the control scheme, the output of the temperature controller is considered to be the total energy required, q_t, to maintain the temperature at its set point. Subtracting q_1 from q_t yields the energy required from the steam, q_s.

$$q_s = q_t - 41 \tag{12-2.31}$$

Finally, dividing q_s by the latent heat of condensation of the steam, h_{fg}, yields the required steam flow, w_s.

$$w_s = \frac{q_s}{h_{fg}}, \text{ lbm/hr} \tag{12-2.32}$$

Block TY51 performs Eqs. 12-2.31 and 12-2.32, and its output is the set point to the flow controller FC50. In Eq. 12-2.32, h_{fg} is assumed constant.
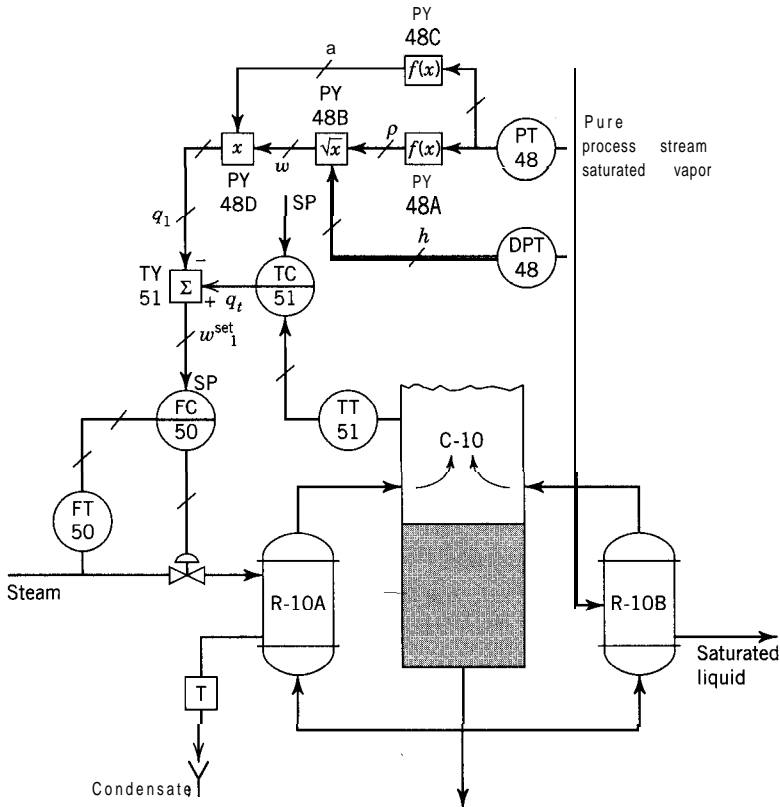
Figure 12-2.28 Implementation of feedforward in distillation column.

Several things must be noted in this feedforward scheme. First, the model of the process is not one equation but several. This model was obtained by using several process engineering principles. This makes process control fun, interesting, and challenging. Second, the feedback compensation is an integral part of the control strategy. This compensation is q_t, or total energy required to maintain set point in TC5 1. Finally, the control scheme shown in Fig. 12-2.28 does not show dynamic compensation. This compensation may be installed later if needed.

12-3 SUMMARY

This section has presented in detail the concept, design, and implementation of feedforward control. The technique has been shown to offer significant improvement over the control performance provided by feedback control. However, you have undoubtedly noticed that the design, implementation, and operation of feedforward control require a significant amount of engineering, extra instrumentation, in-depth understanding, and thorough training of the operating personnel. All of this means that feedforward control is more costly than feedback control and thus must be justified. The reader must also understand that feedforward is not the solution to all control problems. It is another good tool for enhancing feedback control in some cases.

We saw that feedforward control is generally composed of steady-state compensation and dynamic compensation. Not in every case are both compensations needed. The amount of each required compensation depends on the particular process.

There are three means by which to design the steady-state compensator. The best way is based on process engineering principles, usually mass and energy balances. Steady-state simulations-flowsheet simulators-provide still another realistic avenue. Both of these methods were discussed in Section 12-2.5. The process testing method presented in Section 12-2.2 provides a way to design a linear steady-state compensator when neither of the other two methods can be applied. The design of the dynamic compensator-lead/lag and/or dead time-requires dynamic information, which is usually obtained by process testing.

Finally, the reader must remember that *whenever possible, feedforward control should be designed with some amount of feedback compensation.*

REFERENCES

1. Congdon, P. 1981. "Control Alternatives for Industrial Boilers." *InTech,* December.
2. Corripio, A. B. 1990. *Tuning of Industrial Control Systems.* Research Triangle Park N.C.: Instrument Society of America.
3. Dobson, J. G. 1960. "Optimization of Distillation Columns Using Feed-Forward Computer Control." Paper presented at Interkama 1960, Duesseldorf, Germany, October.
4. O'Meara, J. E. 1979. "Oxygen Trim for Combustion Control." *Instrumentation Technology,* March.
5. Scheib, T. J., and T. D. Russell. 1981. "Instrumentation Cuts Boiler Fuel Costs." *Instrumentation and Control Systems,* November.
6. Shinskey, F. G. 1963. "Feedforward Control Applied." *ISA Journal,* November.

PROBLEMS

12-1. Consider the system, shown in Fig. P12-1, to dilute a 50% by mass NaOH solution to a 30% by mass solution. The NaOH valve is manipulated by a controller not shown in the diagram. Because the flow of the 50% NaOH solution can vary frequently, it is desired to design a ratio control scheme to
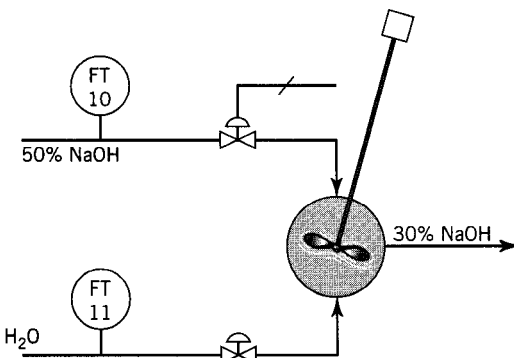
Figure P12-1 Blending process for Problem 12-1.

manipulate the flow of H_2O to maintain the required dilution. The nominal flow of the 50% NaOH solution is 200 lb/h. The flow element used for both streams is such that the output signal from the transmitters is linearly related to the mass flow. The transmitter in the 50% NaOH stream has a range of 0 to 400 lb/h, and the transmitter in the water stream has a range of 0 to 200 lb/h. Specify the computing blocks required to implement the ratio control scheme; use the blocks shown in Table 11-1.2.

12-2. Consider the reactor shown in Fig. P12-2. This reactor is similar to a furnace in that the energy required for the reaction is provided by the combustion of a fuel with air (to simplify the diagram, the temperature control is not completely shown). Methane and steam are reacted to produce hydrogen by the reaction

$$CH, + \ 2H_2O \rightarrow CO, + \ 4H_2$$
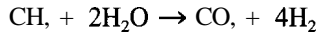
The reaction occurs 'in tubes inside the furnace. The tubes are filled with a catalyst needed for the reaction. It is important that the reactant mixture always be steam-rich to avoid coking the catalyst. That is, if enough carbon deposits over the catalyst, it poisons the catalyst. This situation can be avoided by en- suring that the entering mixture is always rich in steam. However, too much steam is also costly in that it requires more energy (fuel and air) consumption. The engineering department has estimated that the optimum ratio R, (methane to steam) must be maintained. Design a control scheme that ensures that the required ratio will be maintained and that during increases and decreases in production rate, the reactant mixture will be steam-rich. Please note that there is a signal that sets the hydrocarbon flow set point required.
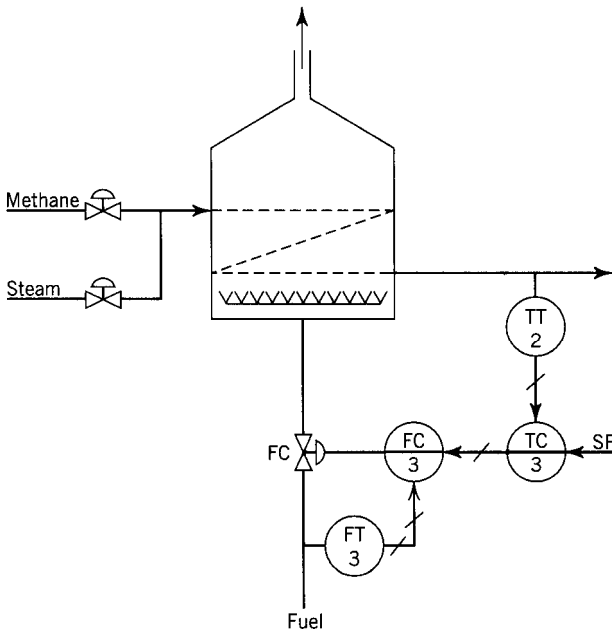
Figure P12-2 Reactor for Problem 12-2.

12-3. Chlorination is used for disinfection of the final effluent of a waste water treat-
ment plant. The EPA requires that a certain chlorine residual be maintained. To
meet this requirement, the free chlorine residual is measured at the beginning
of the chlorine contact basin, as shown in Fig. P12-3. An aqueous solution of
sodium hypochlorite is added to the filter effluent to maintain the free chlorine
residual at the contact basin. The amount of sodium hypochlorite required is
directly proportional to the flow rate of the effluent. The wastewater plant has
two parallel filter effluent streams, which are combined in the chlorine contact
basin. Sodium hypochlorite is added to each stream on the basis of the free
chlorine residual in the basin.

(a) Design a control scheme to control the chlorine residual at the beginning
of the basin.

(b) As a result of a number of reactions occurring in the contact basin, the
chlorine residual exiting the basin is not equal to the chlorine residual en-
tering the basin (the one being measured). It happens that the EPA is in-
terested in the exiting chlorine residual. Thus a second analyzer is added at
the effluent of the contact basin. Design a control scheme to control the
effluent chlorine residual.

12-4. Consider the process shown in Fig. P12-4. In the reactor, the principal reaction
is A + 2B → Product; two other reactions, A + B → Inert and A →
Heavies, also occur but at a lesser rate. All the reactions occur in the gas phase.
Enough cooling is accomplished in the cooler to condense and separate the
heavies. What is left is separated in the separation column. The gases leaving
the column contain A, B, and inerts. The purge is manipulated to maintain the
composition of inerts in the recycle stream at some desired value, 1 mole %. In
the recycle stream there is a temperature transmitter, TT1; a volumetric flow
transmitter, FT3; and two continuous infrared analyzers. One of the analyzers,
AT1, gives the mole fraction of A, y_{AR}, and the other analyzer, AT2, gives the
mole fraction of B, y_{BR}. The process has been designed to have very little
pressure drop between the column and the compressor. The reactants A and B
are pure components and are assumed to be delivered to the valves at some
constant pressure and temperature.

(a) Design a control scheme to control the composition of inerts in the recycle
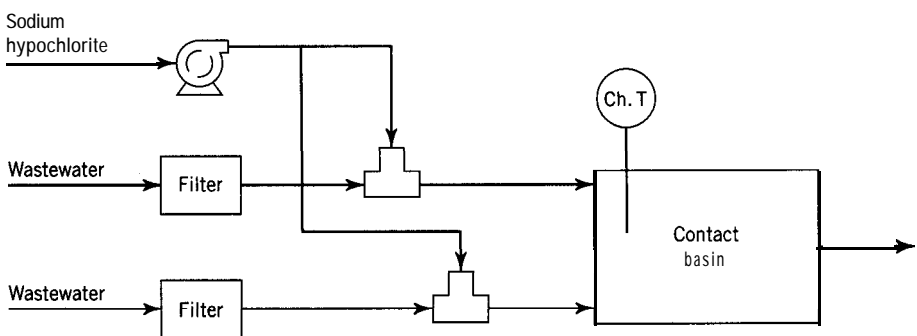stream at 1 mole %.

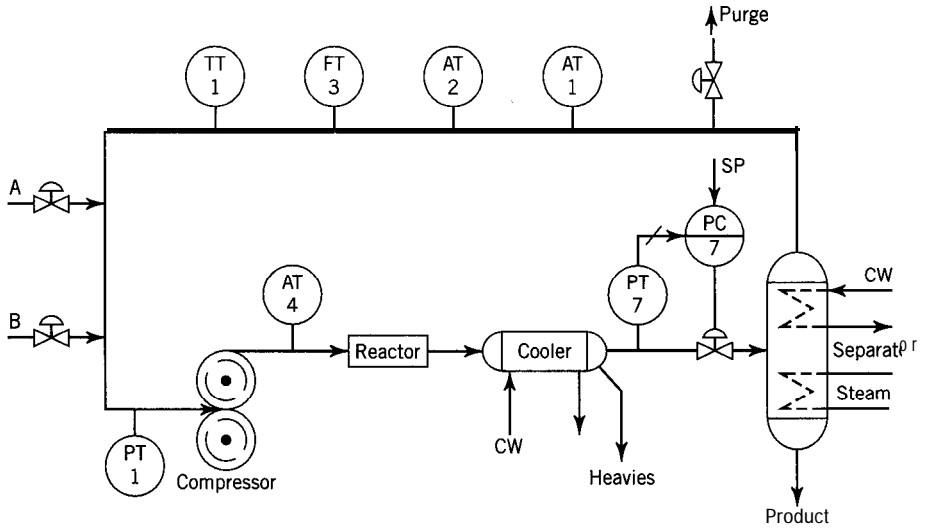Figure P12-3 Chlorine contact basin for Problem 12-3.

Figure **P12-4** Process for Problem 12-4.

(b) Design a control scheme to control the supply pressure to the compressor. It is also very important for the process to maintain the molal ratio of B to A entering the compressor at 2.6. There is one infrared analyzer at the exit of the compressor that provides a signal indicating the mole ratio of B to A.

12-5. Consider the process shown in Fig. **P12-5**. This process is used to manufacture product E from the reaction of A and B. The output from the reactor is product E and some unreacted reactants, mainly A, which are referred to as liquid C. E
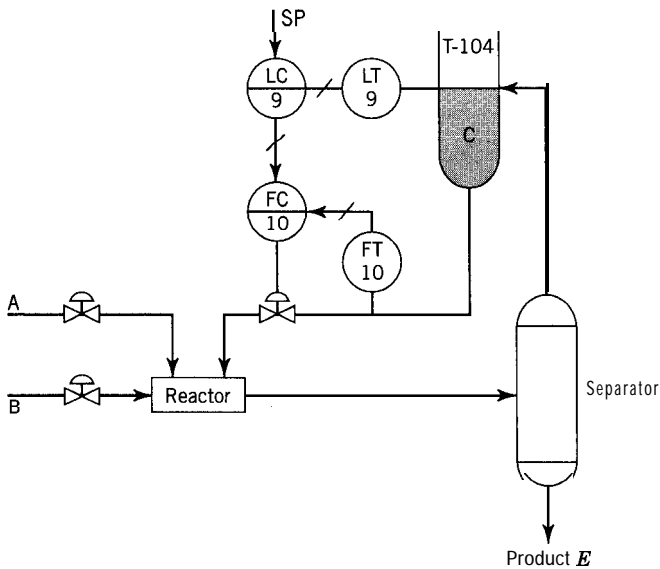
Figure P12-5 Process for Problem 12-5.

and C are separated, and liquid C is recycled back to tank T104 to be fed back to the reactor, as shown in the figure. The amount of B fed to the reactor depends on the amount of A and on the amount of C fed to the reactor. That is, there must be some B to react with the A, given by the ratio $R_1 = B/A$, fed, and some B to react with the C, given by the ratio $R_2 = B/C$, fed. You may assume that all the flowmeters provide a signal related to mass flow. Design a control scheme to control the total flow, T, in lb/min, into the reactor.

12-6. In the production of paper, it is necessary to mix some components in a given proportion to form a stock that will be supplied to a paper machine to produce the final sheet with the desired characteristics. Consider the process shown in Fig. P12-6. For a particular formulation, the final mixture must contain 47 mass % of hardwood slurry, 50 mass % of pine slurry, 2 mass % of additive, and 1 mass % of dye. The nominal system must be designed for a possible maximum production of 2000 lb/h.

 (a) The flowmeters used in this application are magnetic flowmeters, so their output signal is linearly related to the mass flow rate. Specify their range.

 (b) Design a control system to control the level in the blend chest and at the same time maintain the correct formulation. It is also required to know, at the end of the day, the total amount of mass of each stream that has been added to the blend chest. Show, and scale, the necessary instrumentation to implement this control scheme. Use the computing blocks given in Table 11-1.1.

12-7. Fuel cells are used in spacecraft for generating power and heat. The cell produces electric power by the reaction between liquid hydrogen and liquid oxygen:

$$2H_2 + 0_2 \rightarrow 2H_2O$$
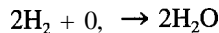
Design a ratio controller to maintain the flows of liquid hydrogen and oxygen into the cell in the exact stoichiometric ratio (both hydrogen and oxygen are valuable in space, so we cannot supply either in excess). Calculate the design flows of hydrogen and oxygen required to produce 0.5 kg/h of water, and give the design ratio of oxygen to hydrogen flow. Sketch a ratio control scheme that
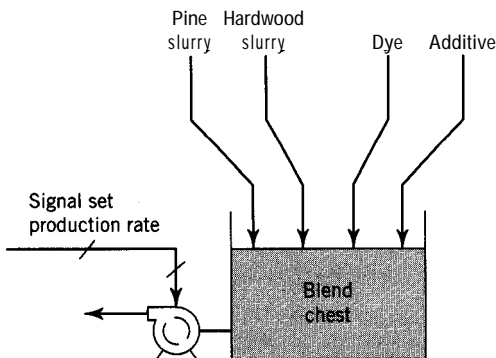

Figure P12-6 Mixing process for Problem 12-6.

will manipulate the flow of oxygen to maintain the exact stoichiometric ratio between the two flows. You may assume that the signals from the flow transmitters are linear with the mass flow rates. Calculate reasonable ranges for the flow transmitters and the ratio in terms of the transmitter signals.

12-8. Consider the furnace shown in Fig. P12-7, which consists of two sections with one common stack. In each section the cracking reaction of hydrocarbons, C_2's and C_3's, with steam takes place. The temperature of the products in each section is controlled by manipulating the fuel to the particular section. The pressure in the stack is controlled by manipulating the speed of a fan installed in the stack. This fan induces the flow of flue gases out of the stack. As the pressure in the stack increases, the pressure controller speeds up the fan to lower the pressure.

   (a) Design a control scheme to ratio the steam flow to the hydrocarbon flow in each section. The operating personnel set the hydrocarbon flow.

   (b) During the last few weeks, the production personnel have noticed that the pressure controller's output is consistently reaching 100%. This indicates that the controller is doing the most possible to maintain pressure control. However, this is not desirable because it means that the pressure is really out of control-not a safe condition. A control strategy must be designed
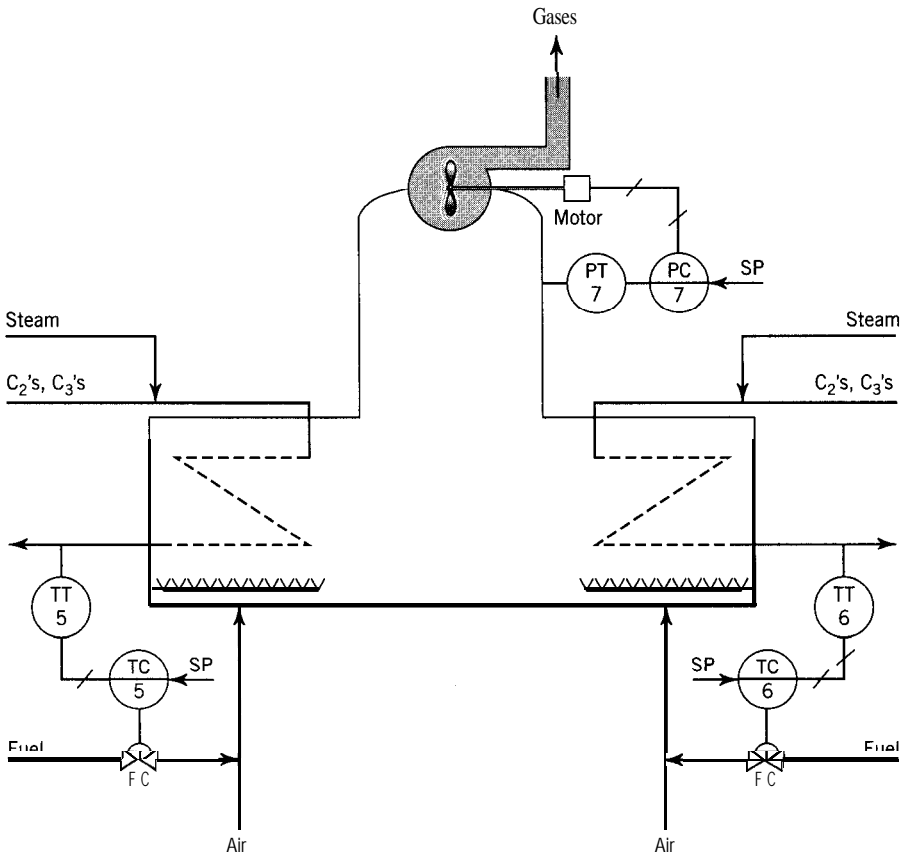
Figure P12-7 Furnace for Problem 12-8.

such that when the speed of the fan is greater than 90%, the flow of hydro-carbons starts to decrease to maintain the fan speed at 90%. As the flow of hydrocarbons decreases, less fuel is required to maintain exit temperature. This in turn reduces the pressure in the stack, and the pressure controller slows down the fan. Whenever the speed is less than 90%, the feed of hydrocarbons can be whatever the operating personnel require.

It is known that the left section of the furnace is less efficient than the right section. Therefore, the correct control strategy to reduce the flow of hydrocarbons calls for reducing the flow to the left section first, up to 35% of the flow set by the operating personnel. If further reduction is necessary, the flow of hydrocarbons to the right section is reduced, also up to 35% of the flow set by the operating personnel. (If even further reduction is nec-essary, an interlock system would then drop off line the furnace.) Design the control strategy to maintain the fan speed below 90%.

(c) If the flow of hydrocarbons changes, the outlet temperature will deviate from set point, and the feedback controller will have to react to bring the temperature back to set point. This situation seems a natural for feedforward control. Design this strategy for each section.

12-9. Consider the process shown in Fig. P12-8. Mud is brought into a storage tank, T-77, from which it is pumped to two filters. The level in the tank is controlled by LIC-1 by manipulating the exit flow. This flow must be split between the two filters in the following known ratio:

$$R = \frac{\text{flow to filter 1}}{\text{total flow}}$$

The two flow transmitters and control valves shown in the figure cannot be moved from their present locations, and no other transmitters or valves can be added. Design a control system that provides the desired exit flow by the level controller, while maintaining the desired flow split between the two filters.
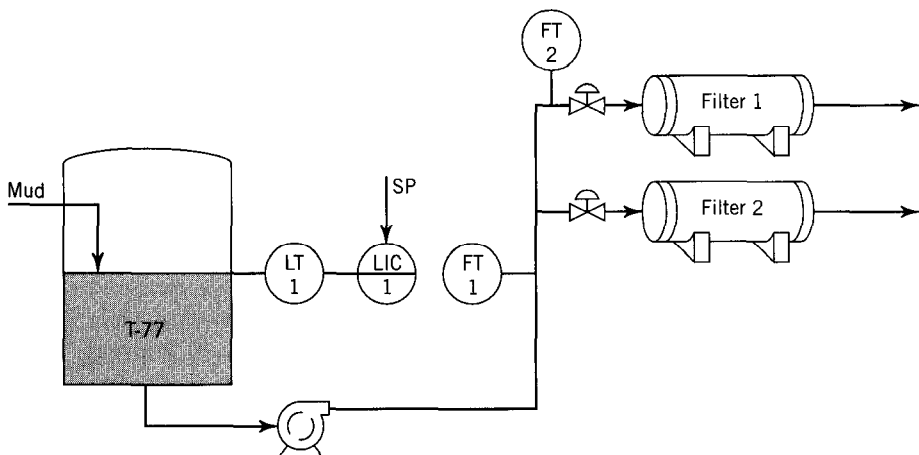

Figure P12-8 Filtering process for Problem 12-9.

12-10. Consider the furnace of Fig. P12-9, wherein two different fuels, a waste gas and fuel oil, are manipulated to control the outlet temperature of a process fluid. The waste gas is free to the operation, so it must be used to full capacity. However, environmental regulations dictate that the maximum waste gas flow be limited to one-quarter of the fuel oil flow. The heating value of the waste gas is HV_{wg}, and that of the fuel oil is HV_{oil}. The ratio of air to waste gas is R_{wg}, and the ratio of air to fuel oil is R_{oil}.

(a) Design a cross-limiting control scheme to control the furnace product temperature, neglecting variation of the heating values of the combustibles.

(b) Assume now that the heating value of the waste gas varies significantly as the composition varies. It is difficult to measure on-line the heating value of this gas. However, laboratory analysis has shown that there is definitely a correlation between the density of the gas and its heating value. There is a densitometer available to measure the density, so the heating value is known. Adjust the control scheme design in part (a) to consider variations in HV,, .

(c) For safety reasons, it is necessary to design a control scheme such that in case of loss of burner flame, the waste gas and fuel oil flows cease; the air valve must open wide. Available for this job is a burner switch whose output is 20 mA as long as the flame is present and whose output drops to 4 mA as soon as the flame stops. Design this control scheme into the previous one.

12-11. Consider the kiln drier sketched in Fig. P12-10. A slurry is fed to the drier and directly contacted with gases from the combustion of fuel and air. The contact of the solid with the hot gases vaporizes the water, the heat of vaporization
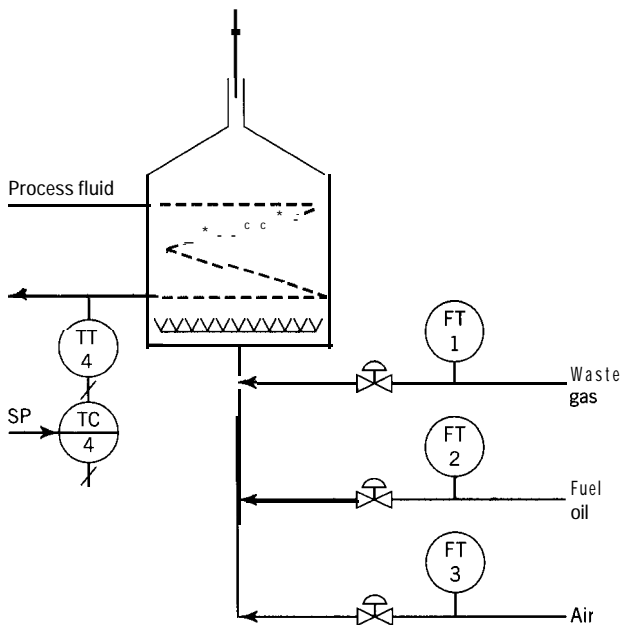
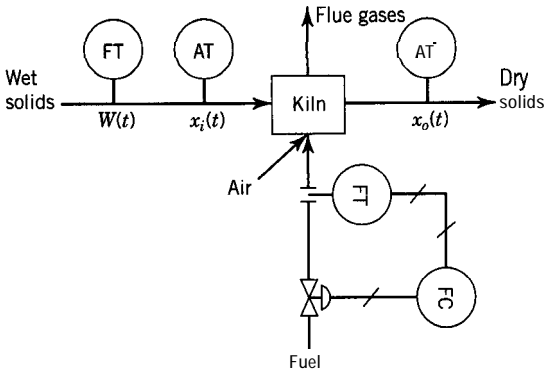**Figure P12-9** Furnace for Problem 12-10.

Figure P12-10 Kiln drier for Problem 12-11.

being provided by the hot gases. It is desired to manipulate the flow of the fuel into the drier to maintain the outlet moisture content of the solid at its set point. The major disturbances are the slurry feed flow, $w_i(t)$; its moisture content, $x_i(t)$; and the heating value of the fuel gas. Of these, only the feed flow and its moisture content can be measured, as shown in the figure. The outlet moisture content of the solid, $x_0(t)$, can also be measured. The design conditions and process parameters are

Slurry feed flow	100 lb/h
Slurry feed composition	0.60 mf of water
Desired outlet moisture content	0.05 mf of water
Design fuel gas flow	80 scfh

The flow transmitter signals are linear with flow and compensated for temperature and pressure. Design a feedforward controller with feedback trim for the drier. Specify which disturbances are measured, what the control is, and which variable is manipulated. Derive the following equation by combining the total and water mass balances on the drier.

$$w_i(t)x_i(t) = [w_i(t) - k_v f_g(t)]x_o(t) + k_v f_g(t)$$

where $f_g(t)$ is the fuel gas flow in scfh, and k_v represents the pounds of water vaporized per scf of fuel gas. Specify how you plan to introduce dynamic compensation and feedback trim if needed. Draw the instrumentation diagram for the feedforward controller.

12-12. It is desired to design a simple linear feedforward controller to compensate for the effect of changes in feed rate to a distillation column on the overhead product composition. The reflux flow is to be manipulated to control the overhead composition. Two steps tests are applied to the column, one on feed rate and one on reflux flow set point; in each case, the response of the overhead composition is recorded and analyzed. Results of the tests are summarized in the following table.

Step test in:	Process Gain, %TO/(klb/h)	Time Constant, min	Dead Time, min
Feed rate, klb/h	0.8	18.0	**8.0**
Reflux flow, klb/h	1.2	**4.0**	1.0

The feed flow transmitter has a calibrated range of 0 to 50 klb/h, and the reflux flow transmitter has a range of 0 to 100 klb/h. Both transmitters generate signals that are linear with flow.

Draw the block diagram that will show the effect of feed rate and reflux flow on the overhead composition. Approximate the transfer functions with first-order-plus-dead-time models, and show the numerical values of all parameters. Using the block diagram, design a linear feedforward controller to compensate for the feed disturbance on the control objective. Draw the instrumentation diagram for your feedforward controller. Include feedback trim, assuming that an on-line analyzer is installed on the overhead product stream.

12-13. Design a feedforward controller to compensate for the feed flow and temperature of the continuous stirred tank heater of Example 6-1.1. Assume that the feed flow transmitter has a range of 0 to 25 ft^3/min and a negligible time constant and that the inlet temperature transmitter has a range of 60 to 120°F and a time constant of 0.6 min. Identify the control objective, the disturbances, and the manipulated variable. Determine the corrections needed in steam flow for the expected changes in disturbances: a 10 ft^3/min change in feed flow and a 20°F change in feed temperature. Discuss the need for feedback trim, and explain how to incorporate it into your design. Also discuss the need for and form of dynamic compensation. Draw the instrumentation diagram for your design. Scaling is not necessary.

12-14. Consider the control scheme for the solid drying system shown in Fig. P12-11. The major disturbance to this process is the moisture content of the incoming solids. For this disturbance, the control system responds quite slowly. It is desired to implement a feedforward system to improve this control. After some initial work, the following data have been obtained:

Step change in inlet moisture = + 2%

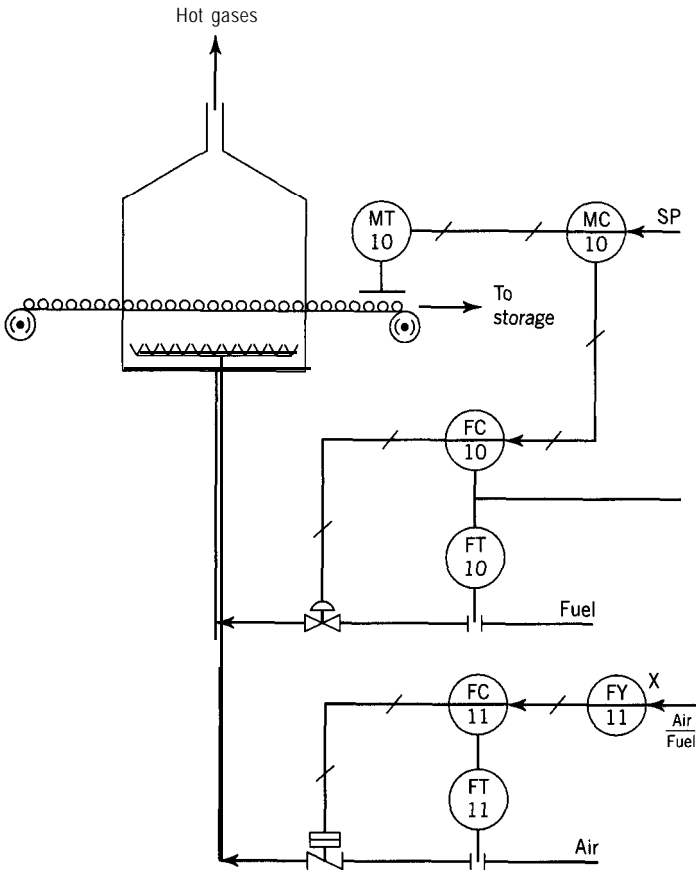| Time, min	Exit Moisture, %	Time, min	Exit Moisture, %
0	**5.0**		
0.5	**5.0**	5.0	**6.6**
1.0	**5.1**	**5.5**	**6.7**
1.5	**5.2**	**6.0**	6.8
2.0	**5.4**	**6.5**	6.9
2.5	**5.7**	**7.0**	7.0
3.0	**5.9**	**7.5**	7.0
3.5	**6.1**	8.0	**6.9**
4.0	**6.3**	**8.5**	7.0

Figure P12-11 Drier for Problem 12-14.

Step change in output signal from moisture controller, MIC-10 = + 25 %CO

Time, min	Exit Moisture, %	Time, min	Exit Moisture, %
0	5.0	5.0	3.81
0.5	5.0	5.5	3.70
1.0	4.95	6.0	3.55
1.5	4.93	6.5	3.45
2.0	4.85	7.0	3.35
2.5	4.70	7.5	3.25
3.0	4.60	8.5	3.10
3.5	4.40	9.5	3.03
4.0	4.20	11.5	3.00

The feedback moisture analyzer has a range of 1% to 7% moisture. There is another analyzer with a range of 10% to 15% moisture that can be used to

measure the inlet moisture. These analyzers have been used before in this par-
ticular process and have proved reliable. Identify the control objective, the dis-
turbances, and the manipulated variable.

(a) On the basis of process engineering principles (mass balances, energy bal-
ances, etc.), develop a feedforward control scheme. The statement of the
problem has not provided all of the necessary information. Assume that you
can obtain this information from the plant's files. Show the implementation
of this scheme.

(b) Draw a complete block diagram for this process that shows the effect of
the inlet moisture on the controlled variable. Include all known transfer
functions.

(c) Develop a feedforward control scheme using the block diagram approach.
Show the implementation of this scheme, including the feedback trim.

12-15. Consider the vacuum filter shown in Problem 7-15. Using the information given
in that problem, design a feedforward/feedback control scheme to compensate
for changes in inlet moisture. You may assume that there is a moisture trans-
mitter with a range of 60% to 95% moisture. Show the complete implementa-
tion, using the computing blocks of Table 1 1- 1.2. Identify the eight steps out-
lined in Section 12-2.6.

12-16. Consider the evaporator system of Problem 7-20. Design a **feedforward/feed-**
back control scheme to compensate for composition changes in the solution
entering the first effect. The sensor to measure inlet composition has a range of
0% to 20% sugar. The data given in the problem are all you need to design a
linear feedforward controller. It is also known that the total evaporation in both
effects is approximately 1.8 lb of water vaporized per pound of steam. Using
this information, design a nonlinear feedforward controller. Use the computing
blocks of Table 11-1.1 to implement the scheme. Identify the eight steps out-
lined in Section 12-2.6.

12-17. Consider the process shown in Problem 7-22 to dry phosphate pebbles. As
mentioned in the problem, an important disturbance to the process is the mois-
ture of the inlet pebbles. Using the information provided, design a feedforward/
feedback control scheme to compensate for its disturbance. There is a moisture
transmitter available to measure the inlet moisture. This transmitter has a range
of 12% to 16% moisture. Use the computing blocks of Table 1 1-l. 1. Identify
the eight steps outlined in Section 12-2.6.

12-18. Design a feedforward controller for controlling the outlet temperature of the
furnace sketched in Fig. P12-12. The controller must compensate for variations
in feed rate and feed temperature. Show all of the instrumentation required,
using Table 11-1.2. The following information is known: gas specific heat:
c_p = 0.26 Btu/lb-°F; fuel heating value = 980 Btu/scf; inlet gas temperature:
T_i = 90°F; outlet gas temperature: T_o = 850°F; furnace efficiency = 0.75;
range of FT42: 0 to 20,000 lb/h, range of TT42: 50°F to 120°F.

12-19. Let us propose that some of the process data, of the furnace of Problem 12-18,
such as the fuel heat of combustion or the gas specific heat, are not known.
Thus it is necessary to design the feedforward controller by using the block
diagram approach. The following data are obtained from the step tests on the
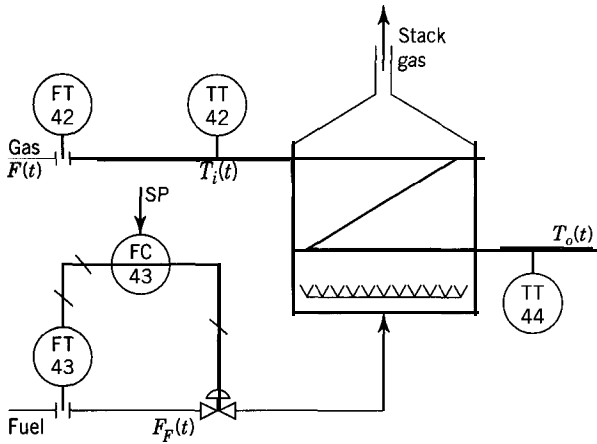furnace.

Figure P12-12 Furnace for Problem 12-18.

Variable	Step Change	Change in T_o, °F	Time Constant, min	Dead Time, min
w	750 lb/h	38	0.30	0.10
T_i	10°F	7.5	0.10	0.40
F_F^{set}	5% of range	20	0.95	0.15

(a) Draw the complete block diagram, including all transfer functions.
(b) Design the feedforward controller to compensate for variations in feed rate and feed temperature. Include the dynamic compensator.
(c) If it is decided to compensate for only one of the disturbances, which one would it be and why? How would your design be modified?

12-20. For the stripping section of a distillation column shown in Fig. P12-13, the objective is to maintain the bottom's purity at a desired value. This objective is commonly attained by controlling the temperature in one of the trays (the column pressure is assumed constant) by using the steam flow to the reboiler as the manipulated variable. A usual "major" disturbance is the feed flow to the column.

(a) Sketch a feedforward/feedback control scheme to compensate for this disturbance; describe it briefly.
(b) Briefly describe the dynamic tests that you would perform on the column in order to tune the feedback controller and the feedforward controller. Would you expect the dynamic compensation on the feedforward controller to be a net lead or a net lag?

12-21. Consider the exothermic reactor shown in Fig. P12-14. The diagram shows the control of the temperature of the reactor by manipulation of the cooling water valve.
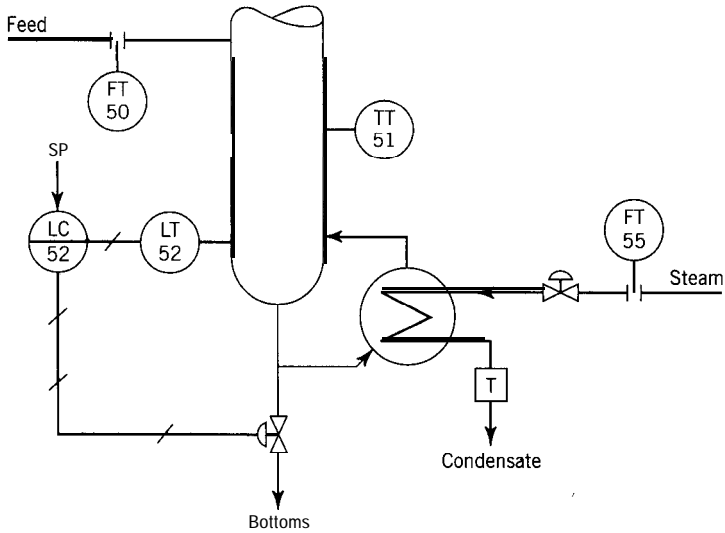
Figure P12-13 Distillation column for Problem 12-20.

(a) Design the control scheme required to control the reactants to the reactor. The flows of A and B can be both measured and controlled. The required ratio between these flows is 2.5 gpm of B to 1 gpm of A. The flowmeter of A has been calibrated between 0 and 40 gpm and that of B between 0 and 200 gpm. Show and scale the necessary instrumentation, using Table 11-1.1.

(b) Operating experience has shown that the inlet temperature of the cooling water varies somewhat. Because of the lags in the system-that is, cooling jacket, metal wall, and reactor volume-this disturbance usually results in
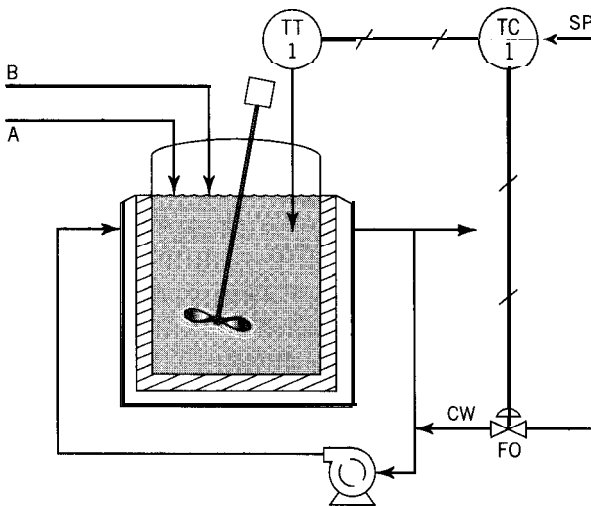
Figure P12-14 Reactor for Problem 12-21.

cycling of the reactor temperature. The engineers in charge of this unit have been wondering whether some other control scheme can help them improve the temperature control. Design such a control scheme.

(c) Operating experience has also shown that under some infrequent conditions, the cooling system does not provide enough cooling. In this case, the only way to control the temperature is by slowing down the flow of reactants. Design a control scheme to do this automatically. The scheme must be such that when the cooling capacity returns to normal, the scheme of part (b) is reestablished.

12-22. Consider the drying process shown in Fig. P12-15. In this process, wet paper stock is being dried to produce the final paper product. The drying is done using hot air; this air is heated in a heater in which fuel is burned to provide the energy. The controlled variable is the moisture of the paper leaving the drier. Fig. P12-15 shows the control scheme that was originally proposed and installed.

(a) A few weeks after start-up, the process engineers noticed that even though the moisture controller was keeping the moisture within certain limits from set point, the oscillations were more than desired. After searching for possible causes and making sure that the moisture controller was well tuned, they found that the temperature of the hot air leaving the heater varied more than had been assumed during the design stage. These variations were attributed to daily changes in ambient temperature and possible disturbances in the combustion chamber of the heater. Design a control scheme to maintain tighter moisture control.

(b) The control scheme just described significantly improved the moisture control. A few weeks later, however, the operators complained that every once in a while, the moisture would go out of set point considerably, though the control scheme would eventually bring it back to set point. This disturbance required that the paper produced during this period be reworked, so it represented a production loss. After searching through the production logs, the
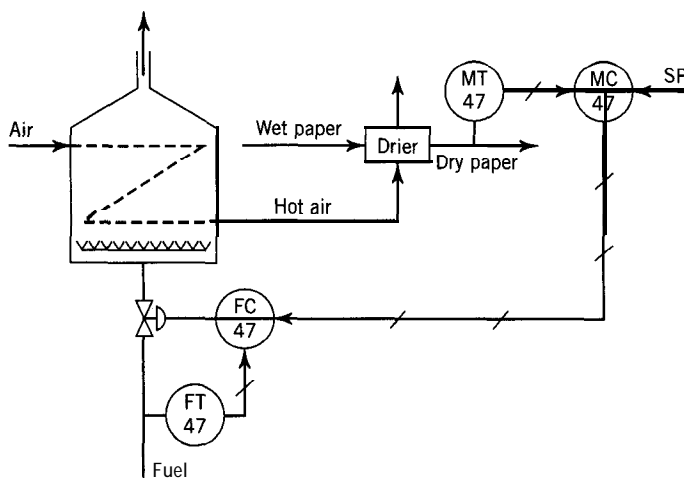
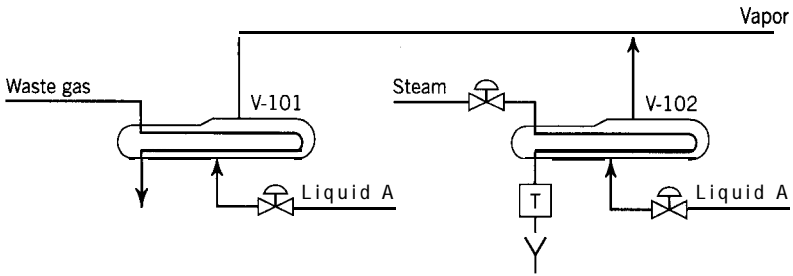**Figure P12-15** Wet paper drier for Problem 12-22.

Figure P12-16 Vaporizers for Problem 12-23.

process engineers discovered that changes in inlet moisture were the cause
of this disturbance. Design a feedforward control scheme that compensates
for these disturbances. There is a moisture transmitter, with a range of 5%
to 20% moisture available to measure the inlet moisture.

12-23. Consider a process, shown in Fig. P12-16, to vaporize a certain liquid A. Two
vaporizers, V-101 and V-102, are used for this purpose. The heating fluid in
V-101 is a waste gas that can be measured but not controlled; steam is used in
V-102 as the heating fluid. The pressure in the vapor header needs to be con-
trolled. It is also necessary to maintain a level of liquid A, above the heating
surface, in each vaporizer.

(a) Design a system to control the vapor header pressure and the level in the
vaporizers.

(b) Operating experience has shown that the waste gas flow, and sometimes its
temperature, vary often enough to swing (upset) the vapor header pressure
significantly. The production engineer wonders whether something can be
done to minimize the pressure swings. Propose a control system to accom-
plish this control.

12-24. Consider the reactor shown in Fig. P12-17, where the "well-known" reaction
A + B → C takes place. This reaction neither releases nor requires any energy.
Two streams mix and enter the reactor. Stream 1 is pure A, and stream 2 is
mainly B with varying amounts of A. The original control scheme calls for
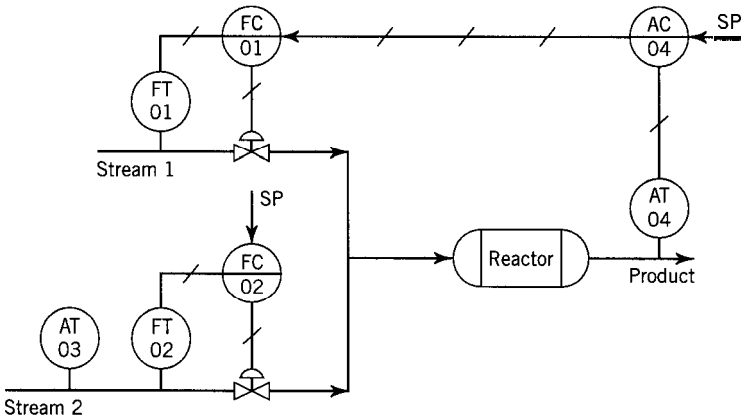
Figure P12-17 Reactor for Problem 12-24.

controlling the rate from this reactor by setting the flow of stream 2. The analysis of the product, the concentration of C in the product stream, manipulates the flow of stream 1; the figure shows this control scheme. After the start-up of this unit, it was noticed that the product concentration deviated significantly from set point once or twice every shift. An analyzer, AT-03, of component A in stream 2 indicated that a few seconds before this deviation in product concentration occurred, a change in stream 2 concentration occurred. Furthermore, the research department indicated that when too much reactant A is present during the reaction, it may start reacting with product C, reducing the amount of C in the product stream. Thus the production department has contacted you, the "well-known" control engineer, for ideas on how to maintain tighter control of the product concentration. The transmitter ranges are as follows:

FT-01: 0 to 800 lb/h   FT-02: 0 to 300 lb/h   AT-03: 0 to 0.3 mf

Design a control scheme, and select and scale the necessary computational blocks, that may provide the control performance required by the production department.

12-25. Consider the reactor shown in Fig. P12-18, in which the irreversible and complete liquid reaction A + B $\rightarrow$ C occurs. Product C is the raw material for several process units downstream from the reactor. Depending on the number of units operating and on their production rate, the production required from the reactor can vary between 4000 and 20,000 kmol/h of product C. Reactant A is available from two sources. Because of a long-term contract, source 1 is less expensive than source 2. However, the contract is written with two limitations: a maximum instantaneous rate of 16,800 kmol/h and a maximum monthly consumption of 3.456 X 10^6 kmol. If either of these limitations is exceeded, a very high penalty must be paid. In this case, it is less expensive to use the excess from source 2. You may assume that the densities of reactants A and B and product C do not vary much and can therefore be assumed constant.

(a) Design a control system that will preferentially use reactant A from source 1 and will not allow the exceeding of any contractual limitations. The feed ratio of A to B is 2:1 in kmole units.

(b) A few weeks after the control strategy designed in part (a) was put in operation, it was noticed that for some unknown reason, the supply pressure from source 2 was cut by the supplier every once in a while. Thus the flow
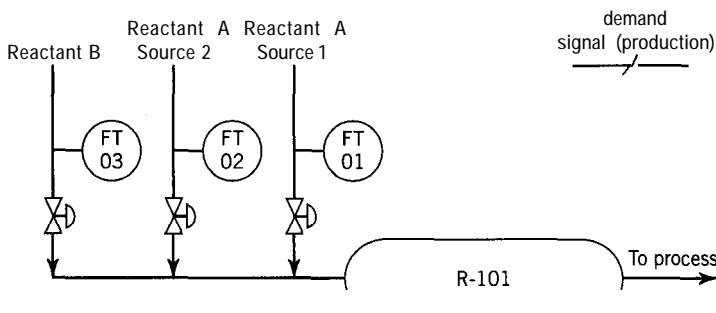

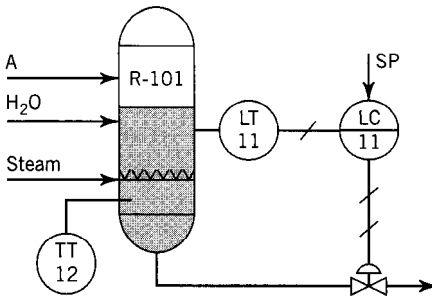Figure P12-18 Reactor for Problem 12-25.

Figure P12-19 Reactor for Problem 12-26.

controller manipulating the flow from source 2 would have to open the valve, and in some instances the valve would go wide open. At this moment, there would not be enough flow from source 2 to satisfy the demand. It was decided that the correct action to take in this case (while the lawyers investigate, which may take a long time), is to obtain from source 1 whatever source 2 does not supply. Design a control strategy to accomplish this action. Be sure that your design is such that, whenever source 2 provides the required flow, the scheme designed for part (a) is in effect.

12-26. Consider the reactor shown in Fig. P12-19, wherein stream A reacts with water. Stream A can be measured but not manipulated. This stream is the by-product of another unit. The water enters the reactor in two different forms, as liquid and as steam. The steam is used to heat the reactor contents. It is necessary to maintain a certain ratio, R, between the total water and stream A into the reactor; the ratio is a function of the concentration of stream A. It is very important to maintain this ratio below value Y. Otherwise, a very thick polymer may be produced, plugging the reactor.

(a) Assuming that the concentration of stream A does not vary much, design a control scheme to control the reactor temperature and at the same time maintain the required ratio.

(b) The operating department has noticed that the concentration of stream A has started to vary quite often. Thus a constant ratio is not appropriate any longer. A concentration sensor has been ordered to be installed on-line. Modify the control scheme of part (a) to take into account the availability of this sensor.

(c) A few months after the control scheme of part (b) had been implemented, the unit providing stream A started to develop a problem. This problem resulted in extended periods of time in which the flow of stream A decreased significantly. When this occurred, the control scheme totally cut the liquid water flow to the reactor to maintain the ratio. However, the steam flow to the reactor, to maintain temperature, still provided more water than was required, and thus the actual ratio of water to stream A entering the reactor was getting dangerously close to Y. Design a control scheme to avoid this condition when the flow of stream A decreases, even if it means that the temperature deviates from set point.

Chapter 13

Multivariable   Process Control

Up to this point in our study of automatic process control, we have considered only processes with a single control objective or controlled variable. Often, however, we encounter processes in which more than one variable must be controlled-that is, multiple control objectives. In such processes, we can still consider each control objective separately from the others as long as they do not interact with each other. In this chapter, we will study and design control systems for processes in which the various control objectives interact with each other. We refer to these systems as *multivariable control systems* or as *multiple-input, multiple-output (MIMO)* control systems. The problem we will be addressing is that of *loop interaction*. We will find that the response and stability of the multivariable system can be quite different from those of its constituent loops taken separately. We will learn how to pair the controlled and manipulated variables to minimize the effect of interaction and how to design decouplers that reduce or eliminate the effect of interaction.

13-1 LOOP INTERACTION

Figure 13-1.1 shows several examples of multivariable control systems. For the blending tank of Fig. 13-1.1*a,* it is necessary to control both the flow and the composition of the outlet stream. To accomplish this objective, we manipulate the flow of each of the two inlet streams. Figure 13-1. lb shows a chemical reactor for which it is necessary to control the outlet temperature and composition. The manipulated variables in this process are the cooling water flow and the process flow. In the evaporator of Fig. 13-1.1*c,* the level, the process flow, and the outlet composition are controlled by manipulating the steam, inlet, and outlet flows. Figure 13-1.1*d* shows a paper-drying machine in which the controlled variables are the moisture and dry basis weight (fibers per unit area) of the final paper product. The two manipulated variables are the stock flow to the machine and the steam flow to the last set of heated drums. Finally, Figure 13-1.1*e* depicts a typical distillation column with the necessary controlled variables: column pressure, product compositions, and the levels in the accumulator and column
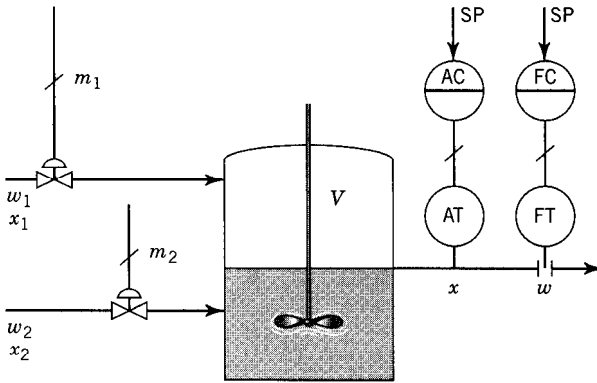
Figure 13-1.1 Examples of multivariable control systems.
(a) Blending tank.

base. The five manipulated variables are the coolant flow to the condenser, distillate flow, reflux flow, bottoms flow, and steam flow to the reboiler.

All of the foregoing are examples of processes with multiple interacting control objectives, the control of which can be quite complex and challenging to the process engineer. There are usually three questions the engineer must ask when faced with a control problem of this type.

1. What is the effect of interaction on the response of the feedback loops?
2. How much interaction exists between the loops, and which is the best way to pair the controlled and manipulated variables to reduce the effect of interaction?
3. Can the interaction between loops be reduced or eliminated through the design of an appropriate control system?

We will now address the first of these questions. The other two questions are addressed in later sections of this chapter.

To understand the effect of interaction, let us consider the blending tank of Fig. 13-1. la. Let us say that, at design conditions, the tank blends a solution containing 10 weight % salt with a concentrated solution containing 35 weight % salt, to produce

(b)
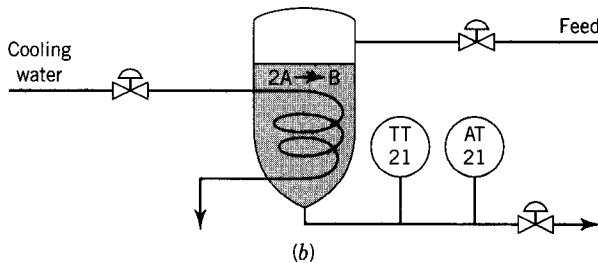
Figure   13-1.1   (Continued)   (b)   Chemical   reactor.

(c)
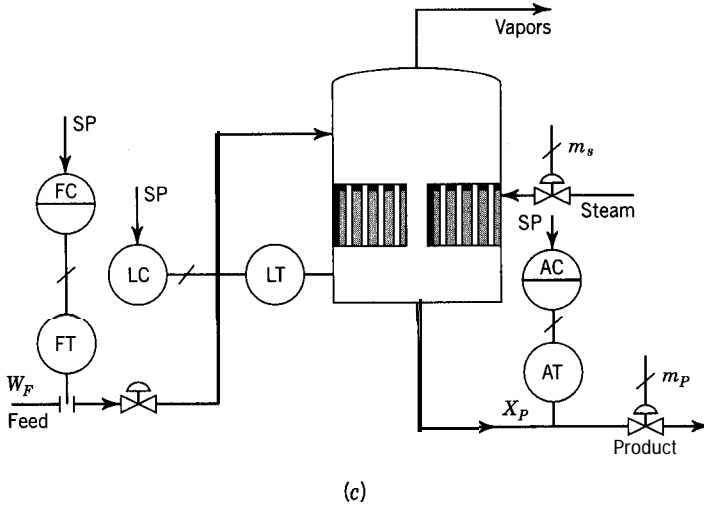
Figure 13-1.1 (Continued) (c) Evaporator.

100 lb/h of a solution containing 20 weight % salt. Steady-state balances on total mass and mass of salt around the tank result in the following two equations:

$$w = w_1 + w_2$$

(13-1.1)

$$wx = w_1 x_1 + w_2 x_2$$

where w is the stream flow in lb/h, x is the mass fraction of salt in each stream, and the subscripts mark the two inlet streams. At the design conditions of $\overline{w} = 100$ lb/h, $\overline{x} = 0.20$, $\overline{x}_1 = 0.1$, and $\overline{x}_2 = 0.35$, solving Eq. 13-1.1 results in the required inlet flows of $\overline{w}_1 = 60.0$ lb/h and $\overline{w}_2 = 40.0$ lb/h.

To show how interaction between the loops affects the response parameters of each loop, specifically the gain, let us consider only the composition control loop and assume that its manipulated variable is the concentrated stream, say stream 2. To obtain the

(d)
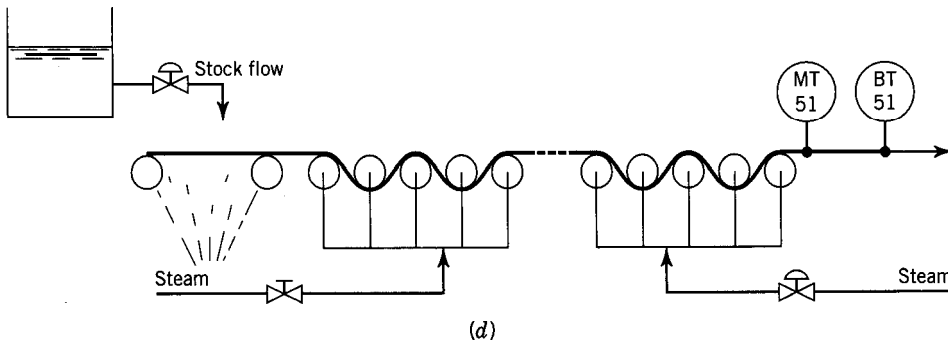
Figure 13-1.1 (Continued) (d) Paper-drying machine.

(e)
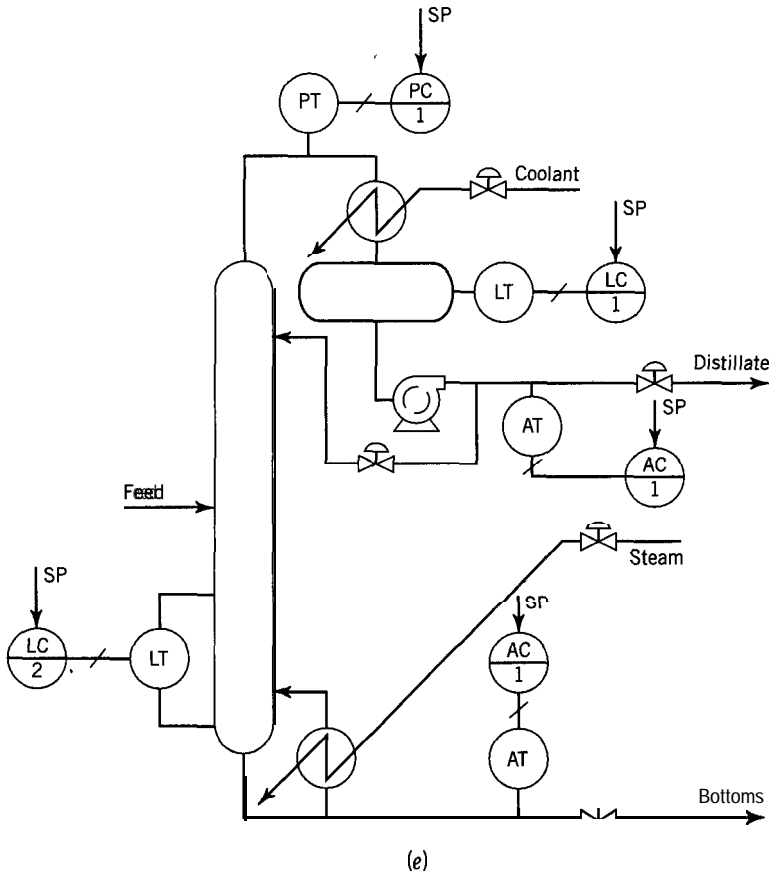
Figure 13-1.1 (Continued) (e) Distillation column.

gain of the loop for this arrangement, we apply a small change in the manipulated variable, say an increase of 2.0 lb/h in w_2 to 42.0 lb/h. From Eq. 13-1.1 we find that the product composition increases to 20.3 weight % salt, resulting in the following steady-state gain:

$$K_{x2} = \frac{20.3 - 20.0}{42.0 - 40.0} = 0.15 \quad \frac{\%\ salt}{lb/h}$$

where K_{x2} is the gain of the flow of stream 2 on the weight % salt in the product stream. At the same time, the total flow increases to $60.0 + 42.0 = 102.0$ lb/h.

Let us now consider the control system of Fig. 13-1.2, in which the product flow is controlled by manipulating the flow of the dilute inlet stream, w,. With this scheme, when the flow of stream 2 is increased from 40.0 to 42.0 lb/h, the product flow controller decreases the flow of the dilute stream to 58.0 lb/h to keep the product flow at 100.0 lb/h. Solving Eq. 13-1.1 for these conditions, we find that the product composition
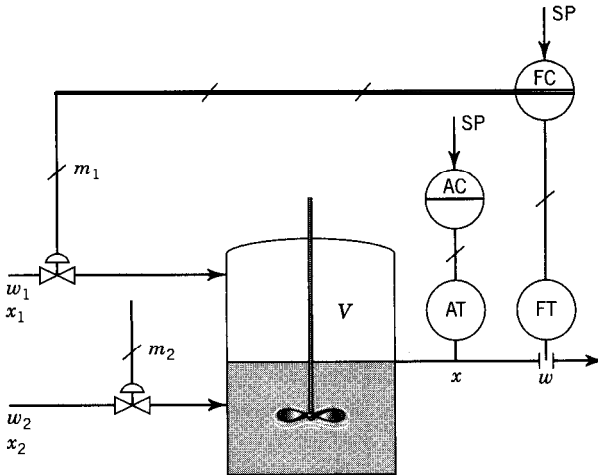
Figure 13-1.2 Blending tank control with product flow controller manipulating stream 1.

now increases to 20.5 weight % salt. The gain of the product composition loop becomes

$$K'_{x2} = \frac{20.5 - 20.0}{42.0 - 40.0} = 0.25 \quad \frac{\% \text{ salt}}{\text{lb/h}}$$

This represents an increase of 67% in the gain of the product composition loop, from 0.15 to 0.25 wt % salt/(lb/h), caused by interaction with the flow controller. To distinguish the two gains, *we* call K'_{x2} the *closed-loop gain,* because it is the gain of the composition loop when the flow loop is closed, whereas K_{x2} is the *open-loop gain,* the composition loop gain when the flow loop is opened.

Interaction occurs because when the composition controller changes the concentrated stream flow, it causes a change in the product flow; this in turn causes the flow controller to change the dilute stream flow, which causes a change in the product composition. This additional change in the product composition would not occur if the product flow were not controlled. It is easy to show that installation of the composition controller causes a similar change in the gain of the flow controller.

Figure 13-1.3 shows the response of the controlled and manipulated variables for the change in concentrated stream flow. The change in flow takes place at point A with the flow controller in *manual.* At point B, the flow controller is switched to *automatic* and brings the product flow back to its set point. The additional increase in product composition is the result of interaction.

Positive and Negative Interaction. In the blending tank example, the interaction is such that the two loops help each other; that is, closing the flow loop causes a change in the product composition that is in the same direction as the original change. This case, known as *positive interaction,* results in an increase in the loop gain when the other loop is closed. When the two loops fight each other, the interaction is said to be *negative,* and the gain of a loop decreases or changes sign when the other loop is closed.
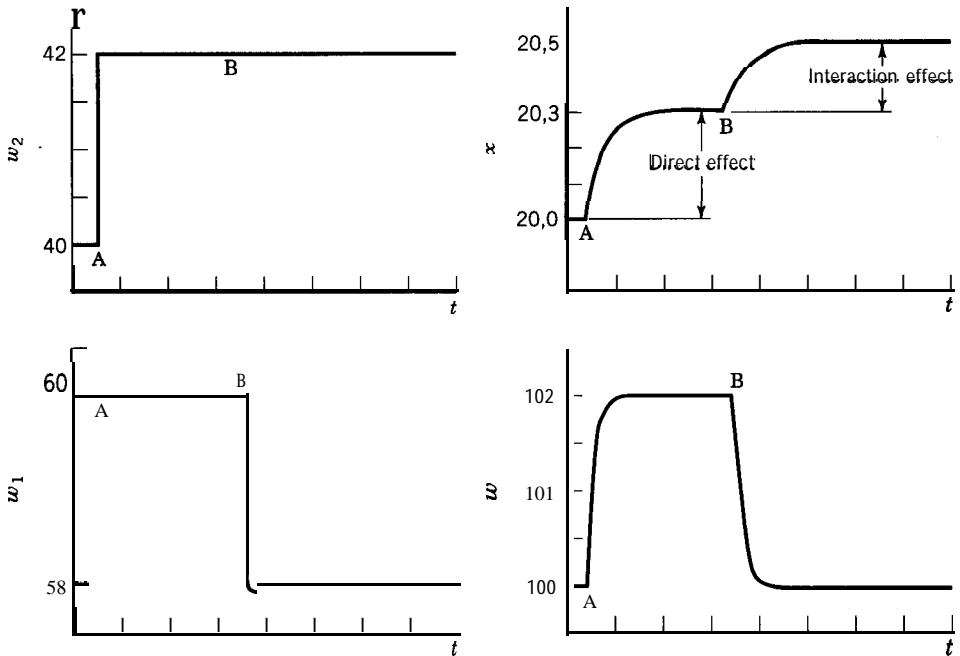
Figure 13-1.3 Response of blender to step change in concentrated stream flow with flow controller manipulating the dilute stream flow.

This is because the interaction causes a change in the controlled variable that is in the opposite direction to the original change. The change in the sign of the gain occurs when the interaction change is greater than the original change. Because of this possible change in the sign of the gain, negative interaction can be a more severe problem than positive interaction.

The preceding presentation of interaction has been limited to its effect on the steady-state gains. A more thorough analysis based on block diagrams will be presented later. We will next look at a quantitative measure of interaction and its application to the proper pairing of controlled and manipulated variables in multivariable control systems.

13-2 PAIRING CONTROLLED AND MANIPULATED VARIABLES

The second question we posed earlier about interaction concerned a quantitative measure of interaction and how to pair controlled and manipulated variables. Often it is simple to decide on the pairing, but many times, such as in the systems of Fig. 13- 1.1, it is more difficult. In this section we will learn to calculate a quantitative measure of the interaction between control loops. Then we will use this interaction measure to select the pairing of controlled and manipulated variables that minimizes the effect of interaction.

Consider the system of Fig. 13-2.1 with two controlled variables, c_1 and c_2, and two manipulated variables, m_1 and m_2. We call this system a 2 X 2 system. It makes sense to pair each controlled variable with the manipulated variable that has the greatest "influence" on it. In this context, influence and gain have the same meaning; to make
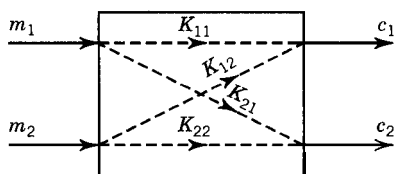
Figure 13-2.1 Schematic of general
2 X 2 process.

a decision, we must find the gain of each manipulated variable on each controlled variable.

Open-Loop   Gains

The four open-loop steady-state gains for the 2 X 2 system are

$$K_{11} = \frac{\Delta c_1}{\Delta m_1}\bigg|_{m_2} \qquad K_{12} = \frac{\Delta c_1}{\Delta m_2}\bigg|_{m_1}$$

$$(13\text{-}2.1)$$

$$K_{21} = \frac{\Delta c_2}{\Delta m_1}\bigg|_{m_2} \qquad K_{22} = \frac{\Delta c_2}{\Delta m_2}\bigg|_{m_1}$$

where K_{ij} is the gain relating the ith controlled variable to the jth manipulated variable. The vertical bars indicate that the gains are determined with the loops opened; that is, a change is made on each manipulated variable while the other manipulated variables are kept constant.

It may appear that the controlled and manipulated variables could be paired by simply comparing the open-loop gains. For example, if K_{12} were larger in magnitude than K_{11}, then m_2 would be chosen to control c_1. This is not quite correct, however, because the gains usually have different units and cannot be compared with each other. Even if the gains were expressed in %TO/%CO, factors such as the transmitter ranges and valve sizes, which have nothing to do with the process interaction, would affect the pairing.

Bristol (1966) proposed a measure of interaction that is independent of units of measure, transmitter ranges, and valve sizes. The measure of interaction is based on what we found in the previous section: in an interacting control system, the gain of a loop changes when the other loop or loops are closed. This means that for each of the open-loop gains of Eq. 13-2.1, there is a corresponding closed-loop gain, to be defined next.

Closed-Loop   Gains

For each pair of controlled and manipulated variables, the **closed-loop gain** is the change in the controlled variable divided by the change in the manipulated variable **when all the other controlled variables are held constant.** Note that this requires that the other manipulated variables be adjusted to bring the other controlled variables back to their base values.

For the 2 X 2 system of Fig. 13-2.1, the four steady-state closed-loop gains are

$$K'_{11} = \frac{\Delta c_1}{\Delta m_1}\bigg|_{c_2} \qquad K'_{12} = \frac{\Delta c_1}{\Delta m_2}\bigg|_{c_2}$$

$$K'_{21} = \frac{\Delta c_2}{\Delta m_1}\bigg|_{c_1} \qquad K'_{22} = \frac{\Delta c_2}{\Delta m_2}\bigg|_{c_1}$$

(13-2.2)

It is important to realize that the closed-loop gains, defined by Eq. 13-2.2, are not the same as the gains of the closed loops that we learned to compute in Chapter 6.

Interaction Measure, or Relative Gain

For a particular pair, the measure of interaction proposed by Bristol (1966) is simply the ratio of the open-loop gain to the closed-loop gain. For the general case,

$$\mu_{ij} = \frac{K_{ij}}{K'_{ij}}$$

(13-2.3)

where μ_{ij} is the interaction measure, or *relative gain,* for the pair c_i-m_j. For the 2 X 2 system of Fig. 13-2.1, the relative gains are

$$\mu_{11} = \frac{K_{11}}{K'_{11}} \qquad \mu_{12} = \frac{K_{12}}{K'_{12}}$$

$$\mu_{21} = \frac{K_{21}}{K'_{21}} \qquad \mu_{22} = \frac{K_{22}}{K'_{22}}$$

(13-2.4)

It is evident that each relative gain is dimensionless, because it is defined as the ratio of two gains relating the same controlled and manipulated variables. This is clearer when we substitute the gains from Eqs. 13-2.1 and 13-2.2, into Eq. 13-2.4. For example, for μ_{12},

$$\mu_{12} = \frac{K_{12}}{K'_{12}} = \frac{\dfrac{\Delta c_1}{\Delta m_2}\bigg|_{m_1}}{\dfrac{\Delta c_1}{\Delta m_2}\bigg|_{c_2}}$$

The relative gain is also independent of such things as transmitter ranges and valve sizes, because these parameters affect the open-loop and closed-loop gains in exactly the same way.

The usefulness of the relative gains as measures of interaction derives directly from their definition. If the gain for a given controlled-variable-manipulated-variable pair were not affected by interaction with other loops, the relative gain for that pair would be unity, because the closed-loop gain would be exactly the same as the open-loop

gain. The more a pair is affected by interaction with the other loops, the farther from unity its relative gain. When the interaction is positive, the relative gain is positive and less than unity, because the closed-loop gain is of the same sign as, and greater than, the open-loop gain. When the interaction is negative, the relative gain is either greater than unity or negative, because the closed-loop gain either is less than the open-loop gain or has the opposite sign.

Pairing Rule. The rule for using the relative gain to pair controlled and manipulated variables is simply stated as follows:

> *To minimize the effect of interaction in a multivariable control system, pair the controlled and manipulated variables so that the relative gain for each pair is closest to unity.*

It is important to realize that this rule considers only the steady-state effect of inter-action. In some systems, particularly those exhibiting negative interaction, dynamic effects may negate the rule's validity.

Pairings with negative relative gains must be avoided at all cost. A negative relative gain means that the action of the process changes when the other loops are opened and closed or when their manipulated variables reach their limits (e.g., fully opened or closed valves). In such cases, the loop with the negative relative gain will become unstable unless the action of the controller is changed, a difficult thing to do automatically.

It is possible, though not recommended, to select a pair with a relative gain of zero. In this case, the manipulated variable has no direct effect on the controlled variable and depends on interaction with the other loops to control it (zero open-loop gain). A pairing with a large relative gain will not work well, because the interaction with the other loops cancels most of the direct effect of the manipulated variable on the controlled variable (closed-loop gain near zero).

To further explore the meaning of the relative gain, let's look at two examples.

EXAMPLE 13-2.1

Consider the following matrix of relative gains:

	m_1	m_2
c_1	0.20	0.80
c_2	0.80	0.20

As we shall soon see, the relative gains for each row and column always add up to unity. Also, for a 2 X 2 system, there are only two possible pairing options with two pairs each: c_1-m_1, c_2-m_2, and c_1-m_2, c_2-m_1; the relative gains for the two pairs in each option are exactly the same.

The relative gain for one pairing, $\mu_{11} = \mu_{22} = 0.2 = 1/5$, indicates that for this **pair-**

ing, the gain of each loop increases by a factor of 5 when the other loop is closed. The relative gain for the other pairing, $\mu_{12} = \mu_{21} = 0.8 = 4/5$, indicates that for this pairing, the gain increases only by a factor of 1.25 when the other loop is closed. Obviously, the pairing c_1-m_2, c_2-m_1 results in less sensitivity to interaction than the other pairing.

EXAMPLE 13-2.2

Consider the following matrix of relative gains:

	m_1	m_2
c_1	2.0	− 1.0
c_2	− 1.0	2.0

In this case, the relative gain $\mu_{11} = \mu_{22} = 2.0 = 1/0.50$ indicates that the gain of each loop is cut in half when the other loop is closed, whereas the relative gain $\mu_{12} = \mu_{21} = -1.0 = 1/-1$ indicates that the gain of each loop changes sign when the other loop is closed. Certainly, this last case is undesirable, because it means that the action of the controller depends on whether the other loop is closed or open. The correct pairing is obviously c_1-m_1, c_2-m_2.

It is convenient to arrange the relative gains in a matrix, as demonstrated in the previous examples, because doing so graphically associates the values with the controlled and manipulated variables. This can also be done with the open-loop and the closed-loop gains. The first of the two examples involved a process with positive interaction, and the second had negative interaction.

We will next look at the calculation of the relative gains from the open-loop gains. First, we will derive a simple formula for the relative gains of 2 X 2 systems. Then we will derive a formula for a general n X n system-that is, for a system with n interacting control objectives.

13-2.1 Calculating the Relative Gains for a 2 x 2 System

The major advantage of the relative gain analysis presented here is that it requires only steady-state process parameters-specifically, the steady-state gains. In many cases, these gains can be calculated from simple steady-state material and energy balances on the process. More complex processes such as distillation columns and reactors may require a simulation of the process, but steady-state flowsheet simulations are readily available and commonly used to design the process and to analyze its performance. In this section, we show how to compute the relative gains from just the open-loop gains for a 2 X 2 system.

Relative gain analysis is based on a linear approximation of the process around some operating conditions. As we have said many times in this text, most processes are

nonlinear. This means that just as the steady-state gains vary with operating conditions, so do the relative gains.

For the 2 X 2 system of Fig. 13-2.1, the steady-state changes in the controlled variables caused by simultaneous changes in both manipulated variables are

$$Ac, = K_{11} Am, + K_{12} Am,$$

$$(13-2.5)$$

$$Ac, = K_{21} Am, + K_{22} Am,$$

where we have used the principle of superposition and assumed a linear approximation for small changes in the manipulated variables. The gains are the open-loop steady-state gains.

To determine the closed loop gain for the pairing c_1-m_1 , we must introduce a feedback controller connecting c_2 with m_2 , as in Fig. 13-2.2. If this controller has integral mode, when we apply a change in $m,$, it will adjust m_2 to bring c_2 back to its set point, making the steady-state change in c_2 equal to zero.

$$Ac, = K_{21} Am, 4 K_{22} Am, = 0$$

The change in m_2 required to compensate for the change in m_1 can be calculated from this formula.

$$Am, = - \frac{K_{21}}{K_{22}} \Delta m_1$$

Substitute this change in m_2 into Eq. 13-2.5 to obtain the total change in c_1 caused by the change in $m,$ when c_2 is kept constant.

$$Ac, = K,, Am, - K_{12} \frac{K_{21}}{K_{22}} Am,$$

This is the value we need to compute the closed-loop gain; from its definition, Eq. 13-2.2,

$$K'_{11} = \frac{\Delta c_1}{\Delta m,}\bigg|_{c_2} = K_{11} - K_{12} \frac{K_{21}}{K_{22}}$$
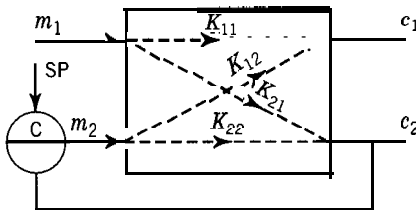
Figure 13-2.2 Schematic of general 2 X 2 process with one loop closed.

This formula shows that the closed-loop gain can be calculated from the four **open**-loop gains. The relative gains for each of the other three pairs of variables are obtained by appropriately rearranging the connections of the feedback controller of Fig. 13-2.2. The resulting formulas are, after minor rearrangement,

$$K'_{11} = \frac{K_{11}K_{22} - K_{12}K_{21}}{K_{22}} \qquad K'_{12} = \frac{K_{11}K_{22} - K_{12}K_{21}}{-K_{21}}$$

$$K'_{21} = \frac{K_{11}K_{22} - K_{12}K_{21}}{-K_{12}} \qquad K'_{22} = \frac{K_{11}K_{22} - K_{12}K_{21}}{K_{11}}$$

(13-2.6)

The relative gains can now be obtained from their definition, Eq. 13-2.3.

$$\mu_{11} = \frac{K_{11}K_{22}}{K_{11}K_{22} - K_{12}K_{21}} \qquad \mu_{12} = \frac{-K_{12}K_{21}}{K_{11}K_{22} - K_{12}K_{21}}$$

$$\mu_{21} = \frac{-K_{12}K_{21}}{K_{11}K_{22} - K_{12}K_{21}} \qquad \mu_{22} = \frac{K_{11}K_{22}}{K_{11}K_{22} - K_{12}K_{21}}$$

(13-2.7)

As we noted earlier, the relative gain for the pair c_1-m_1 is equal to that for the pair c_2-m_2, ($\mu_{11} = \mu_{22}$), because they represent the same pairing option. Similarly, $\mu_{12} = \mu_{21}$. Also, it can readily be shown that $\mu_{11} + \mu_{12} = 1$ and that $\mu_{12} + \mu_{22} = 1$; that is, the sum of the relative gains in any row or column is equal to 1. This last property extends to $n \times n$ systems.

You can easily memorize the formulas for the relative gains by noticing that the denominator is the determinant of the matrix of open-loop gains and that the numerator is the product term of the denominator that contains the open-loop gains for the pair of interest.

Having shown that the relative gains can be calculated from the open-loop gains, we shall next demonstrate the procedure for calculating the relative gains.

EXAMPLE 13-2.3 *RELATIVE GAINS OF BLENDING PROCESS*

The equations relating the variables of the blending process of Fig. 13-1.1a were presented earlier in this chapter (Eq. 13-1.1). These equations are so simple that the gains can be obtained analytically by using differential calculus. Determine the general formulas for the open-loop and relative gains for the blending process, and derive a general strategy for pairing variables in any blending process.

SOLUTION

The first step is to solve for the controlled variables in terms of the manipulated variables. From Eq. 13-1.1,

$$\text{w} = w_1 + w_2$$

$$x = \frac{w_1 x_1 + w_2 x_2}{w_1 + w_2} \tag{13-2.8}$$

where w and x are the controlled variables and w_1 and w_2 are the manipulated variables. Note that we use the manipulated flows instead of the control valve positions as manipulated variables. This is because the gains of the valves have no effect on the relative gains; they affect the open-loop and closed-loop gains in exactly the same way.

Differential calculus can be used to determine the open-loop gains: take limits of the formulas for the gains as the change in manipulated variable approaches zero.

$$K_{ij} = \lim_{\Delta m_j \to 0} \frac{\Delta c_i}{\Delta m_j}\bigg|_{\substack{m_k \\ k \neq j}} = \frac{\partial c_i}{\partial m_j} \tag{13-2.9}$$

Applying this definition to Eq. 13-2.8, we obtain the following open-loop gains:

$$K_{w1} = 1 \qquad\qquad K_{w2} = 1$$

$$K_{x1} = \frac{-w_2(x_2 - x_1)}{(w_1 + w_2)^2} \qquad K_{x2} = \frac{w_1(x_2 - x_1)}{(w_1 + w_2)^2} \tag{13-2.10}$$

The relative gains are then obtained using Eq. 13-2.7.

$$\mu_{w1} = \frac{w_1}{w_1 + w_2} \qquad \mu_{w2} = \frac{w_2}{w_1 + w_2}$$

$$\mu_{x1} = \frac{w_2}{w_1 + w_2} \qquad \mu_{x2} = \frac{w_1}{w_1 + w_2} \tag{13-2.11}$$

Both relative gains are positive and less than unity, which indicates positive interaction. The correct pairing is the one that results in the relative gain being closest to unity. Because the blending process is nonlinear, the relative gains and therefore the correct pairing are a function of the operating conditions. The correct pairing depends on which of the two inlet flows is larger.

- If $w_1 > w_2$, then the relative gains closer to 1 are μ_{w1} and μ_{x2} (they are equal). The correct pairing is w-w_1, x-w_2.
- If $w_2 > w_1$, then the relative gains closer to 1 are μ_{w2} and μ_{x1} (they are equal). The correct pairing is w-w_2, x-w,.

Note that this results in a general pairing strategy for blending processes:

Control the product flow with the inlet stream that has the larger flow, and control the composition with the inlet stream that has the smaller flow.

The composition represents any intensive property of the product stream that is to be controlled. It could be, for example, temperature, as in the case of a household shower, where cold and hot water streams are mixed to obtain the desired shower temperature.

For processes in which the level in the tank is controlled, the stream with the largest flow should be manipulated to control the level. In this example, the stream with the largest flow is the product stream, but it may be desirable to have the operator directly set the product flow. In such a case, the product stream is flow-controlled, the inlet stream with the larger flow is manipulated to control the level, and the inlet stream with the smaller flow controls the intensive product property (e.g., composition, temperature).

EXAMPLE 13-2.4   CONTROL OF DISTILLATION PRODUCT PURITIES

The distillation column of Fig. 13-2.3 is designed to separate a mixture consisting of 60 mole % benzene and 40 mole % toluene into a distillate product with 95 mole %
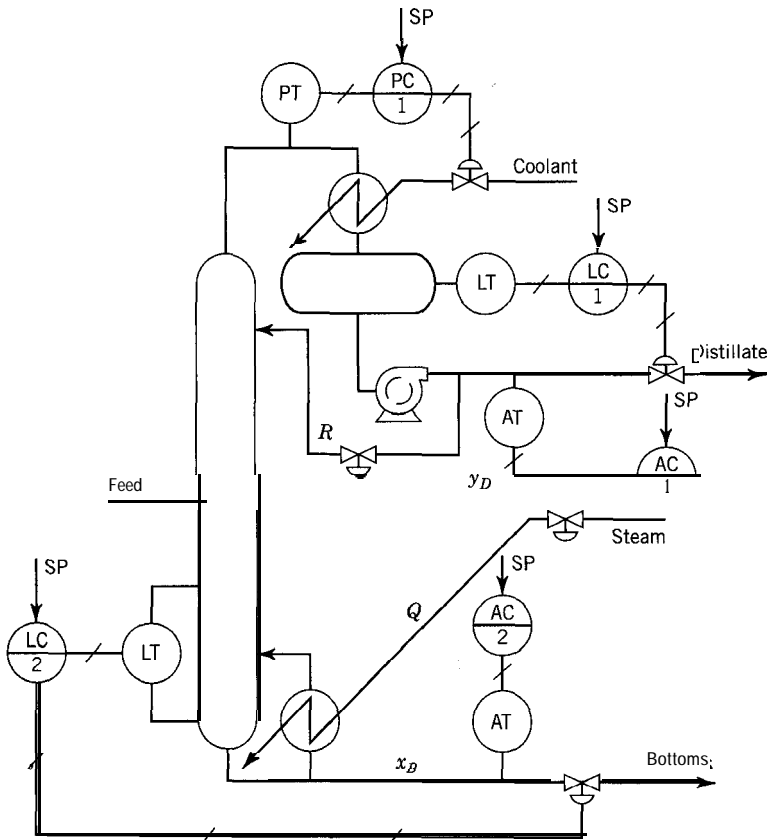
Figure 13-2.3 Distillation column for Example 13-2.4.

benzene and 5 mole % toluene and a bottoms product with 5 mole % benzene and 95 mole % toluene. The feed rate is 1000 lbmol/h and enters on the seventh tray from the top. The column has 13 sieve trays, a partial reboiler, and a total condenser, and the Murphree tray efficiency is 70%. The distillate composition, measured as mole % toluene, and the bottoms composition, measured as mole % benzene, are to be controlled by manipulating the reflux flow and the heat rate to the reboiler. Determine the open-loop gains for the 2 X 2 system, the relative gains, and the pairing of the manipulated and controlled variables that minimizes the effect of interaction. Assume that the pressure controller maintains the column pressure constant and that the two level controllers maintain the material flows in balance at the top and bottom by manipulating the two product flows.

SOLUTION

A ChemSep (Taylor and Kooijman, 1995) simulation of the column was used to determine the steady-state gain. Normally, three runs of the simulation are needed to estimate the steady-state gains: one at the design conditions, one in which the reflux flow is changed while the reboiler heat rate is kept constant, and one in which the reboiler heat rate is changed at constant reflux. However, because distillation columns exhibit asymmetric responses when the variables are increased and decreased, it is necessary to make five runs so that the reflux flow and reboiler heat rate can both be increased and decreased from their design values. The results of these runs are shown in the accompanying table.

	Base Case	Test 1+	Test 1−	Test 2+	Test 2−
Reflux, klb/h	131.94	132.59	131.29	131.94	131.94
Heat rate, MBtu/h	30.10	30.10	30.10	30.23	29.97
Distillate mole %	5.00	4.43	5.69	5.87	4.34
Bottoms mole %	5.00	5.84	4.25	4.05	6.06

Note the asymmetric response of the compositions. Although the change of reflux flow in each direction is of the same magnitude, 0.65 klb/h, the composition changes are different for the increase and decrease in flow, and the same happens for the changes in heat rate. These differences are small in this case, because the changes in flow and heat rate are small, as they should be. Nevertheless, these small differences are enough to throw the calculation of the gains off. This is why two changes, one up and one down, are required for each variable. The open-loop gains are calculated as follows:

$$K_{DR} = \frac{4.43 - 5.69}{132.59 - 131.29} = -0.97 \qquad K_{DQ} = \frac{5.87 - 4.34}{30.23 - 29.97} = 5.88$$

$$K_{BR} = \frac{5.84 - 4.25}{132.59 - 131.29} = 1.22 \qquad K_{BQ} = \frac{4.05 - 6.06}{30.23 - 29.97} = -7.73$$

where the first subscript is D for the distillate composition or B for the bottoms com-position, and the second subscript is R for the reflux flow or Q for the reboiler heat rate. Gains K_{DR} and K_{BR} have units of mole %/(klb/h), and gains K_{DQ} and K_{BQ} have units of mole %/(MBtu/h).

The relative gains are, from Eq. 13-2.7,

$$\mu_{DR} = \mu_{BQ} = \frac{(-0.97)(-7.73)}{(-0.97)(-7.73) - (1.22)(5.88)} = 23.1$$

$$\mu_{BR} = \mu_{DQ} = \frac{-(1.22)(5.88)}{(-0.97)(-7.73) - (1.22)(5.88)} = -22.1$$

or, in matrix form,

	R	Q
y_D	23.1	-22.1
x_B	-22.1	23.1

The relative gains are either greater than unity or negative, indicating negative inter-action. To avoid negative relative gains, the pairing should be the obvious one: the reflux flow controls the distillate composition, and the reboiler heat rate controls the bottoms composition. However, even for the best pairing, the high value of the relative gain, 23.1, shows that the two loops fight each other to the point of essentially canceling each other's actions. For example, assume the distillate composition controller changes the reflux flow by 0.2 klb/h. If the bottoms composition loop is open, then the change in distillate composition, using the open-loop gain, is

$$\Delta y_D = K_{DR}\ \text{AR} = (-0.97)(0.20) = -0.194 \text{ mole \% toluene}$$

and the bottoms composition will change by

$$\Delta x_B = K_{BR}\ \text{AR} = (1.22)(0.20) = 0.244 \text{ mole \% benzene}$$

If the bottoms composition controller is switched to automatic, then it will bring the bottoms composition back to set point. The required change in reboiler heat rate is

$$\Delta Q = \frac{-\Delta x_B}{K_{BQ}} = \frac{-0.244}{-7.73} = 0.03\ 16 \text{ MBtu/h}$$

The total change in distillate composition will then be

$$\Delta y_D = K_{DR}\ AR + K_{DQ}\ AQ = (-0.97)(0.20) + (5.88)(0.0316)$$

$$= -0.194 + 0.186 = -0.008 \text{ mole \% toluene}$$

Note that the change caused by interaction between the two controllers, 0.186, essentially cancels the effect the manipulated variable had on its controlled variable, -0.194. The gain of each controller is reduced by a factor of 23.1, the relative gain.

13-2.2 Calculating the Relative Gains for an n x n System

A system of more than two interacting control loops is called an n X n system, where n is the number of interacting control objectives. Such a system consists of n controlled variables and n manipulated variables. The definition of the relative gain is the same as for 2 X 2 systems: the ratio of the open-loop gain to the closed-loop gain for each controlled-variable-manipulated-variable   pair:

$$\mu_{ij} = \frac{\left.\dfrac{\Delta c_i}{\Delta m_j}\right|_{\substack{m_k \\ k \neq j}}}{\left.\dfrac{\Delta c_i}{\Delta m_j}\right|_{\substack{c_k \\ k \neq i}}} = \frac{K_{ij}}{K'_{ij}} \tag{13-2.12}$$

where K_{ij} is the open-loop gain (the gain when all other manipulated variables are kept constant), and K'_{ij} is the closed-loop gain (the gain when all other controlled variables are kept constant).

The formulas for the relative gain matrix for a 2 X 2 system, Eq. 13-2.7, are simple. For an n X n system, it is easier to develop a procedure based on matrix operations. The procedure, as proposed by Bristol (1966), is as follows:

Obtain the transpose of the inverse of the steady-state gain matrix and multiply each term of the resulting matrix by the corresponding term in the original matrix. The terms thus obtained are the relative gains.

Mathematically, let B be the inverse of the steady-state gain matrix K, that is, let $\mathbf{B} = \mathbf{K}^{-1}$. Then

$$\mu_{ij} = B_{ji}K_{ij} \tag{13-2.13}$$

where K_{ij} and B_{ji} are elements of matrices K and B, respectively. With today's personal calculators and programs such as MATLAB (1994) and MathCad (1994), the required matrix operations can be readily performed. For a 2 X 2 system, this procedure yields the same relative gains as Eq. 13-2.7.

The derivation of Eq. 13-2.13 is carried out as follows. In matrix notation, the change in the controlled variables caused by simultaneous changes in the manipulated variables is

$$\mathbf{A}c = \mathbf{K}\mathbf{A}m \tag{13-2.14}$$

To solve for the changes in the manipulated variables required to produce specific changes in the controlled variables, we multiply by the inverse of the steady-state gain matrix K.

$$Am = K^{-1} Ac = B Ac \tag{13-2.15}$$

We now write the jth element of this matrix equation.

$$\Delta m_j = B_{j1} Ac_1 + \ldots + B_{ji} \Delta c_i + \cdots + B_{jn} Ac_1 \tag{13-2.16}$$

from which we notice that

$$B_{ji} = \frac{\Delta m_j}{\Delta c_i}\bigg|_{\substack{c_k \\ k \neq i}} = \frac{1}{K'_{ij}} \tag{13-2.17}$$

In other words, the elements of the inverse matrix B are the inverse of the closed-loop gains, transposed. Combining Eq. 13-2.17 with Eq. 13-2.12, we obtain Eq. 13-2.13. Let us demonstrate the use of these formulas with an example.

EXAMPLE 13-2.5   *CONTROL OF GASOLINE BLENDING SYSTEM*

Figure 13-2.4 shows an in-line blender to produce gasoline of a given formulation by mixing three refinery products: alkylate (stream 1), light straight run (stream 2), and reformate (stream 3). It is desired to control the research octane x, Reid vapor pressure y, and flow f of the gasoline product by manipulating the flow set points on the inlet streams. Calculate the steady-state gains and the relative gains, and decide which inlet stream should control each product variable so that the effect of interaction is mini-mized. The design values of the process variables are as shown in the accompanying table.

		Octane, x	RVP, $_Y$	Flow, kbl/day
Alkylate	1	97.0	5.00	7.50
Straight run	2	80.0	11.0	28.12
Reformate	3	92.0	3.00	24.38
Gasoline		87.0	7.00	60.00

SOLUTION

Assuming that the properties of the product gasoline are the average of the feed streams, weighted by the volume rate of each stream, and that the densities of all streams are
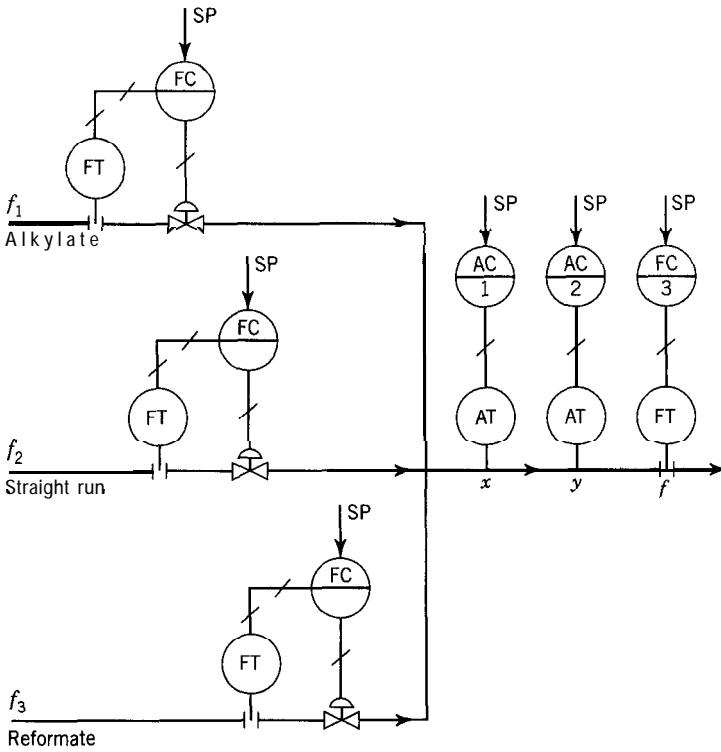
Figure 13-2.4 In-line gasoline blending process for Example 13-2.5.

approximately equal, we obtain the following equations relating the controlled variables to the manipulated flows.

$$x = \frac{f_1 x_1 + f_2 x_2 + f_3 x_3}{f_1 + f_2 + f_3}$$

$$y = \frac{f_1 y_1 + f_2 y_2 + f_3 y_3}{f_1 + f_2 + f_3}$$

$$f = f_1 + f_2 + f_3$$

Note that the design values given in the statement of the problem satisfy these equations, as they must. This means that three of the design values are not independent of the rest.

As in Example 13-2.3, we obtain the steady-state gains by taking partial derivatives of the controlled variables with respect to the manipulated variables. The results are

$$K_{1j} = \frac{\partial x}{\partial f_j} = \frac{f_1(x_j - x_1) + f_2(x_j - x_2) + f_3(x_j - x_3)}{(f_1 + f_2 + f_3)^2}$$

$$K_{2j} = \frac{\partial y}{\partial f_j} = \frac{f_1(y_j - y_1) + f_2(y_j - y_2) + f_3(y_j - y_3)}{(f_1 + f_2 + f_3)^2}$$

$$K_{3j} = \frac{\partial f}{\partial f_j} = 1$$

Substitute the design values to obtain the open-loop gain matrix.

$$K = \begin{bmatrix} 0.167 & -0.117 & 0.083 \\ -0.033 & 0.067 & -0.067 \end{bmatrix}$$

The inverse of this matrix is

$$B = K^{-1} = \begin{bmatrix} 7.500 & 11.250 & 0.125 \\ 5.625 & -15.938 & 0.406 \\ 1.875 & -4.688 & 0.469 \end{bmatrix}$$

where the inverse was calculated using MathCad (1994). From Eq. 13-2.13, the relative gains are

	f_1	f_2	f_3
x	1.250	0.219	− 0.469
Y	− 0.375	0.312	1.063
f	0.125	0.469	0.406

The boxed values are the relative gains closest to 1, which minimize the effect of interaction. The proper pairing is then to control the octane (x) with the alkylate (stream 1), the Reid vapor pressure (y) with the reformate (stream 3), and the gasoline flow (f) with the light straight run (stream 2). Two of the relative gains are slightly greater than unity, which means the interaction is negative for each of these loops and thus their gains decrease when the other loops are closed. The other relative gain is 0.469, which means the loop gain is more than doubled when the other loops are closed. This loop is helped by the combination of the other two loops (positive interaction). Note that the sum of the relative gains in each row and each column of the matrix is unity, as it should be. Because an in-line mixer has very fast response, it should be sufficient to use the steady-state analysis of the interactions to decide on the best pairing.

   In this example, as in the two-stream blender of Example 13-2.3, the correct pairing has the inlet stream with the largest flow controlling the product flow. This makes sense, because the largest stream has the greatest relative influence on the product flow.

13-3 DECOUPLING OF INTERACTING LOOPS

The third and final question to answer is whether the interaction between loops can be reduced or eliminated through the design of an appropriate control system. The answer

is of course yes, and the simplest way to do it is by *decoupling*. Decoupling can be a profitable, realistic possibility when applied carefully. The relative gain matrix provides an indication of when decoupling can be beneficial. If for the best pairing option, one or more of the relative gains is far from unity, decoupling may help. For existing systems, operating experience usually helps in making the decision.

This section presents the design of decouplers, which is very similar to the design of feedforward controllers presented in Chapter 12. Decouplers can be designed from block diagrams or from basic engineering principles. The basic difference is that, unlike feedforward controllers, decouplers form part of feedback control of loops. Because of this, they must be selected and designed with great care. One basic characteristic of decoupling is that *in interacting systems, decoupling does to each loop what the other loops were going to do anyway.* The difference is that the performance of each loop becomes independent of the tuning and open/closed condition of the other loops, provided the action of the decoupler is not blocked.

13-3.1 Decoupler Design from Block Diagrams

Figure 13-3.1 presents the block diagram for a 2 X 2 interacting system. This block diagram shows graphically that the interaction between the two loops is caused by the process "cross" blocks with transfer functions $G_{,,}(s)$ and $G_{,,}(s)$. To circumvent this interaction, two decoupler blocks with transfer functions $D_{,,}(s)$ and $D_{,,}(s)$ are installed, as shown in Fig. 13-3.2. The purpose of the decouplers is to cancel the effects of the process cross blocks so that each controlled variable is not affected by changes in the manipulated variable of the other loop. In other words, decoupler $D_{,,}(s)$ cancels the effect of manipulated variable $M_{,}(s)$ on controlled variable $C_{,}(s)$, and $D_{,,}(s)$ cancels
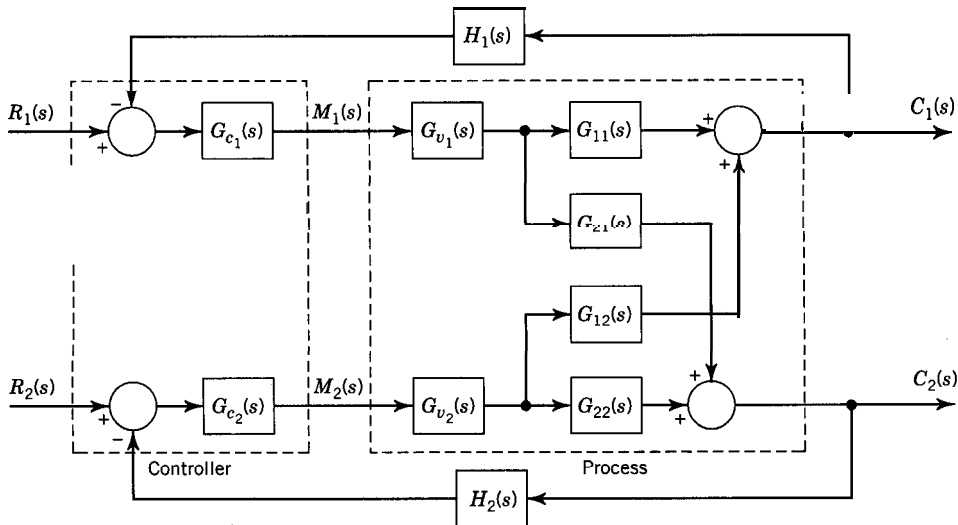
Figure **13-3.1** Block diagram of general 2 **X** 2 control system.
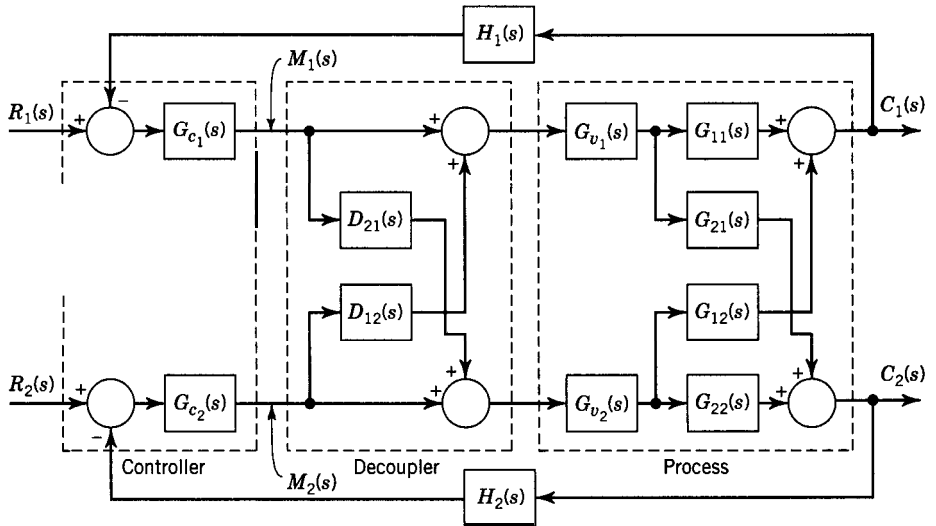
Figure 13-3.2 Block diagram of 2 X 2 control system with decouplers.

the effect of $M_2(s)$ on controlled variable C,(s). From block diagram algebra, these effects are

$$\frac{C_1(s)}{M_2(s)} = D_{12}(s)G_{v_1}(s)G_{11}(s) + G_{v_2}(s)G_{12}(s) = 0$$

(13-3.1)

$$\frac{C_2(s)}{M_1(s)} = D_{21}(s)G_{v_2}(s)G_{22}(s) + G_{v_1}(s)G_{21}(s) = 0$$

Note that this is similar to designing feedforward controllers in which M, (s) is a disturbance to $C_1(s)$ and M, (s) is a disturbance to $C_2(s)$. To obtain the design formulas for the decouplers, solve for the decoupler transfer functions from Eq. 13-3.1.

$$D_{12}(s) = -\frac{G_{v_2}(s)G_{12}(s)}{G_{v_1}(s)G_{11}(s)}$$

(13-3.2)

$$D_{21}(s) = -\frac{G_{v_1}(s)G_{21}(s)}{G_{v_2}(s)G_{22}(s)}$$

These formulas can be used to design linear decouplers for any 2 X 2 system as long as the transfer functions of the process and valves can be derived from basic principles, as we learned to do in Chapters 3, 4, 5, and 6, or approximated from the step responses of the controlled variables to the manipulated variables, as we learned in Chapter 7. In the latter approach, the transfer functions of the valve, process, and transmitter are

lumped into a single transfer function for each combination. It can easily be shown that the decoupler formulas then become

$$D_{,,}(s) = -\frac{G_{P_{12}}(s)}{G_{P_{11}}(s)} \qquad D_{21}(s) = -\frac{G_{P_{21}}(s)}{G_{P_{22}}(s)} \tag{13-3.3}$$

where $G_{P_{ij}}(s) = G_{,}(s) \; G_{ij}(s)H_i(s)$.

As pointed out earlier, the decouplers form part of the loops, as can be seen from the block diagram of Fig. 13-3.2. The relationship between each controlled variable and its manipulated variable in the decoupled diagram is obtained by block diagram algebra:

$$\frac{C_1(s)}{M_1(s)} = G_{v_1}(s)G_{11}(s) + D_{21}(s)G_{v_2}(s)G_{12}(s)$$

$$\tag{13-3.4}$$

$$\frac{C_2(s)}{M_2(s)} = G_{v_2}(s)G_{22}(s) + D_{12}(s)G_{v_1}(s)G_{21}(s)$$

The two terms in each formula tell us what the block diagram of Fig. 13-3.2 tells us graphically: that in the decoupled system, each manipulated variable affects its controlled variable through two parallel paths. As with the interacting system, these two paths may help each other if their effects are additive (positive interaction), or they may fight each other if the effects are of opposite sign (negative interaction).

If the decouplers can be implemented exactly as designed by Eq. 13-3.2, then substitution into Eq. 13-3.4 gives, after some simplification,

$$\frac{C_1(s)}{M_1(s)} = G_{v_1}(s)\left[G_{,,}(s) - \frac{G_{21}(s)G_{12}(s)}{G_{22}(s)} \right]$$

$$\tag{13-3.5}$$

$$\frac{C_2(s)}{M_2(s)} = G_{v_2}(s)\left[G_{22}(s) - \frac{G_{12}(s)G_{21}(s)}{G_{11}(s)} \right]$$

It is easy to show that the steady-state gains of the terms in brackets, obtained by letting $s = 0$ in the transfer functions, are exactly the same as the closed-loop gains of Eq. 13-2.6. In other words, the effect of perfect decouplers on the gain of the loop is the same as the effect of closing both loops without decouplers. This makes sense, because each decoupler is keeping each controlled variable constant when the manipulated variable in the other loop changes. This is exactly what the controller on the other loop does, at steady state, if it has integral mode.

EXAMPLE 13-3.1 DECOUPLER DESIGN FOR A BLENDING TANK

Design a linear decoupler for the blending tank of Fig. 13-1. la. Neglect the time constants of the control valves, because they are small relative to the time constant of the tank.

SOLUTION

To obtain the transfer functions for the tank, we write total and salt balances. Assuming perfect mixing, constant volume and density, and constant inlet compositions,' the differential equations are

$$w_1(t) + w_2(t) - w(t) = \frac{d}{dt}(\rho V) = 0$$

(13-3.6)

$$x_1 w_1(t) + x_2 w_2(t) - x(t)w(t) = \frac{d}{dt}[\rho V x(t)]$$

where ρ is the density of the solution, lb/ft^3, and V is the volume of solution in the tank, ft^3. The flows $w_1(t)$, $w_2(t)$, and w(t), are in lb/min, and the compositions are in mass fraction (mf) salt. We next linearize and take **Laplace** transforms, as in Chapter 3, to obtain the transfer functions.

$$W(s) = W_1(s) + W_2(s)$$

(13-3.7)

$$X(s) = \frac{K_{x1}}{\tau s + 1} W_1(s) + \frac{K_{x2}}{\tau s + 1} W_2(s)$$

where

$$\tau = \frac{\rho V}{\overline{w}} \text{ min}$$

$$K_{x1} = \frac{x_1 - \overline{x}}{\overline{w}} \frac{\text{mf}}{\text{lb/min}}$$

$$K_{x2} = \frac{x_2 - \overline{x}}{\overline{w}} \frac{\text{mf}}{\text{lb/min}}$$

These gains are the same as those of Eq. 13-2.10 because, given that the base conditions are at steady state, it can be shown that

$$\frac{x_2 - x_1}{\overline{w}} = \frac{x_2 - \overline{x}}{\overline{w}_1} = -\frac{x_1 - \overline{x}}{\overline{w}_2}$$

and $\overline{w} = \overline{w}_1 + \overline{w}_2$.

The transfer functions we have developed match the block diagram of Fig. 13-3.2 with $C_1(s) = W(s)$, and $C_2(s) = X(s)$. When we neglect the time constants of the control

---

[1] The inlet compositions are the disturbance inputs, but we will consider them constant to keep the analysis simple.

valves, from Eq. 13-3.7, the transfer functions shown in the block diagram are

$$G_{11}(s) = 1 \qquad G_{12}(s) = 1$$

$$G_{21}(s) = \frac{K_{x1}}{\tau s + 1} \qquad G_{22}(s) = \frac{K_{x2}}{\tau s + 1}$$

$$G_{v_1}(s) = K_{v_1} \qquad G_{v_2}(s) = K_{v_2}$$

where K_{v_1} and K_{v_2} are the valve gains in (lb/h)/%CO.

The decouplers are now obtained by substituting the foregoing transfer functions into the decoupler design formulas, Eq. 13-3.2.

$$D_{12}(s) = -\frac{K_{v_2}G_{12}(s)}{K_{v_1}G_{11}(s)} = \frac{K_{v_2}\;\%CO_1}{K_{v_1}\;\%CO_2}$$

$$D_{21}(s) = -\frac{K_{v_1}K_{x1}}{K_{v_2}K_{x2}}\frac{K_{v_1}\,\overline{w}_2\;\%CO_2}{K_{v_2}\,\overline{w}_1\;\%CO_1}$$

(13-3.8)

where we have substituted the gains from Eq. 13-2.10. Figure 13-3.3 shows the instrumentation diagram for the tank with the linear decoupler. In the diagram, FY-1 implements $D_{,,}(s)$ and AY-2 implements $D_{21}(s)$.

Both decouplers are simple gains. This makes sense, because both inlet flows have exactly the same dynamic effects on the outlet flow and composition. Decoupler $D_{,,}(s)$ (FY-1) has a negative gain, because its purpose is to keep the outlet flow constant when the second inlet stream changes. This requires that the first flow change in the opposite direction by exactly the same amount. The decoupler gain corrects for the different capacities of the two valves.
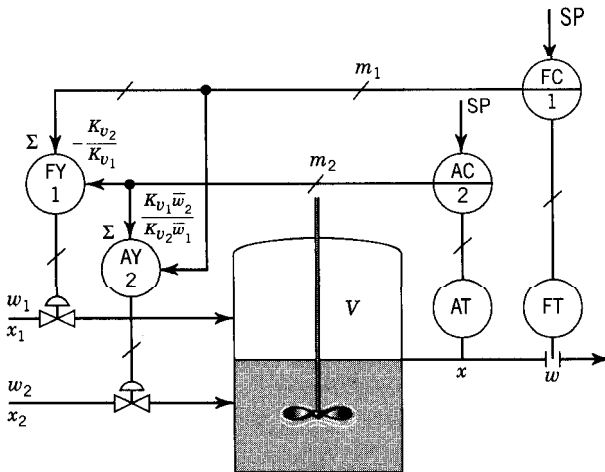
Figure 13-3.3 Instrumentation diagram for linear decoupler of blending tank.

Decoupler D,,(s) (AY-2) is positive, because its purpose is to keep the outlet composition constant when the first inlet flow changes. To keep the outlet composition constant, the concentrated stream flow must change in the same direction as the dilute stream flow, and the ratio of the two streams must remain constant. Note that this is exactly what the decoupler does; its gain is the ratio of the two streams corrected for the capacities of the two valves.

To illustrate, suppose that valve 2 has half the gain (capacity) of valve 1 ($K_{v_1} = 2K_{v_2}$) and that, at the base conditions, flow 1 is 50% greater than flow 2 ($\overline{w}_1 = 1.5\ \overline{w}_2$). Then, if the composition controller changes its output by 1%, decoupler D,,(s), with a gain of $-1/2$, changes the signal to valve 1 by -0.5%. Because valve 1 has twice the gain of valve 2, the stream 1 flow will decrease by exactly the same amount stream 2 increased, keeping the total flow constant. If the flow controller now changes its output by 1%, decoupler D,,(s), with a gain of $(2)(1/1.5) = 1.33$, changes the signal to valve 2 by 1.33%. Because valve 2 has half the gain of valve 1, the change in the flow of stream 2 is $1.33/2 = 0.67$ of the change in the flow of stream 1. This maintains the flow ratio of stream 2 to stream 1 at 0.67. That is, the flow stream 1 continues to be 50% greater than the flow of stream 2, which keeps the outlet composition constant.

Variable Pairing and Decoupling

We have shown that decoupling has the same effect on each loop as the interacting loops had before decoupling. Because of this, and because the decoupler action could be blocked if the valve it actuates were driven to saturation, it is important to pair the controlled and manipulated variables properly even when a decoupler is used.

In the preceding example the variables are properly paired, with stream 1, the larger flow, controlling the product flow and with stream 2, the smaller flow, controlling the composition. Note that this makes the ratio of the two streams less than unity, so the action of the flow controller in changing stream 1 has a greater effect on the product flow than the action of the decoupler. Imagine what would happen if the ratio of stream 2 to stream 1 were 10 and the same pairing were used: the decoupler would have 10 times greater effect on the product flow than the direct action of the controller. It can be shown that the same would happen to the composition controller.

Partial Decoupling

It is not always necessary to decouple all the interactions. When control of one of the variables is more important than control of the others, better control of that variable results when only the decoupler terms that keep that variable constant are implemented. Suppose that in the blending tank example it is more important to keep the product composition constant than it is to keep the product flow constant. We would then implement only decoupler D,,(s) (AY-2 in Fig. 13-3.3). This keeps the composition from changing when the flow controller takes action. When the composition controller takes action, the flow is allowed to vary, and the flow controller must take action to bring the flow back to set point. However, this action of the flow controller does not affect the composition because the decoupler compensates for it.

The following example shows how to design a decoupler from simple process models.

EXAMPLE 13-3.2 **DECOUPLER** *DESIGN FOR AN* EVAPORATOR

In the evaporator of Fig. 13-3.4, the product composition, x_P, and feed flow, w_F, are to be controlled by manipulating the signals to the control valves on the steam and product lines, m_S and m_P. As shown in the diagram, although there is a control valve on the feed line, this valve must control the level in the evaporator (LC). This is because the feed is the largest of the three flows in the system and thus has the greatest influence on the level. The level in a calandria-type evaporator must be controlled very tightly, because it has great influence on the heat transfer rate.

Step tests are performed on the signals to the steam and product valves, m_S and m_P, one at a time, and the responses of the feed flow and product compositions are carefully recorded. The resulting first-order-plus-dead-time models are

$$G_{FS}(s) = \frac{0.84e^{-0.30s}}{1.05s + 1} \qquad G_{FP}(s) = \frac{0.20e^{-0.03s}}{0.65s + 1}$$

$$G_{xS}(s) = \frac{1.68e^{-0.80s}}{2.70s + 1} \qquad G_{xP}(s) = \frac{-1.60e^{-0.50s}}{2.97s + 1}$$

where the first subscript is F for feed rate or x for product mass percent, and the second subscript is S for steam or P for product. Each transfer function is in units of %TO/%CO; that is, they each combine the transfer functions of the control valve, the process, and the sensor/transmitter. Calculate the relative gains, pair the controlled and manipulated variables in such a way as to minimize the effect of interaction, and design the decouplers.

SOLUTION

The open-loop steady-state gains are obtained by setting $s = 0$ in the model transfer functions, and the relative gains are calculated from Eq. 13-2.7. In matrix form,
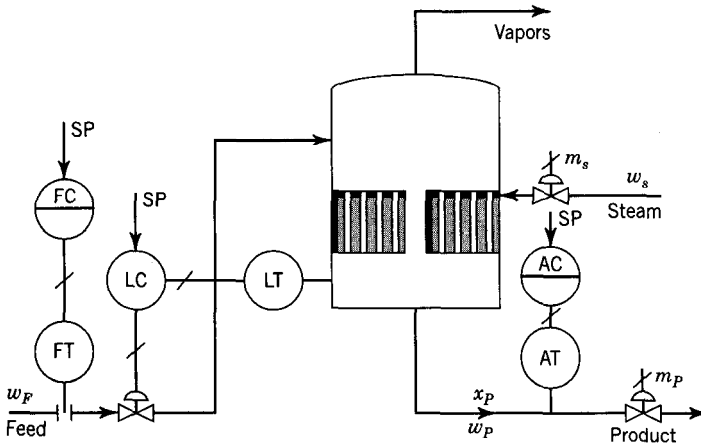
Figure 13-3.4 Evaporator control system for Example 13-3.2.

they are

K	m_S	m_P		μ	m_S	m_P
w_F	0.84	0.20		w_F	0.80	0.20
x_P	1.68	-1.60		x_P	0.20	0.80

The pairing that minimizes interaction has the steam valve control the feed flow and has the valve on the product stream control the product composition. The relative gain for this pairing is 0.8.

The decouplers are designed using Eq. 13-3.3, because we have the combined transfer functions for all the field devices. The transfer functions of the decouplers are

$$D_{12}(s) = -\frac{G_{FP}(s)}{G_{FS}(s)} = -\frac{0.20}{0.84}\left(\frac{1.05s + 1}{0.65s + 1}\right)e^{0.27s}$$

$$D_{21}(s) = -\frac{G_{xS}(s)}{G_{xP}(s)} = \frac{1.68}{1.60}\left(\frac{2.97s + 1}{2.70s + 1}\right)e^{-0.30s}$$

The first decoupler has a positive dead time that cannot be implemented. To provide this necessary lead in the action of the decoupler, we increase the lead to $1.05 + 0.27 = 1.32$ min. This reduces the decoupler to a simple lead-lag unit with a gain of $-0.20/0.84 = -0.24$, a lead of 1.32 min and a lag of 0.65 min. Although the dead-time compensation of the second decoupler is negative and could be implemented, using the dead time in a decoupler is not recommended. This is because, as we saw earlier, the decoupler forms part of the feedback loop, and dead time decreases the controllability of feedback loops. To remove the dead-time term and keep the same total lag, we decrease the lead to $2.97 - 0.30 = 2.67$ min. This makes the lead almost the same as the lag, which means that the second decoupler can be a simple gain. The resulting decouplers are

$$D_{12}(s) = -0.24\left(\frac{1.32s + 1}{0.65s + 1}\right)\frac{\%CO_S}{\%CO_P}$$

$$D_{21}(s) = \frac{1.68}{1.60} = 1.05 \quad \frac{\%CO_P}{\%CO_S}$$

Figure 13-3.5 shows the instrumentation diagram for the evaporator decouplers. The purpose of the first decoupler, $D_{21}(S)$ (FY-2 in the diagram), is to keep the feed flow constant when the product flow is changed. Its gain is negative because an increase in product flow requires a decrease in steam flow to reduce the rate of vapor generation. If these two effects cancel each other, as they will if the decoupler gain is correct, then the level in the evaporator, and consequently the feed flow, do not change. This decoupler requires a net lead (lead-lag FY-1) because the vapor rate has a lag to the steam flow change whereas the product flow has an immediate effect on the level.

The purpose of the second decoupler, $D_{21}(s)$ (AY), is to maintain the product com-
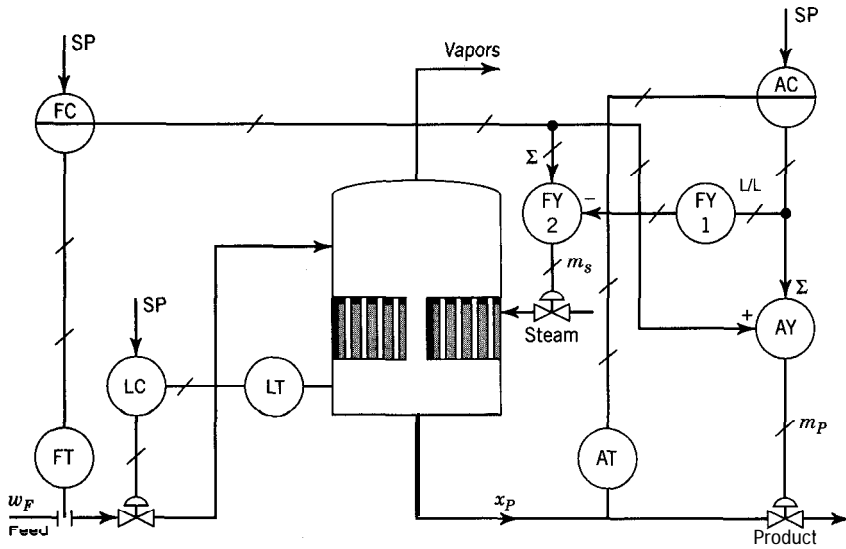
Figure   13-3.5   Instrumentation   diagram   for   linear   decoupler   of   evaporator.

position constant when the steam flow changes. Its gain is positive because when the steam flow increases, so does the vapor rate, causing the product concentration to increase. To keep this increase from happening, the decoupler will increase the product flow just enough to cause the level controller to increase the flow of dilute feed into the evaporator. This increased feed flow will decrease the product composition, balancing the effect of the increase in vapor rate. No dynamic compensation is required in this decoupler, because the lag between the steam and vapor flows is approximately matched by the mixing lag of the feed flow on the composition.

In practice, it is more important to maintain the product composition constant than the feed flow. Thus we choose partial decoupling, using only the term $D_{21}(s)$ (AY) and leaving out the other decoupler term (FYs).

Static Decoupling

As with feedforward control, decoupling can be accomplished statically and dynamically. Static decoupling is accomplished by leaving out the dynamic terms. The advantage is that only the steady-state gains of the transfer functions need be used in designing a static decoupler. The steady-state gains can be obtained off-line by the methods we discussed when we talked about determination of the relative gains.

13-3.2   Decoupler Design for $n \times n$ Systems

The simple procedure for designing the decouplers for 2 X 2 systems, Eq. 13-3.1, cannot be extended to systems with more than two interacting control objectives. Once more, we must turn to matrix notation to outline a design procedure for the general $n \times n$ system. However, even with matrix notation, the design of decouplers is not

simple for these higher-order systems because of the large number of decouplers required. For example, for a 4 x 4 system, twelve decouplers must be designed and implemented.

By the rules of matrix multiplication, it can be shown that the relationship between the controlled and manipulated variables in a decoupled n x n system is

$$
\begin{bmatrix} C_1(s) \\ C_2(s) \\ \vdots \\ C_n(s) \end{bmatrix} = \begin{bmatrix} G_{P_{11}}(s) & G_{P_{12}}(s) & | & | & | & G_{P_{1n}}(s) \\ G_{P_{21}}(s) & G_{P_{22}}(s) & | \cdot & | & G_{P_{2n}}(s) \\ ||| & ||| & ||| & ||| & ||| \\ G_{P_{n1}}(s) & G_{P_{n2}}(s) & | & | & | & G_{P_{nn}}(s) \end{bmatrix} \begin{bmatrix} 1 & D_{12}(s) & \cdots & D_{1n}(s) \\ D_{21}(s) & 1 & \cdots & D_{2n}(s) \\ ||| & ||| & ||| & ||| \\ D_{n1}(s) & D_{n2}(s) & | & | & | & 1 \end{bmatrix} \begin{bmatrix} M_1(s) \\ M_2(s) \\ \cdots \\ M_n(s) \end{bmatrix}
$$

or, equivalently,

$$
\mathbf{c}(s) = \mathbf{G}_P(s)\mathbf{D}(s)\mathbf{m}(s) \tag{13-3.9}
$$

where the terms of matrix $G_s(s)$ combine the transfer functions of all the field elements. The objective of the decoupler matrix $\mathbf{D(s)}$ is to obtain the following decoupled system:

$$
\mathbf{c}(s) = \mathbf{G}'_P(s)\mathbf{m}(s) = \begin{bmatrix} G'_{P_{11}}(s) & 0 & \cdots & 0 \\ 0 & G'_{P_{22}}(s) & \cdots & 0 \\ ||| & ||| & ||| & ||| \\ 0 & 0 & ||| & G'_{P_{nn}}(s) \end{bmatrix} \mathbf{m}(s)
$$

In other words, a diagonal matrix. Comparing this desired result with Eq. 13-3.9, we find that

$$
\mathbf{G}_P(s)\mathbf{D}(s) = \mathbf{G}'_P(s)
$$

Next we solve for the decoupler matrix by premultiplying by the inverse of the process transfer function matrix.

$$
\mathbf{D(s)} = \mathbf{G}_P^{-1}(s)\mathbf{G}'_P(s) \tag{13-3.10}
$$

This is the design formula for decoupling an n x n system. The matrix equation represents n^2 independent equations. The n^2 unknowns are the $(n^2 - n)$ nondiagonal elements of $\mathbf{D(s)}$ and the n elements of the diagonal matrix $\mathbf{G}'_P(s)$.

Equation 13-3.10 is not easy to solve, because it requires the inversion of a matrix of transfer functions. Because of this, modem computer multivariable control systems are designed by different methods. Two of these methods are Dynamic Matrix Control, DMC (Cutler and Ramaker, 1979), and Identification Command, or IdCom (see Reference 6). The first of these will be discussed in Chapter 15.

If the inverse $\mathbf{B(s)} = \mathbf{G}_P^{-1}(s)$ can be obtained, then the elements of the decoupler matrix are calculated as follows:

$$
D_{ij}(s) = \frac{B_{ij}(s)}{B_{jj}(s)} \tag{13-3.11}
$$

This calculation forces the diagonal terms of the decoupler matrix to be unity. The following example shows how Eq. 13-3.10 is used to design a static decoupler.

EXAMPLE 13-3.3   *STATIC DECOUPLER FOR GASOLINE BLENDING SYSTEM*

Design a static decoupler for the gasoline blending control system of Example 13-2.5.

SOLUTION

Because the decoupler is static, the steady-state open-loop gains can be substituted for the process transfer function in Eq. 13-3.10. However, before this equation is applied to design the decoupler, we must rearrange the gain matrix to conform to the pairing that minimizes interaction. Recall from Example 13-2.5 that the correct pairing has stream 2 controlling the third variable and stream 3 controlling the second variable. With this pairing, columns 2 and 3 of the open-loop gain matrix must be swapped.

	f_1	f_3	f_2
x	0.167	0.083	-0.117
y	-0.033	-0.067	0.067
f	1	1	1

When this matrix is inverted, the inverse is the same matrix as in Example 13-2.5, with the last two rows swapped.

$$B = \mathbf{K}^{-1} = \begin{bmatrix} 7.500 & 11.250 & 0.125 \\ -5.625 & -13.938 & 0.406 \\ -1.875 & -4.688 & 0.469 \end{bmatrix}$$

Equation 13-3.11 gives us the decoupler matrix.

$$D = \begin{bmatrix} 1 & -0.71 & 0.27 \\ -0.75 & 1 & 0.87 \\ -0.25 & -0.29 & 0.87 \end{bmatrix}$$

where all the numbers are in (kbl/day)/(kbl/day). Assuming that the control signals are in these units instead of in %CO, the equations required to carry out the decoupling are

$$f_1^{set} = m_1 - 0.71m_2 + 0.27m_3$$

$$f_3^{set} = -0.75m_1 + m_2 + 0.87m_3$$

$$f_2^{set} = -0.25m_1 - 0.29m_2 + m_3$$

where m_1, m_2, and m_3 are in kbl/day and so are the flow set points. Figure 13-3.6 presents the instrumentation diagram required to implement these calculations. Note that stream
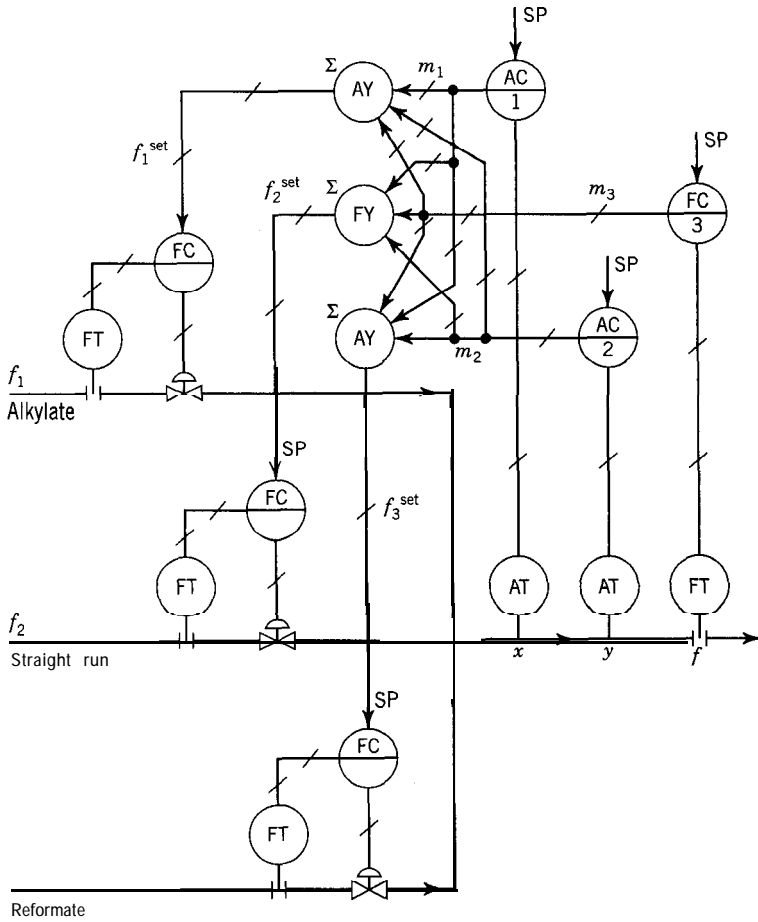
Figure 13-3.6 Instrumentation diagram, for static linear decoupler on gasoline blending system.

3, the reformate, is directly manipulated by m_2, the output of the Reid vapor pressure controller, and that stream 2, the light straight run, is directly manipulated by m_3, the output of the gasoline flow controller. This is the pairing we found in Example 13-2.5.

If each manipulated stream has the same dynamic effect on its controlled variable as the others, then this static decoupler is all that is needed to decouple the three control objectives. The decoupled transfer function matrix is then the matrix product **KD**, which is a diagonal matrix with diagonal terms 0.133, $-$ 0.063, and 2.133. The transfer functions of the decoupled system are then

$$X(s) = 0.133M_1(s)$$

$$Y(s) = -0.063M_2(s)$$

$$F(s) = 2.133M_3(s)$$

These gains are dimensional, because they come from the process gains that are in engineering units. This means we cannot compare them with each other.

13-3.3 Decoupler Design from Basic Principles

The decoupler design techniques discussed previously were derived from the block diagram of the multivariable control system. Because block diagrams show only linear relationships between the variables, only linear decouplers can be designed by that procedure. This section presents an example of the design of decouplers from process models that are obtained by applying basic engineering principles to the process. The procedure is fundamentally the same as the one presented in Chapter 12 for designing feedforward controllers. The only difference is that, in decoupler design, the disturbances to one loop are the manipulated variables for the other loops.

EXAMPLE 13-3.4 *NONLINEAR DECOUPLER FOR BLENDING TANK*

Design a decoupler for the blending tank of Fig. 13- 1. la from the basic model equations that represent the tank.

SOLUTION

In designing the decoupler we ignore the disturbances, which for the tank are the inlet stream compositions. After identifying the controlled and manipulated variables, we write the simple mass and energy balances that relate them. For the blending tank, total mass and salt balances result in Eq. 13-1.1, which can be slightly rearranged into the following form:

$$w = w_1 + w_2$$

$$x = \left(\frac{w_1}{w_1 + w_2}\right) x_1 + \left(\frac{w_2}{w_1 + w_2}\right) x_2$$

(13-3.12)

Shinskey (1981) proposes the idea of assigning to the controller output the physical significance of the process variable or combination of variables that has the most influence on the controlled variable. We see from the preceding formula that the product flow w is most affected by the sum of the two inlet flows, rather than by either one of them. Also, the product composition is most affected by the ratios of the input flows to the total flow. However, these two ratios are not independent of each other, because they must add up to unity:

$$\frac{w_1}{w_1 + w_2} + \frac{w_2}{w_1 + w_2} = 1$$

Therefore, we must pick one of them. According to the results of the relative gain analysis for the tank, the smaller of the two flows should control the composition of

the product (see Example 13-2.3). This result extends to the ratios, so that the smaller of the two ratios should be the one to control the composition. Let us assume that stream 2 is smaller than stream 1. Then the outputs of the two controllers must be made to have the following significance:

$$m_, = w_1 + w_2$$

(13-3.13)

$$m_2 = \frac{w_2}{w_1 + w_2} = \frac{w_2}{m_1}$$

where m_1 and m_2 are the outputs of the product flow and composition controllers, respectively, assumed to be in the appropriate engineering units.

The implementation of any nonlinear control scheme is more precise if flow controllers are installed on the manipulated streams so that the controller outputs are the set points of these flow controllers. The decoupler scheme then consists of calculating the set points to the two inlet streams, w_1^{set} and w_2^{set}. From Eq. 13-3.13, these set points are

$$w_1^{set} = m_1 - w_2$$

(13-3.14)

$$w_2^{set} = m_1 m_2$$

The implementation of these equations is shown in Fig. 13-3.7a. Note that the actual flow w_2, from the flow transmitter signal, is used in the calculation of w_1^{set}. This ensures

(a)
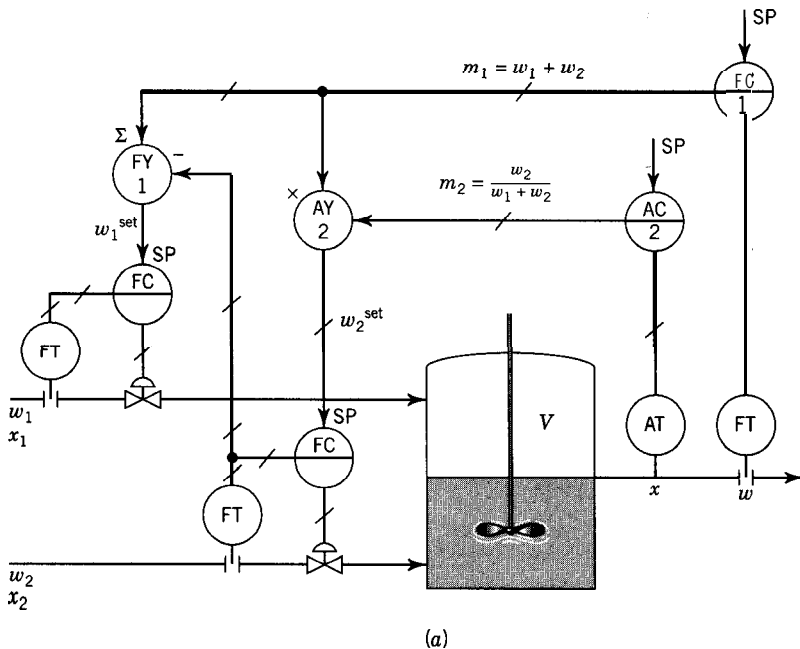
Figure   13-3.7   Nonlinear   decoupler   for   blending   tank.   (a)   Combining   the   controller outputs.
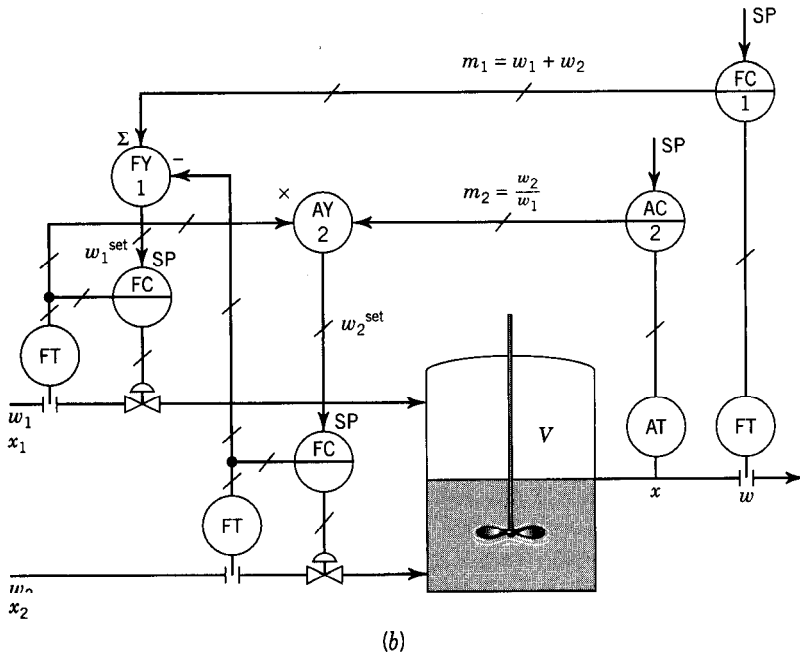
Figure 13-3.7 (Continued) (b) Using a ratio controller.

that the correct set point of stream 1 is computed even when the flow controller on stream 2 is unable to keep its set point. You might ask, "Why not also use the measured total flow w instead of m_1 in the formulas?" Shinskey (198 1) points out that if the outlet flow is used in calculating the set points of the inlet flows, then a positive feedback loop is created around the tank (an increase in the outlet flow increases the inlet flow, which increases the outlet flow further, and so on). Such positive feedback may result in instability and should not be allowed.

The decoupling scheme of Fig. 13-3.7a may be simplified to the scheme shown in Fig. 13-3.7b. The difference is that the output of the composition controller is the ratio w_2/w_1 between the two inlet streams. It is easy to show that this simpler ratio still decouples the two loops, because it makes the product composition independent of total flow. This scheme also makes it possible to use only partial decoupling by leaving out the summer that calculates w_1^{set}. If this is done, then the product composition is independent of the total flow, but the product flow is affected by changes in composition.

13-4 MULTIVARIABLE CONTROL VS. OPTIMIZATION

Often in the design of control systems for multivariable systems, a feedback controller is installed for every flow in the process that can be manipulated. This is not always good practice, especially when it is not necessary to maintain the controlled variable at a specific set point. An alternative to providing a controlled variable for each manipulated variable is to use the additional manipulated variables to optimize the operation of the process.

Consider the distillation column of Example 13-2.4. In that example, we found that

the scheme that tries to control both distillate and bottoms composition has little chance for success because the action of each controller is almost exactly canceled by the interaction with the other. Suppose that it was not really important to maintain the bottoms composition at 5 mole % benzene but that this controller had been included just to keep the losses of benzene in the bottom from becoming too high. If benzene is the valuable product from the column, excessive losses of benzene in the bottom represent lower profits because of lower distillate rates. However, the economic value of the benzene lost in the bottom stream must be balanced against the cost of the steam required to provide the heat to the reboiler. Lower benzene compositions in the bottom require higher reboiler heat rates and thus entail higher steam costs.

A simple optimization scheme can be applied to the column by removing the bottoms composition controller and setting the steam rate at the value that maximizes the profit rate. The profit rate is the difference between the values of the products and the costs of the feed, steam, and condenser coolant.

$$\text{Profit rate} = V_D w_D + V_B w_B - V_R Q_R - V_C Q_C - V_F w_F$$

where V are the economic values in \$/klb or \$/MBtu, w are the mass rates in klb/h, Q are the heat rates in MBtu/h, and the subscripts refer to the distillate, bottoms, reboiler, condenser, and feed.

Several methods have been proposed to determine the reboiler heat rate that maximizes the profit rate, but they are outside the scope of this text. See, for example, Moore and Corripio (1991).

13-5 DYNAMIC ANALYSIS OF MULTIVARIABLE SYSTEMS

To analyze the dynamic response of a multivariable control system, we must determine the transfer functions of the system. In this section, we show that the characteristic equation of a set of interacting control loops does not have the same roots as the characteristic equations of the individual loops taken separately. But before we proceed with the analysis of the complex block diagrams that represent multivariable systems, we will introduce the technique of signal flow graphs. This technique is more compact than regular block diagram algebra for complex diagrams involving more than one loop.

13-5.1 Signal Flow Graphs (SFG)

Block diagram algebra, presented in Chapter 3, is useful to analyze the response of linear control systems as long as the diagram is not too complex. When the diagram involves two or more interacting loops, a technique known as signal flow graphs (SFG) simplifies the algebraic manipulations required by block diagram algebra. Like block diagrams, signal flow graphs are graphical representations of the transfer functions that describe control systems. This section presents a brief introduction to the SFG technique. For a more detailed treatment, see any textbook on linear systems analysis, such as Kuo (1991).

Figure 13-5.1 shows the graphical representation of a feedback control loop by block diagram and by signal flow graph (SFG). The following definitions are important for understanding and using SFG.

Block diagram
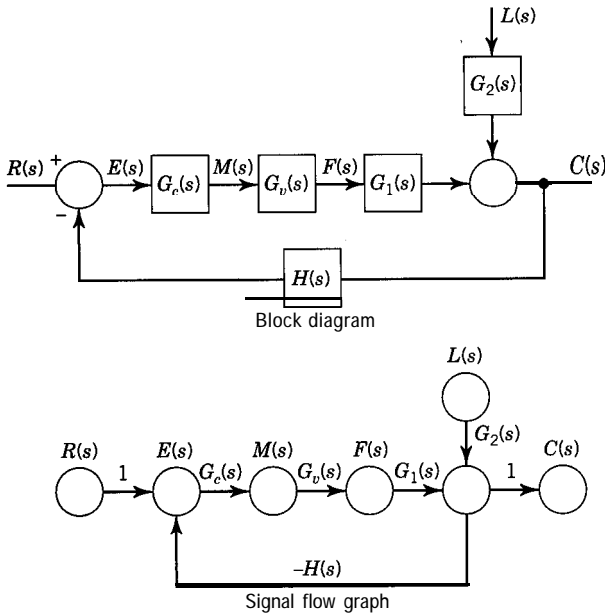
Signal flow graph

Figure 13-5.1 Block diagram and signal flow graph represen-
tation of a feedback loop.

- A node, shown as a circle or a period, represents a variable or signal.
- A **branch**, shown as a line with an arrow, represents the transfer function that relates the nodes it joins.
- A **path** is a continuous unidirectional succession of branches along which no node is passed more than once.
- A **path transferfunction** is the product of the branch transfer functions encountered in traversing the path.
- A **loop** is a path that originates and terminates at the same node.

Referring to the SFG of Fig. 13-5.1, we find the following values.

Path: $\qquad R(s) \rightarrow E(s) \rightarrow M(s) \rightarrow F(s) \rightarrow C(s) \rightarrow C(s)$

Path transfer function: $\qquad G_c(s)G_v(s)G_1(s)$

Loop: $\qquad E(s) \rightarrow M(s) \rightarrow F(s) \rightarrow C(s) \rightarrow E(s)$

Figure 13-5.2 shows some rules necessary for applying SFG algebra. With these rules and the previous definitions, the transfer functions of a graph can be obtained. The Mason's Gain Formula provides a compact guide to the development of the transfer functions of a complex graph.

$$T = \frac{\sum_i P_i \Delta_i}{\Delta}$$

(13-5.1)
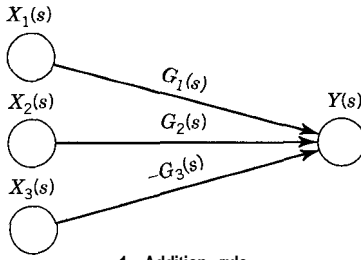
1. Addition rule

$$Y(s) = G_1(s)X_1(s) + G_2(s)X_2(s) - G_3(s)X_3(s)$$
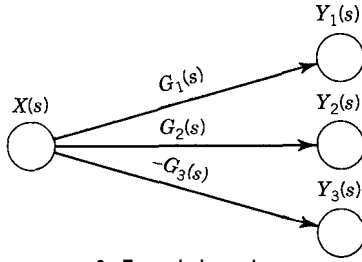
2. Transmission rule

$$Y_1(s) = G_1(s)X(s)$$
$$Y_2(s) = G_2(s)X(s)$$
$$Y_3(s) = -G_3(s)X(s)$$
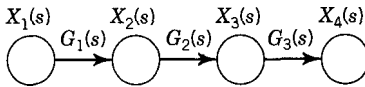
3. Multiplication rule

$$X_4(s) = G_3(s)X_3(s) = G_3(s)G_2(s)X_2(s) = G_3(s)G_2(s)G_1(s)X_1(s)$$

Figure 13-5.2 Rules for signal flow graph algebra.

where

T = transmittance (transfer function) between input and output nodes

P_i = product of the transfer functions in the ith forward path between input and output nodes

A = determinant of the graph, given by

$$A = 1 - \sum L_1 + \sum L_2 - \sum L_3 + \ldots \qquad (13\text{-}5.2)$$

A_i = A evaluated with loops touching P_i eliminated

L_1 = product of the transfer functions in each loop. ΣL_1 is the sum of all the loops in the graph.

L_2 = product of the transfer functions of two nontouching loops. ΣL_2 is the sum of all possible combinations of nontouching loops **taken** two at a time.

L_3 = product of the transfer functions of three nontouching loops. ΣL_3 is the sum of all possible combinations of nontouching loops taken three at a time.

and so on. Touching loops are those that have at least one node in common.

Let us look at several examples of the use of SFG to obtain the desired transfer functions.

EXAMPLE 13-5.1

Consider the simple block diagram shown in Fig. 13-5.1. Using SFG, determine the closed-loop transfer functions

$$\frac{C(s)}{R(s)} \quad \text{and} \quad \frac{C(s)}{L(s)}$$

SOLUTION

For the first transfer function, the only path between the input node, $R(s)$, and the output node, $C(s)$ is

$$P_1 = G_c(s)G_v(s)G_1(s)$$

The only loop in the graph is

$$L_1 = - G_c(s)G_v(s)G_1(s)H(s)$$

and, from Eq. 13-5.2, the determinant of the graph is

$$A = 1 + G_c(s)G_v(s)G_1(s)H(s)$$

Because the only loop is touching the forward path, A, = 1. Then, from Eq. 13-5.1,

$$\frac{C(s)}{R(s)} = \frac{G_c(s)G_v(s)G_1(s)}{1 + G_c(s)G_v(s)G_1(s)H(s)}$$

which is the same result we obtain by block diagram algebra.

For the second transfer function, the only path between the input node, $L(s)$, and the output node, $C(s)$, is

$$P_1 = G_2(s)$$

Because the loop in the graph also touches this path, the transfer function, from Eq. 13-5.1,

$$\frac{C(s)}{L(s)} = \frac{G_2(s)}{1 + G_c(s)G_v(s)G_1(s)H(s)}$$

This is also the same result we obtain by block diagram algebra.

EXAMPLE 13-5.2

For the block diagram of the cascade control system of Fig. $13\text{-}5.3a$, obtain $\dfrac{C_2(s)}{R(s)}$,

SOLUTION

The first step is to draw the SFG as shown in Fig. $13\text{-}5.3b$. The only path between $R(s)$ and $C_2(s)$ is

$$P_1 = G_{c_2}(s)G_{c_1}(s)G_v(s)G_1(s)G_2(s)$$

The two loops in the graph are

$$L_{1_1} = - G_{c_1}(s)G_v(s)G_1(s)H_1(s)$$

$$L_{1_2} = - G_{c_2}(s)G_{c_1}(s)G_v(s)G_1(s)G_2(s)H_2(s)$$

These two loops touch each other, so they do not form a pair of nontouching loops. The determinant of the graph, from Eq. 13-5.2, is

$$\Delta = 1 + G_{c_1}G_v(s)G_1(s)H_1(s) + G_{c_2}(s)G_{c_1}(s)G_v(s)G_1(s)G_2(s)H_2(s)$$
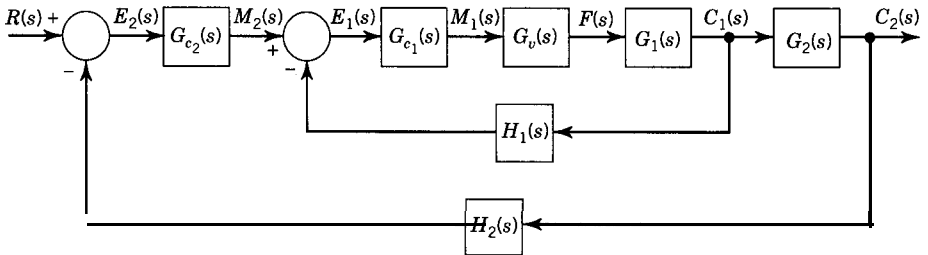
(a)
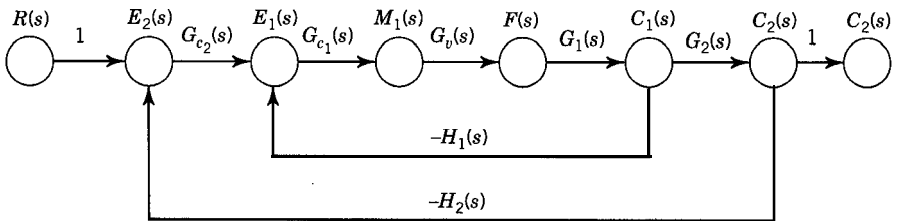
(b)

Figure 13-5.3 Block diagram and signal flow graph of a cascade control system.

Because both loops touch the path, $\Delta_1 = 1$, and the transfer function, from Eq. 13-5.1, is

$$\frac{C_2(s)}{R(s)} = \frac{G_{c_2}(s)G_{c_1}(s)G_v(s)G_1(s)G_2(s)}{1 + G_{c_1}(s)G_v(s)G_1(s)H_1(s) + G_{c_2}(s)G_{c_1}G_v(s)G_1(s)G_2(s)H_2(s)}$$

This is the same result we obtained by block diagram algebra in Chapter 3 after considerable algebraic manipulation.

13-5.2   Dynamic Analysis of a 2 x 2 System

Figure 13-5.4a shows the general block diagram for a 2 X 2 control system in which the field devices, valve, process, and transmitter have been combined into the process transfer functions G,,(s). For simplicity, the disturbances have been omitted from the diagram.

(a)

(b)
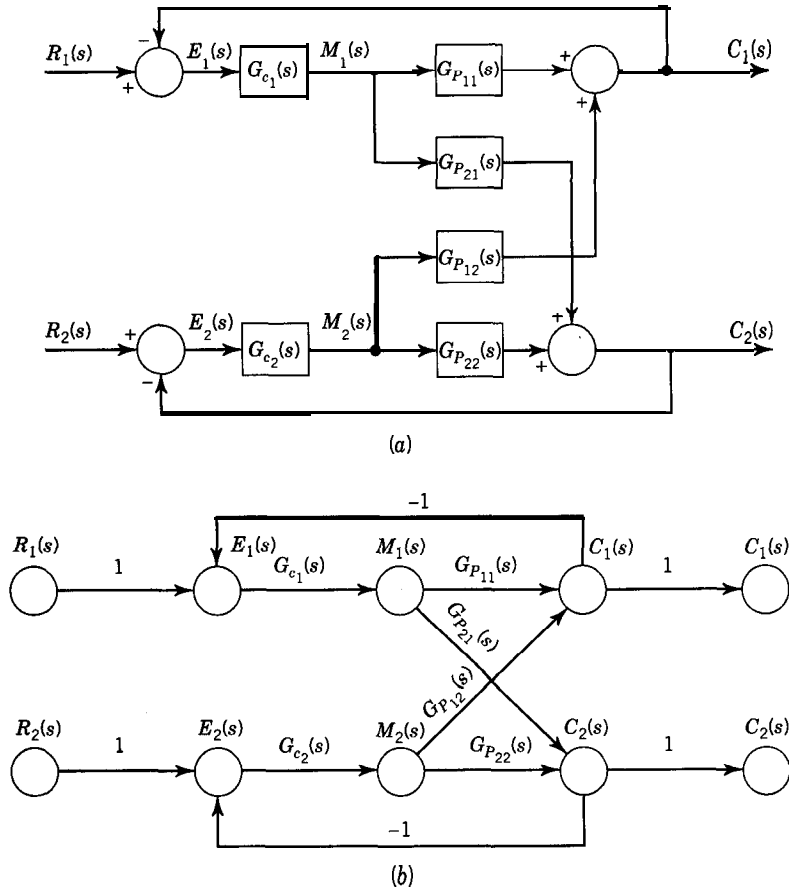
Figure 13-5.4 General 2 X 2 system. (a) Block diagram. (b) Signal flow graph.

To obtain the closed-loop transfer functions for the diagram, we first draw the corresponding SFG, Fig. 13-5.46. The graph has three loops, two of which do not touch each other.

$$L_{1_1} = -G_{c_1}(s)G_{P_{11}}(s)$$

$$L_{1_2} = -G_{c_2}(s)G_{P_{22}}(s)$$

$$L_{1_3} = G_{c_1}(s)G_{P_{21}}(s)G_{c_2}(s)G_{P_{12}}(s)$$

Loops L_{1_1} and L_{1_2} are the familiar feedback loops. Loop L_{1_3} is more complex and goes through both controllers and the "cross" process transfer functions. The determinant of the graph is then, from Eq. 13-5.2,

$$\begin{aligned}
\Delta = 1 &+ G_{c_1}(s)G_{P_{11}}(s) + G_{c_2}(s)G_{P_{22}}(s) \\
&- G_{c_1}(s)G_{P_{21}}(s)G_{c_2}(s)G_{P_{12}}(s) \\
&+ G_{c_1}(s)G_{P_{11}}(s)G_{c_2}(s)G_{P_{22}}(s)
\end{aligned} \qquad \textbf{(13-5.3)}$$

where the last term is the product of the two nontouching loops.

There are two paths between $R_,(s)$ and $C_,(s)$:

$$P_1 = G_{c_1}(s)G_{P_{11}}(s)$$

$$P_2 = -G_{c_1}(s)G_{P_{21}}(s)G_{c_2}(s)G_{P_{12}}(s)$$

The first of these paths does not touch the bottom loop, and the other one touches all three loops.

$$\Delta_1 = 1 + G_{c_2}(s)G_{P_{22}}(s)$$

$$\Delta_2 = 1$$

From Eq. 13-5.1, the transfer function is

$$\frac{C_1(s)}{R_1(s)} = \frac{G_{c_1}(s)G_{P_{11}}(s)[1 + G_{c_2}(s)G_{P_{22}}(s)] - G_{c_1}(s)G_{P_{21}}(s)G_{c_2}(s)G_{P_{12}}(s)}{A}$$

There is only one path between $R_,(s)$ and $C_,(s)$, and it touches all three loops in the graph. Thus the transfer function, from Eq. 13-5.1, is

$$\frac{C_2(s)}{R_1(s)} = \frac{G_{c_1}(s)G_{P_{21}}(s)}{A}$$

By the same procedure, we can obtain the transfer functions between $R_2(s)$ and the two controlled variables. They are

$$\frac{C_1(s)}{R_2(s)} = \frac{G_{c_2}(s)G_{P_{12}}(s)}{\Delta}.$$

$$\frac{C_2(s)}{R_2(s)} = \frac{G_{c_2}(s)G_{P_{22}}(s)[1 + G_{c_1}(s)G_{P_{11}}(s)] - G_{c_2}(s)G_{P_{12}}(s)G_{c_1}(s)G_{P_{21}}(s)}{A}.$$

As with any dynamic system, the response is determined by the location of the roots of the denominator polynomial or characteristic equation. To obtain the characteristic equation, just set the determinant of the graph equal to zero.

$$\Delta = 0$$

It is enlightening to rearrange the determinant, Eq. 13-5.3, into the following form:

$$\Delta = [1 + G_{c_1}(s)G_{P_{11}}(s)][1 + G_{c_2}(s)G_{P_{22}}(s)]$$
$$- G_{c_1}(s)G_{P_{12}}(s)G_{c_2}(s)G_{P_{21}}(s) = 0 \quad \textbf{(13-5.4)}$$

The roots of this equation determine the stability and response of the interacting 2 X 2 system. As in any other system, the response will oscillate if there is at least one pair of complex conjugate roots, and it will be unstable if any of the real roots is positive or if a pair of complex conjugate roots has a positive real part (see Chapter 2). Furthermore, because the denominator is the same for all four transfer functions, the response characteristics are the same for the two controlled variables.

Equation 13-5.4 tells us the following:

- The tuning of each controller affects the response of both controlled variables, because it affects the roots of the common characteristic equation.
- The effect of interaction on one loop may be eliminated by interrupting the other loop. For example, if controller 2 is switched to manual, then $G_{,,}(s) = 0$, and the characteristic equation for the system becomes

$$1 + G_{c_1}(s)G_{P_{11}}(s) = 0 \qquad (13\text{-}5.5)$$

This is the characteristic equation for control loop 1 taken by itself. The change in the characteristic equation means that if we tune controller 1 with controller 2 in manual, then the response of control loop 1 will change when loop 2 is switched back to automatic. Similarly, if controller 1 is switched to manual, then $G_{,,}(s) = 0$, and the characteristic equation of the system becomes

$$1 + G_{c_2}(s)G_{P_{22}}(s) = 0 \qquad (13\text{-}5.6)$$

This is the characteristic equation of control loop 2 taken separately.

- For interaction to affect the response of the loops, it must act both ways. That is, each manipulated variable must affect the controlled variable of the other loop. This is because if either of the "cross" process transfer functions, $G_{P_{12}}(s)$ or $G_{P_{21}}(s)$, is equal to zero, then the characteristic equation of the system becomes

$$[1 + G_{c_1}(s)G_{P_{11}}(s)][1 + G_{c_2}(s)G_{P_{22}}(s)] = 0$$

The roots of this characteristic equation are the same as the roots of the characteristic equations of the feedback loops taken separately, Eqs. 13-5.5 and 13-5.6. Thus there would be no interaction. Note, however, that one of the manipulated variables will still affect both controlled variables.

The reader is encouraged to study Eq. 13-5.4 to gain an understanding of how these important properties of interacting systems follow from the characteristic equation of the system.

EXAMPLE 13-5.3

Study the effect of interaction on the dynamic response of the blending tank of Fig. 13-1.1a. The tank holds 1000 lb of solution, and the inlet stream compositions are 10 and 30 mass % salt, respectively. It is desired to produce 100 lb/min of product solution with a concentration of 20 mass % salt. The control valves are linear with constant pressure drop and sized to pass 150 lb/min when fully opened. They have negligible time constants. The flow transmitter is linear with a range of 0 to 150 lb/min and negligible time constant. The analyzer transmitter has a range of 5 to 35 mass % salt and can be represented by a first-order lag with a time constant of 5.0 min. Assume the composition controller is PI tuned for quarter decay ratio response with the flow controller in manual. How does the decay ratio change when the flow controller, also PI, is switched to automatic?

SOLUTION

From the steady-state mass balances, Eq. 13-1.1, we determine that each of the design inlet flows is equal to 50 lb/min. The relative gains, determined in Example 13-2.3, are, from Eq. 13-2.11, equal to 0.50, no matter how the loops are paired. We will assume that the flow controller manipulates the dilute stream and that the composition controller manipulates the concentrated stream.

The transfer functions of the tank were determined in Example 13-3.1, Eq. 13-3.7, giving a time constant for the tank $\tau = (1000 \text{ lb})/(100 \text{ lb/min}) = 10.0$ min and gains $K_{x_1} = (10 - 20)/100 = -0.10$ mass %/(lb/min) and $K_{x_2} = (30 - 20)/100 = 0.10$ mass %/(lb/min). The valve and transmitter gains are calculated as in Sections 5-1 and 5-2.

$$K_{v_1} = K_{v_2} = \frac{150}{100} = 1.50 \frac{\text{lb/min}}{\%CO}$$

$$K_{Tw} = \frac{100}{150} = 0.667 \; \frac{\% \, TO}{lb/min}$$

$$K_{Tx} = \frac{100}{35 - 5} = 3.333 \; \frac{\% \, TO}{mass \, \%}$$

These numbers are then used to determine the transfer functions of the valves and transmitters, following the assumptions in the statement of the problem.

$$G_{,,}(s) = G_{,,}(s) = 1.50 \; \frac{lb/min}{\%CO}$$

$$H_{,,,}(s) = K_{Tw} = 0.667 \frac{\% \, TO}{lb/min}$$

$$H_x(s) = \frac{K_{Tx}}{\tau_T s + 1} = \frac{3.333}{5.0s + 1} \; \frac{\% \, TO}{mass \, \%}$$

The combined process transfer functions, all in %TO/%CO, are

$$G_{P_{11}}(s) = G_{v_1}(s)G_{11}(s)H_w(s) = (1.5)(1.0)(0.667) = 1.0$$

$$G_{P_{12}}(s) = G_{v_2}(s)G_{12}(s)H_w(s) = (1.5)(1.0)(0.667) = 1.0$$

$$G_{P_{21}}(s) = G_{v_1}(s)G_{21}(s)H_x(s) = \frac{-0.50}{(100s + 1)(50s + 1)}$$

$$G_{P_{22}}(s) = G_{v_2}(s)G_{22}(s)H_x(s) = \frac{0.50}{(10.0s + 1)(5.0s + 1)}$$

Typical tuning of flow controllers includes gains less than unity and fast reset times, say $K_{c_1} = 0.50$ %CO/%TO, and $\tau_{I_1} = 0.10$ min. With the composition loop opened, the characteristic equation of the flow control loop is

$$1 + G_{c_1}(s)G_{P_{11}}(s) = 1 + 0.5\left(1 + \frac{1}{0.1s}\right)1.0 = 0$$

This equation has a root at $s = -3.333 \; min^{-1}$, which corresponds to a closed-loop time constant of 0.30 min.

With the flow controller opened, the composition controller is tuned for quarter decay ratio response. The resulting tuning parameters are $K_{c_2} = 25$ %CO/%TO and $\tau_{I_2} = 12$ min. The characteristic equation when only the composition loop is closed is

$$1 + 25\left(1 + \frac{1}{12s}\right)\frac{0.5}{(10s + 1)(5.0s + 1)} = 0$$

The roots for this equation are at -0.083 min^{-1} and $0.109 \pm i0.490$ min^{-1}. The real root essentially cancels a numerator term caused by the controller reset mode, so the response is dominated by the complex conjugate roots. The period of the oscillations (see Chapter 2) is $2\pi/0.490 = 12.8$ min, the decay ratio is 0.25, and the 1% settling time is $-5/(-0.109) = 46$ min.

When both controllers are switched to automatic, the characteristic equation is obtained by substituting the process and controller transfer functions into Eq. 13-5.4. The resulting roots are at -0.083, -3.38, and $-0.083 \pm i0.698$ min^{-1}. Once again, the first real root is essentially canceled by a numerator term in the transfer function that is caused by the reset mode of the composition controller. The second real root is essentially the same as that of the flow control loop when it was the only one closed, but now the complex roots will cause both the composition and the flow to oscillate. The period is $2\pi/0.698 = 9.0$ min, the decay ratio is 0.47, and the 1% settling time is 60 min.

In the preceding example, the basic response of the flow controller was not affected by the interaction. This is because the flow controller is very fast and controllable so that its response is insensitive to the process gain. The composition controller is affected by the interaction in that it oscillates faster and the oscillations decay more slowly. This is because its gain increases by a factor of 2 when the flow control loop is closed.

The results obtained in this example assume linear approximations of the blending tank. Nonlinear behavior may cause the system to go unstable when both loops are closed and the process conditions vary significantly from the design values.

134.3   Controller Tuning for Interacting Systems

It is evident from the discussion in the preceding section that interaction makes the tuning of multivariable control systems a difficult task. However, proper procedure can overcome the difficulties. This section presents a brief outline of the tuning procedure. For a more detailed discussion, see Shinskey (1988).

The first step in the tuning procedure, after proper pairing of the controlled and manipulated variables, is to determine the relative speed of the loops. If one loop is much faster than the other-say, having a dominant time constant 5 times smaller— then it is tuned first with the other loop open. Then the slower loop is tuned with the faster loop closed. The assumption is that the closing of the slower loop has a minor effect on the performance of the faster loop, because the faster loop can respond rapidly to the actions of the slower loop. When the slower loop is tuned with the faster loop closed, the tuning takes the interaction fully into account. Flow and pressure loops are usually much faster than temperature and composition loops. For example, in the blending tank of Example 13-5.3, the flow loop had a closed-loop time constant of 0.3 min, whereas the analyzer loop had a dominant time constant of 10 min. In that example, we saw that the time constant of the flow control loop did not change when the composition loop was closed.

If the two loops are of about the same speed of response, then de-tuning one loop by setting a small gain and a long reset time will have the same effect of making that loop slow. This will reduce its effect on the response of the other loop, because the de-

tuned loop will appear to be open. Such an approach is followed when control of one variable is more important than control of the other. The controller on the less important variable is the one that is de-tuned.

Finally, when both loops are of about the same speed of response and of equal importance, each controller should be tuned with the other loop open. Then the tuning should be adjusted to account for interaction. If the interaction is positive (see Section 13-1), then the two loops help each other. This means that the dynamic effects of interaction are usually insignificant. In such a case, adjust each controller gain from its initial (open-loop) tuning by multiplying it by the relative gain for the pair. The principle is that such an adjustment will keep a constant overall gain for each loop. For the loop gain to remain the same,

$$K_c K_{P_{ii}} = K'_c K'_{P_{ii}}$$

where the primes denote the controller and process gains when the other loop is closed. Solving for the controller gain with the other loop closed yields

$$\boxed{K'_{c_i} = K_{c_i} \frac{K_{P_{ii}}}{K'_{P_{ii}}} = K_{c_i} \mu_{ii}}$$ (13-5.5)

where we have used the definition of the relative gain, Eq. 13-2.3.

For example, in the blending tank of Example 13-5.3, the relative gain for either pairing is 0.5. This means that the gain obtained for each controller when the other loop is open should be reduced by half to account for the increase in the process gain when the other loop is closed. In that case, however, the flow control loop is so fast that it is not necessary to adjust its gain. The reader is invited to check that when the gain of the composition controller is reduced to $(25)(0.5) = 13$ %CO/%TO, the roots of the characteristic equation become -0.083, -3.36, and $-0.0952 \pm i0.501$ min^{-1}, with a period of oscillation of 12.5 min, a decay ratio of 0.3, and a 1% settling time of 52 min. These numbers are close to those for each loop taken separately when the composition controller gain is 25 %CO/%TO.

If the interaction is negative, then its effect on the dynamic response is quite significant because the two loops fight each other. This means that inverse response is possible when the two loops are closed. Inverse response causes the controlled variable to move the wrong way first and is quite detrimental to the performance of the loop. Because of this, adjustment of the controller gains must be done by trial and error, after both loops are closed.

Tuning of n × n Systems. The tuning procedure we have just presented for 2 X 2 systems can be extended to any number of loops. What needs to be done is first to rank the loops in order of speed of response and then to tune those loops that are much faster than the others. In doing this, one must adjust the gains of the fast loops to account for interaction between them. Then the slower loops must be ranked in order of decreasing importance and the less important loops de-tuned if possible. Finally, the gains of the slow loops must be adjusted to compensate for the effect of interaction. Equation

13-5.5 is suitable for doing the adjustment for those loops for which the interaction is positive. Recall that the interaction is positive when the relative gain for a pair is between 0 and + 1.

13-6   SUMMARY

This chapter has presented the design and tuning of multivariable feedback control systems. We first explained the effect of interaction and then introduced its measurement from the open-loop gains. The relative gains are used not only as quantitative measures of interaction but also to pair the controlled and manipulated variables in such a way as to minimize the effect of interaction. The proper pairing has each pair with a relative gain near unity. The design of decouplers was presented next. Decoupling can totally or partially eliminate the effect of interaction. Signal flow graphs were introduced and used to determine the closed-loop transfer functions of multivariable systems. Finally, a procedure for tuning interacting loops was discussed.

REFERENCES

1. Bristol, E. H. 1966. "On a New Measure of Interaction for Multivariable Process Control." *IEEE Transactions on Automatic Control* AC-l 1 (January):133–134.
2. Cutler, C. R., and B. L. Ramaker. 1979. "DMC-A Computer Control Algorithm." Paper read at AIChE 1979 Houston Meeting, Paper #516, AIChE, New York.
3. Hulbert, D. G., and E. T. Woodburn. 1983. "Multivariable Control of a Wet Grinding Circuit." *AIChE Journal,* March.
4. IdCom (Identification Command), SetPoint, Incorporated, Houston, TX.
5. Kuo, B. *C.* 1991. *Automatic Control Systems.* 6th ed. Englewood Cliffs, N.J.: Prentice-Hall.
6. MathCad. 1994. MathSoft, Incorporated, Cambridge, Mass.
7. MatLab. 1994. The MathWorks, Incorporated, Natick, Mass.
8. Moore, R. D., and A. B. Corripio. 1991. "On-Line Optimization of Distillation Columns in *Series.*" *Chemical Engineering Communications* 106:71–86.
9. Shinskey, F. *G.* 1981. *Controlling Multivariable Processes.* Research Triangle Park, N.C., ISA.
10. Shinskey, F. *G.* 1988. *Process Control Systems, Application, Design, and Tuning.* 3rd ed. New York: McGraw-Hill.
11. Taylor, R. and H. Kooijman. 1995. *ChemSep.* Potsdam, N.Y.: Chemical Engineering Department, Clarkson University.

PROBLEMS

13-1. For each of the processes of Fig. 13-1.1, using your knowledge of process mechanisms and basic principles, explain what causes interaction between the variables shown. Is the interaction negative or positive? List your assumptions.

13-2. Control of Caustic Dilution Process. Shown in Fig. P13-1 is a tank used to mix a 50% solution of caustic with demineralized water to produce a dilute solution. The specified product composition is 30 mass % NaOH, and at design conditions, the product flow is 40 klb/h. It is desired to control the product flow, w_P (FC-3), and composition, x_P (AC-3), by manipulating the set points on the flow
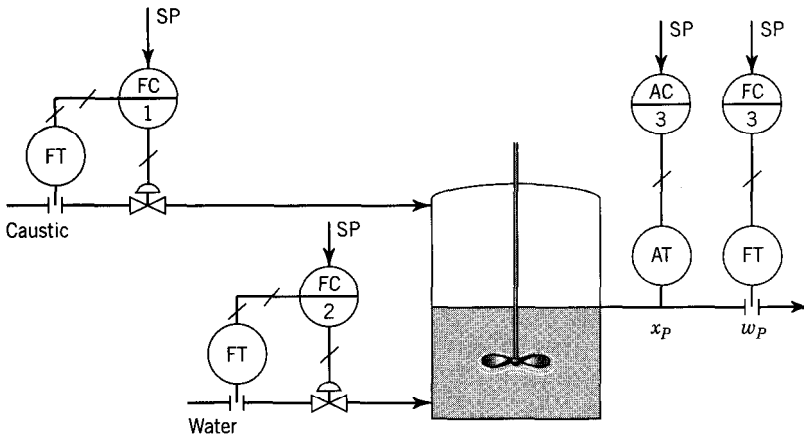
Figure P13-1 Caustic blending tank for Problem 13-2.

controllers on the two inlet streams (FC-1 and FC-2). The tank is assumed perfectly mixed, and the total mass M of the solution in the tank is constant. Obtain the steady-state open-loop gains for the 2 X 2 system and the relative gains at design conditions. Is the interaction positive or negative? Which pairing of the controlled and manipulated variables minimizes the effect of interaction? By how much does the gain of the product composition loop increase or decrease (specify which) when the product flow loop is closed?

13-3. *Automatic Control of a Household Shower.* The design of automated showers for Mr. Trump's mansion requires control of the temperature and flow of the water to the shower head. The design calls for a system that can deliver 3 gpm of water at 110°F by mixing hot water at 170°F with "cold" water at 80°F. The density of water is 8.33 lb/gal, and its specific heat is 1.0 Btu/lb-°F. Control valves are used to manipulate the flows of cold and hot water. A flow transmitter and a temperature transmitter are also used in each shower.

(a) Show that steady-state mass and energy balances result in the following relationships, assuming constant density and specific heat of the water.

$$f_0 = f_1 + f_2$$

$$f_0 T_0 = f_1 T_1 + f_2 T_2$$

(b) Calculate the required flows of hot and cold water at design conditions and the steady-state open-loop gains. Specify the units.

(c) Calculate the relative gains for the system, and decide which of the two inlet flows is to control the flow and which is to control the temperature so that the effect of interaction is minimized.

13-4. *Control of an Evaporator.* Consider the evaporator of Fig. 13-3.4. The level in the evaporator is controlled very tightly by manipulating a valve on the feed line, because this is the largest of the inlet and outlet streams (LC). It is desired to control the feed flow, w_F (FC) and the product composition, x_P (AC), by manipulating the signals to the valves on the product, m_P, and steam, m_S,

streams. The evaporator economy is approximately constant at E lb of vapor per lb of steam, and the inlet composition is x_F, weight % solute.

(a) Show, from steady-state balances on the evaporator, that the relationships between the variables are

$$w_F = w_P + Ew,$$

$$w_P x_P = x_F(w_P + Ew_S)$$

(b) Derive the general formulas for the steady-state open-loop gains and for the relative gains. Is the interaction positive or negative?

(c) Develop a general pairing strategy that will minimize the effect of interaction for any evaporator depending on the design values of the process variables.

13-5. The evaporator of Problem 13-4 is designed to concentrate a 30% solution of caustic to produce a product solution with SO mass % NaOH, and at design conditions, the feed flow is 80 klb/h. The evaporator economy is 0.9 klb of vapor per klb of steam. Solve the problem for these design conditions. By how much does the gain of the product composition control loop increase or decrease (specify which) when the feed flow control loop is closed?

13-6. *Distillation Product Composition Control.* An example of highly interactive control loops is control of the top and bottom compositions in a distillation column. Two tests were performed on a simulation of a butane-pentane column: one was a change in steam flow keeping the reflux flow constant, and the other was a change in reflux flow keeping the steam flow constant. The following table summarizes the flows and compositions for both of the tests, after steady state is established.

	Base	Test 1	Test 2
Steam flow, klb/h	24.0	25.0	24.0
Reflux flow, lbmol/h	70.0	70.0	75.0
Bottoms percent butane	6.22	3.08	8.77
Distillate percent butane	93.50	91.79	96.88

The controlled variables are the percent butane in the bottoms and distillate products, and the manipulated variables are the steam and reflux flow set points. Calculate the steady-state open-loop gains, and the relative gains for the two possible pairings. Is the interaction positive or negative? Which variable should be manipulated to control the distillate composition so that the interaction between the two loops is minimized? What is the steady-state gain for the distillate composition loop when the bottoms composition loop is closed?

13-7. For the gasoline blending control system of Example 13-2.5, calculate the required flows of the three feed streams if the specified gasoline octane number is 89.0 instead of 87.0. Then recalculate the steady-state open-loop gains and the relative gains. Does the pairing of controlled and manipulated variables that minimizes interaction change from the one obtained in the example? Design a

static **decoupler** for this system and calculate the gains of the decoupled system, as in Example 13-3.3.

13-8. Control of Wet *Grinding Circuit.* Hulbert and **Woodburn** (1983) present the multivariable control of a wet grinding circuit such as the one shown in Fig. **P13-2.** The controlled variables are the torque required to turn the mill (TOR), the density of the cyclone feed (DCF), and the flow from the mill (FML). These variables were selected on the basis of their importance to the metallurgical process. The torque and flow from the mill describe the operation of the mill, and the density of the cyclone feed describes the operation of the slurry tank. The flow from the mill is not measured directly but is derived from a combination of other measurements and a mass balance around the mill. However, for simplicity, a single transmitter is shown in the diagram.

The manipulated variables are the flows of the three feed streams: solids (SF) and water to the mill (MW) and water to the slurry tank (SW). The feed rate of solids to the mill is manipulated by adjusting the speed of the conveyor carrying **the** solids.

Open-loop step tests in each manipulated variable result in the following transfer functions:

	SF, **kg/s**	MW, kg/s	SW, **kg/s**
TOR, **N·m**	$\dfrac{119}{217s + 1}$	$\dfrac{153}{337s + 1}$	$\dfrac{-21}{10s + 1}$
FML, **m³/s**	$\dfrac{0.00037}{500s + 1}$	$\dfrac{0.000767}{33s + 1}$	$\dfrac{-0.000050}{10s + 1}$
DCF, **kg/m³**	$\dfrac{930}{500s + 1}$	$\dfrac{-667e^{-320s}}{166s + 1}$	$\dfrac{-1033}{47s + 1}$

where all time parameters are in seconds.
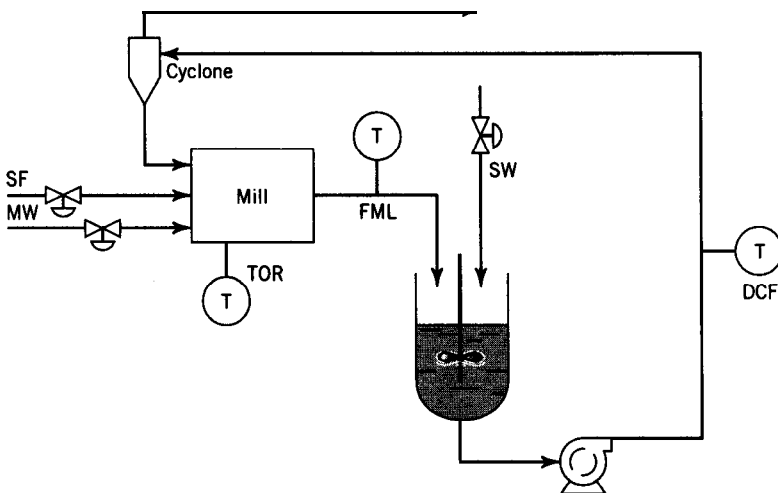
Figure P13-2 Wet grinding circuit for Problem 13-8.

(a) Calculate the relative gains, and select the pairings of the controlled and manipulated variables that minimize the effect of interaction.

(b) Using the pairing of part (a), draw the block diagram and design a linear decoupler for the system. Show the instrumentation diagram for the complete control system.

13-9. *Decoupler Design for Distillation Column.* In the distillation column of Fig. 13-2.3, the flows of reflux and steam to the reboiler are manipulated to control the distillate and bottoms product purities. Open-loop step tests on each of these manipulated variables result in the following transfer functions:

$$Y_D(s) = -3.0 \left(\frac{1 - 0.11s}{1 + 0.35s}\right) M_R(s) + \left(\frac{0.75}{1 + 0.35s}\right) M_S(s)$$

$$X_B(s) = \left(\frac{2.5}{1 + 0.35s}\right) M_R(s) - 3.5 \left(\frac{1 + 0.25s}{1 + 0.35s}\right) M_S(s)$$

where $Y,(s)$ is the composition of the heavy key in the distillate, $X_B(s)$ is that of the light key in the bottoms, in %TO, and $M_R(s)$ and $M_S(s)$ are the signals to the reflux and steam valves, respectively, in %CO. The time parameters are in hours.

(a) Calculate the relative gains for this system, and determine the correct way to pair the controlled and manipulated variables. Do the loops help or fight each other?

(b) Draw the block diagram and design the decouplers for this system. Briefly discuss any implementation problems and suggest modifications to ensure that the control system is stable.

13-10. For the distillation column of Problem 13-9, draw the signal flow graph of the un-decoupled system. Obtain the four closed-loop transfer functions between the two set points and the two controlled variables, and give the characteristic equation of the system. Assuming PI controllers with gains of 1.0 and reset times of 0.35 h, obtain the roots of the characteristic equation of the system with either controller on manual and with both loops closed.

13-11. *Decoupler Design for* 2 X 2 *Process.* A process has two controlled variables that are affected by two manipulated variables and one disturbance, U(s). The transfer functions are

$$C_1(s) = G_{P_{11}}(s)M_1(s) + G_{P_{12}}(s)M_2(s) + G_{P_{1U}}(s)U(s)$$

$$C_2(s) = G_{P_{21}}(s)M_1(s) + G_{P_{22}}(s)M_2(s) + G_{P_{2U}}(s)U(s)$$

where

$$G_{P_{11}} = \frac{0.81e^{-0.6s}}{1.4s + 1} \qquad G_{P_{12}} = \frac{1.2e^{-1.1s}}{2.4s + 1} \qquad G_{P_{1U}} = \frac{0.5}{2.2s + 1}$$

$$G_{P_{21}} = \frac{1.1e^{-0.3s}}{1.5s + 1} \qquad G_{P_{22}} = \frac{0.90e^{-s}}{2.0s + 1} \qquad G_{P_{2U}} = \frac{-1.5}{1.8s + 1}$$
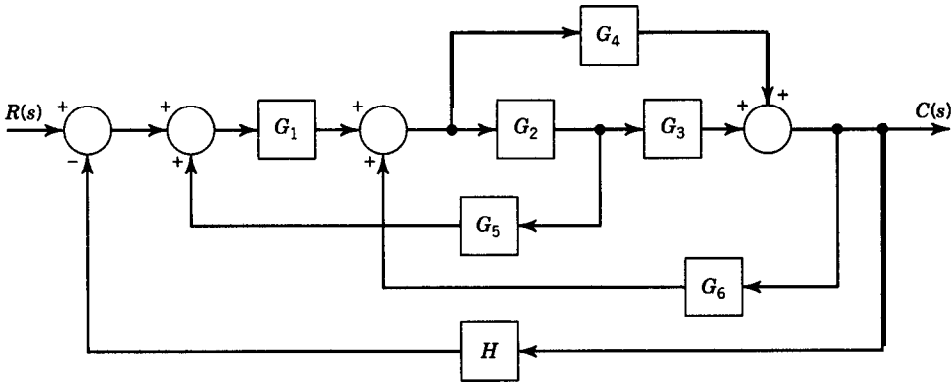
Figure P13-3 Block diagram for Problem 13-13.

(a) Calculate the relative gains, and select the correct pairing for this system.
(b) Draw the complete block diagram and the signal flow graph for the system with the correct pairing, and write the characteristic equation. Assume PI controllers with quarter decay ratio tuning.
(c) Determine the closed-loop transfer functions between the disturbance and the two controlled variables.
(d) Design a decoupler for this system, and show its implementation on the block diagram.

13-12. Show that if the exact decoupler is implemented on the system represented by the block diagram of Fig. 13-3.2, the system is indeed decoupled. Hint: Draw the signal flow graph and obtain the transfer functions between the controlled variables and their set points.

13-13. Draw the signal flow graph corresponding to the block diagram of Fig. P13-3, and obtain the closed-loop transfer function $C(s)/R(s)$.

13-14. Do Problem 13-13 for the block diagram of Fig. P13-4.

13-15. Do Problem 13-13 for the block diagram of Fig. P13-5.

13-16. Do Problem 13-13 for the block diagram of Fig. P3-11.

13-17. Do Problem 13-13 for the block diagram of Fig. P3-12.
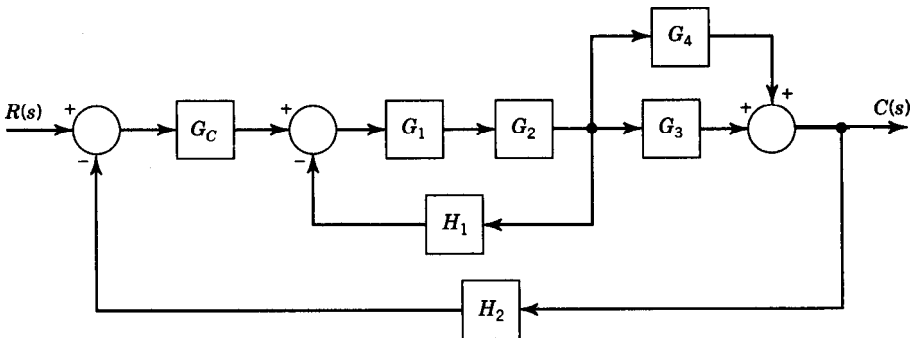
Figure P13-4 Block diagram for Problem 13-14.
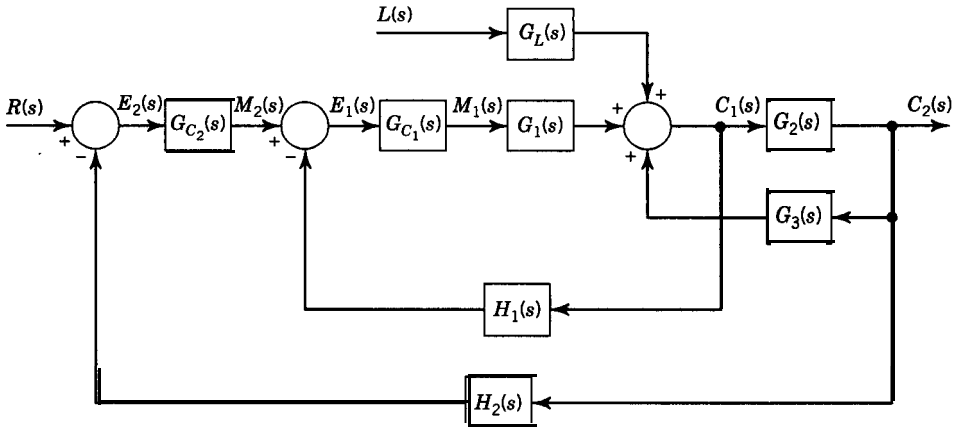
Figure P13-5 Block diagram for Problem 13-15.

13-18. Do Problem 13-13 for the block diagram of Fig. **P3-13.**

13-19. Design a decoupler for the blending tank described in Example 13-5.3. Show the instrumentation diagram for implementing the decoupler.

13-20. Design, from basic principles, a nonlinear static decoupler for the evaporator of Problem 13-4.

Chapter **14**

Mathematical Tools for Computer Control Systems

Up to this point, we have discussed control systems as though all the signals varied continuously with time. The original mechanical, hydraulic, pneumatic, and electronic control systems consisted of dedicated instruments, each operating continuously on its input signals and generating its output signal continuously in time. Modem control systems use digital devices that perform their control functions sequentially, one at a time, dedicating a small fraction of the time to each task. This requires that the process variables, which vary continuously with time, be sampled in time for sequential operations to be performed on them. The outputs to the process must be updated whenever their values are computed and must be held constant until the next update. This is because the process requires a sustained signal, not pulses, for its operation. In this chapter, we will learn the mathematical tools required to analyze *sampled-data control systems.* These systems are represented by difference (as opposed to differential) equations and *can* be analyzed by *z-transforms,* which are to sampled dynamic systems what Laplace transforms are to continuous systems. We will see that under certain conditions, sampled-data devices can be represented by ***pulse transfer functions*** in the z-transform variable, which are similar to Laplace transfer functions. We will also analyze sampled-data feedback control systems and their stability. Modified z-transforms will also be presented, because they provide precise representation of systems with dead time.

We assume in this chapter that the signals are processed at a uniform interval of time, the *sample time.* For many common digital control systems, the sample time is of the order of 1 s or less. This sampling frequency is so fast, compared with the response of most processes, that their control systems can be treated as though they were continuous without significant loss of accuracy. On the other hand, many advanced control techniques are best carried out on computers at sampling intervals of between 0.1 and 20 min. At such slow sampling rates, continuous analysis is not appropriate. Analyzer controllers also require sampled-data design techniques, because their cycles are of the order of 5 to 20 min-too slow to ignore for most processes. The techniques we will

learn in this chapter are essential for designing and analyzing sampled-data control systems.

14-1 COMPUTER PROCESS CONTROL

Any computer, from the smallest hand-held personal computer to the most sophisticated parallel processor, can be used for process control. All that is required is the proper interface to the process instrumentation and an operating system that can perform operations at regular intervals of time, based on the computer's internal clock. Computers and microprocessors that are specifically designed for process control operate on a basic processing frequency known as their heartbeat, usually 1 s. Different tasks are then scheduled at a set number of heartbeats. For example, flow and pressure control loops may be executed every heartbeat, level control loops every 5 heartbeats, some temperature control loops every 15 heartbeats, and other temperature loops every 30 to 60 heartbeats. Analyzer controllers are usually set to operate whenever an analysis cycle is completed and the measured composition is available, but even then, the period between executions of the control tasks is relatively constant. The point is that computers execute control functions at regular intervals of time that we will call the *sample time.*

A typical computer process control system is sketched in Fig. 14-1.1. The process interface consists of two parts, the analog-to-digital (A/D) converter and the **digital-to-**analog (D/A) converter. The A/D converter converts the continuous process signals (e.g., **mA,** psig, **mV**) into digital numbers that are stored in the computer memory. These represent the inputs to the control system. The process signals are continuously available for sampling whenever the computer needs to know what they are. The D/A converter converts the computer outputs into electronic signals, in volts or **mA,** and holds them constant until their next update. These signals normally go to **current-to-**pressure (I/P) transducers and position control valves. Figure 14- 1.2 shows the continuous input and output signals; the sample values are marked on the input signal.

A typical computer control calculation includes the following sequence, executed at each sample instant.

1. The computer requests a value from the A/D converter. The A/D converter samples the process signal, converts it to a number, and stores it in the computer memory or a register.

2. The computer performs the control calculations on the sampled process signal(s) and computes the output(s) to the process.
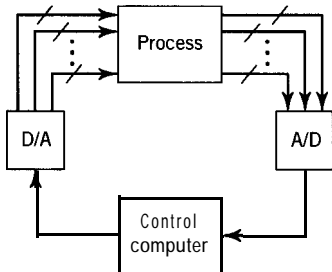
Figure   14-1.1   Computer   feedback control   system.
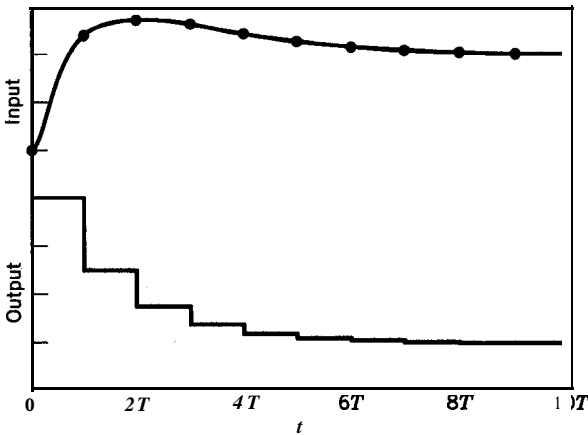
Figure 14-1.2 Continuous input signal, showing the sampled values, and computer output signal.

3. The computer output is sent to the D/A converter, which converts it to an electronic signal, updates the output, and holds it constant until the next update.

With current computer technology, all these operations can take place in a single **panel**-mounted instrument or in a complete network involving several processors and the transfer of signals through a data highway or field bus.

Note the order of the operations. The process signal is sampled first, and the output to the process is updated last. The time required for the computer to carry out step 2 is usually very small compared with the sample time, so the updating of the output signal happens almost simultaneously with the sampling of the input signal, but never before. This means that even if the process responded instantaneously to the action of the computer, the computer would have to wait an entire sample before it could receive feedback on its action. In other words, *the process signal is always a minimum of one sample behind the computer output update.*

Having shown that the operation of computers in process control is performed on sampled signals at discrete intervals of time, we will next see how the z-transform can represent operations on sampled signals.

14-2 THE z-TRANSFORM

This section presents the definition and properties of the z-transform and describes two methods for calculating its inverse: partial fractions expansion and long division.

14-2.1 Definition of the z-Transform

Consider the continuous function of Fig. **14-2.1,** which is sampled at uniform intervals of time, *T.* The z-transform of the function is defined by

$$\mathcal{Z}[f(t)] = \sum_{n=0}^{\infty} f(nT)z^{-n}$$

$$= f(0) + f(T)z^{-1} + f(2T)z^{-2} + \cdots \tag{14-2.1}$$
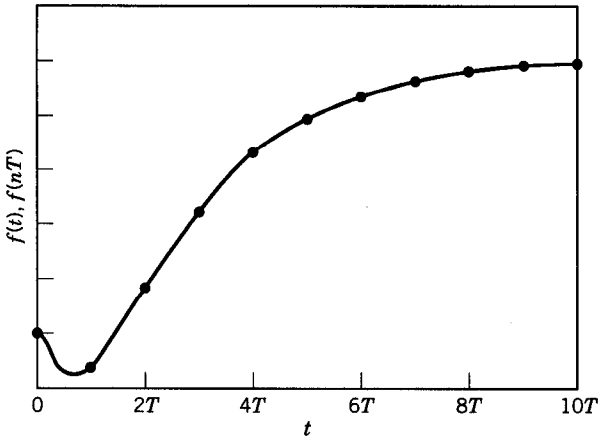
Figure **14-2.1** Continuous function sampled at uniform intervals of time, T.

where

z = transform variable
T = sample interval or sample time
$f(nT)$ = value of the function at $t = nT$

From the definition, we see that the z-transform of a function contains only the values at the sampling instants and that these values depend on the sample time, T. Also, the transform does not contain information on the function for negative time. The following notation is more convenient to use for the z-transform of a function:

$$F(z) = \sum_{n=0}^{\infty} f(nT)z^{-n} \qquad \textbf{(14-2.2)}$$

where $F(z)$ = z-transform of $f(t)$ sampled at intervals of time, T.
    We will next use Eq. 14-2.2 to derive the z-transform of some common functions.

EXAMPLE 14-2.1

Derive the z-transform of (a) a unit step function, (b) a decaying exponential, and (c) a cosine wave.

SOLUTION

(a) Unit step function:

$$f(t) = u(t)$$

(a)
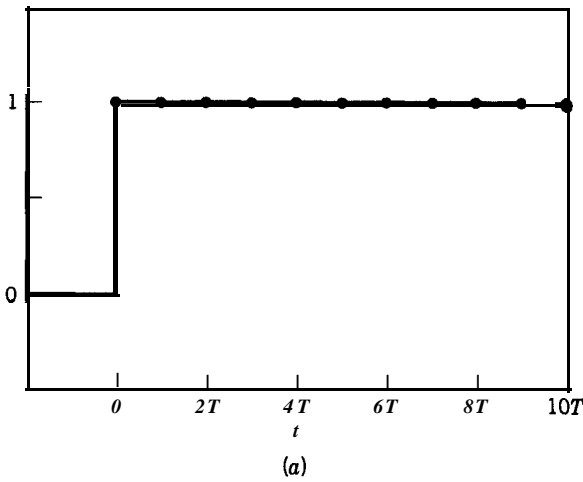
Figure 14-2.2 Sampled functions for Example 14-2.1.
(a) Unit step function, u(r).

The sampled function is sketched in Fig. **14-2.2a.** Applying Eq. 14-2.2 yields

$$F(z) = \sum_{n=0}^{\infty} u(nT)z^{-n} = 1 + z^{-1} + z^{-2} + \cdots$$

We note that the resulting function is an infinite series. You can easily verify that $(1 - z^{-1})F(z) = 1$ and thus

$$\boxed{F(z) = \frac{1}{1 - z^{-1}}}$$

$$(14\text{-}2.3)$$

(b)
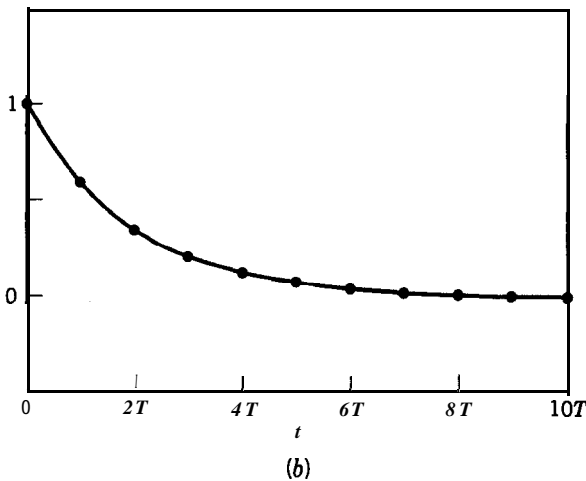
Figure 14-2.2 (Continued) (b) Decaying exponential.

(c)
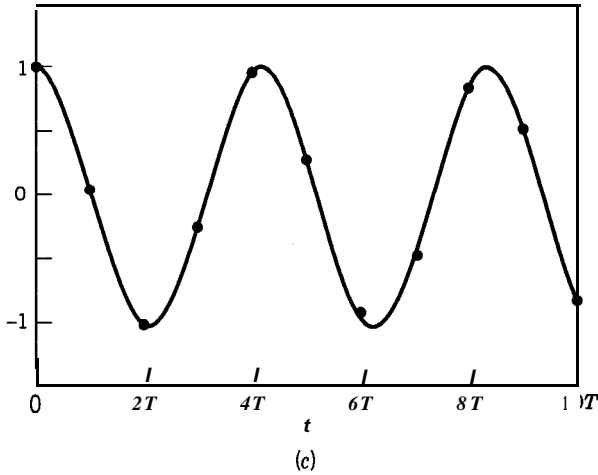
Figure 14-2.2 *(Continued) (c)* Cosine wave.

(b) Decaying exponential:

$$f(t) = e^{-t/\tau}$$

The sampled function is sketched in Fig. 14-2.2*b*. From Eq. 14-2.2,

$$F(z) = \sum_{n=0}^{\infty} e^{-nT/\tau}z^{-n} = 1 + e^{-T/\tau}z^{-1} + e^{-2T/\tau}z^{-2} + \cdots$$

Again we can again verify that $(1 - e^{-T/\tau}z^{-1})F(z) = 1$, so

$$F(z) = \frac{1}{1 - e^{-T/\tau}z^{-1}} \tag{14-2.4}$$

(c) Cosine wave:

$$f(t) = \cos \omega t$$

The sampled function is sketched in Fig. 14-2.2*c*. From Eq. 14-2.2,

$$F(z) = \sum_{n=0}^{\infty} \cos n\omega T z^{-n}$$

First, substitute the exponential form of the cosine.

$$\cos n\omega T = \frac{1}{2}(e^{in\omega T} + e^{-in\omega T})$$

where $i = \sqrt{-1}$. Substitute to obtain

$$F(z) = \sum_{n=0}^{\infty} \frac{1}{2} (e^{in\omega T} + e^{-in\omega T})z^{-n}$$

$$= \frac{1}{2} \left(\sum_{n=0}^{\infty} e^{in\omega T}z^{-n} + \sum_{n=0}^{\infty} e^{-in\omega T}z^{-n} \right)$$

Now, as in part (b), the sums can be expressed as

$$F(z) = \frac{1}{2} \left(\frac{1}{1 - e^{i\omega T}z^{-1}} + \frac{1}{1 - e^{-i\omega T}z^{-1}} \right)$$

$$= \frac{1}{2} \left[\frac{2 - (e^{i\omega T} + e^{-i\omega T})z^{-1}}{1 - (e^{i\omega T} + e^{-i\omega T})z^{-1} + z^{-2}} \right]$$

Substitute the cosine function and simplify, to obtain

$$\boxed{F(z) = \frac{1 - z^{-1}\cos \omega T}{1 - 2z^{-1}\cos \omega T + z^{-2}}} \qquad (14\text{-}2.5)$$

We will now use the z-transform of the cosine function to show that more than one function can share the same z-transform if they coincide at the sampling instants. For example, all cosine waves with frequencies

$$\omega = \frac{2k\pi}{T}, \ k = 1, 2, 3, \ . \ . \ .$$

have the z-transform

$$F(z) = \frac{1 - z^{-1}\cos 2k\pi}{1 - 2z^{-1}\cos 2k\pi + z^{-2}} = \frac{1 - z^{-1}}{1 - 2z^{-1} + z^{-2}} = \frac{1 - z^{-1}}{(1 - z^{-1})^2} = \frac{1}{1 - z^{-1}}$$

This is also the z-transform derived in part (a) of this example for a unit step function. Figure 14-2.3 shows that these two functions coincide at the sampling instants.

Table 14-2.1 contains the z-transforms of some common functions, along with the time functions and their **Laplace** transforms.

14-2.2   Relationship to the **Laplace** Transform

We will now establish the relationship between the z-transform and the **Laplace** transform by taking the **Laplace** transform of the sampled function. Figure 14-2.4 shows a diagram of the sampler with a continuous input **function** $f(t)$. The output of the sampler
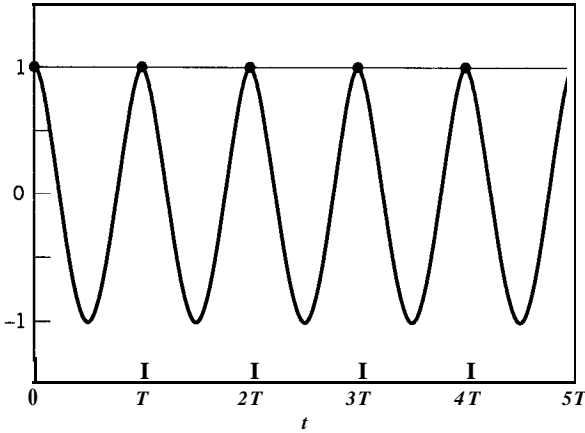
Figure 14-2.3 The cosine function sampled once per cycle
coincides at the sampling instants with the unit step function.

is the sampled function $f^*(t)$. The Laplace transform of the sampled function is, from Eq. 2-1.1,

$$F^*(s) = \mathscr{L}[f^*(t)] = \int_0^\infty f^*(t)e^{-st}\, dt \qquad (14\text{-}2.6)$$

where $s =$ Laplace transform variable.

To evaluate the integral in Eq. 14-2.6, we must determine the nature of the sampled function. In practice, the sampled function is zero except for a negligibly small but finite period of time around the sampling instant. During this time, the sampler output settles to the value of the continuous input function. For an exact mathematical function such as the Laplace transform to be valid, we must assume that the sampler acts instantaneously, but this would make the integral identically zero because the sampler
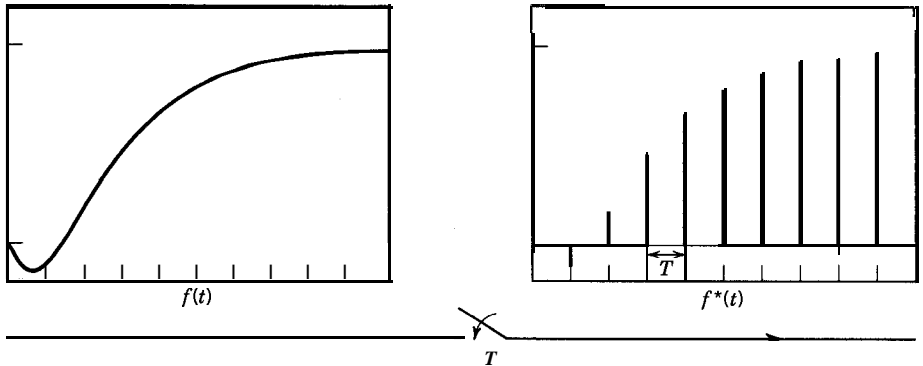
Figure 14-2.4 A sampler and the sampled function with sample time T.

Table 14-2.1 z-Transforms and Modified z-Transforms

$F(t)$ Time Function	$F(s)$ Laplace Transform	$F(z)$ z-Transform	$F(z, m)$ Modified z-Transform
$u(t)$	$\dfrac{1}{s}$	$\dfrac{1}{1 - z^{-1}}$	$\dfrac{z^{-1}}{1 - z^{-1}}$
t	$\dfrac{1}{s^2}$	$\dfrac{Tz^{-1}}{(1 - z^{-1})^2}$	$\dfrac{mTz^{-1}}{1 - z^{-1}} + \dfrac{Tz^{-2}}{(1 - z^{-1})^2}$
t^n	$\dfrac{n!}{s^{n+1}}$	$\lim_{a \to 0} (-1)^n \dfrac{\partial^n}{\partial a^n} \dfrac{1}{1 - e^{-aT}z^{-1}}$	$\lim_{a \to 0} (-1)^n \dfrac{\partial^n}{\partial a^n} \left(\dfrac{e^{-amT}z^{-1}}{1 - e^{-aT}z^{-1}} \right)$
$e^{-t/\tau}$	$\dfrac{\tau}{\tau s + 1}$	$\dfrac{1}{1 - e^{-T/\tau}z^{-1}}$	$\dfrac{e^{-mT/\tau}z^{-1}}{1 - e^{-T/\tau}z^{-1}}$
$u(t)(1 - e^{-t/\tau})$	$\dfrac{1}{s(\tau s + 1)}$	$\dfrac{(1 - e^{-T/\tau})z^{-1}}{(1 - z^{-1})(1 - e^{-T/\tau}z^{-1})}$	$z^{-1} \left(\dfrac{1}{1 - z^{-1}} - \dfrac{e^{-mT/\tau}}{1 - e^{-T/\tau}z^{-1}} \right)$
$\sin \omega t$	$\dfrac{\omega}{s^2 + \omega^2}$	$\dfrac{z^{-1}\sin \omega T}{1 - 2z^{-1}\cos \omega T + z^{-2}}$	$z^{-1} \left(\dfrac{\sin m\omega T + z^{-1}\sin(1 - m)\omega T}{1 - 2z^{-1}\cos \omega T + z^{-2}} \right)$
$\cos \omega t$	$\dfrac{s}{s^2 + \omega^2}$	$\dfrac{1 - z^{-1}\cos \omega T}{1 - 2z^{-1}\cos \omega T + z^{-2}}$	$z^{-1} \left(\dfrac{\cos m\omega T - z^{-1}\cos(1 - m)\omega T}{1 - 2z^{-1}\cos \omega T + z^{-2}} \right)$
$e^{-at}\sin \omega T$	$\dfrac{\omega}{(s + a)^2 + \omega^2}$	$\dfrac{z^{-1}e^{-aT}\sin \omega T}{1 - 2z^{-1}e^{-aT}\cos \omega T + e^{-2aT}z^{-2}}$	$z^{-1}e^{-amT} \left(\dfrac{\sin m\omega T + z^{-1}e^{-aT}\sin(1 - m)\omega T}{1 - 2z^{-1}e^{-aT}\cos \omega T + e^{-2aT}z^{-2}} \right)$
$e^{-aT}\cos \omega T$	$\dfrac{s + a}{(s + a)^2 + \omega^2}$	$\dfrac{1 - z^{-1}e^{-aT}\cos \omega T}{1 - 2z^{-1}e^{-aT}\cos \omega T + e^{-2aT}z^{-2}}$	$z^{-1}e^{-amT} \left(\dfrac{\cos m\omega T - z^{-1}e^{-aT}\cos(1 - m)\omega T}{1 - 2z^{-1}e^{-aT}\cos \omega T + e^{-2aT}z^{-2}} \right)$

output pulses would have zero duration and therefore zero area. How do we get around this problem? By the definition of the ***ideal sampler.***

Ideal   Impulse   Sampler. An ideal sampler outputs not a series of pulses of finite magnitude and zero duration but rather a series of impulses of infinite magnitude, zero duration, and area equal to the magnitude of the continuous function at the sampling instants. The sampled function is then the following sum of impulse functions:

$$f^*(t) = \sum_{n=0}^{\infty} f(nT)\, \delta(t - nT) \qquad\qquad (14\text{-}2.7)$$

where $\delta(t - nT)$ is the unit impulse (unit area) occurring at $t = nT$. Substitute the ideal sampled function, Eq. 14-2.7, into the **Laplace** transform of the sampled function, Eq. 14-2.6. to obtain

$$F^*(s) = \int_0^{\infty} \sum_{n=0}^{\infty} f(nT)\, \delta(t - nT)\, e^{-st}\, dt$$

$$= \sum_{n=0}^{\infty} f(nT) \int_0^{\infty} \delta(t - nT)\, e^{-st}\, dt$$

The impulse function in each integral is zero except at the sampling instant $t = nT$, where the area of the impulse is concentrated. Each integral in the sum is then

$$\int_0^{\infty} \delta(t - nT)\, e^{-st}\, dt = e^{-nTs}$$

Substitute into the summation to obtain

$$F^*(s) = \sum_{n=0}^{\infty} f(nT) e^{-nTs} \qquad\qquad (14\text{-}2.8)$$

When we compare Eq. 14-2.8 with Eq. 14-2.2, the definition

$$\boxed{z = e^{Ts}} \qquad\qquad (14\text{-}2.9)$$

makes the z-transform equal to the **Laplace** transform of the sampled function:

$$F(z) = F^*(s)$$

This result is significant, because it shows that the z-transform is a special case of the **Laplace** transform.

14-2.3   Properties of the z-Transform

Linearity. The z-transform is linear:

$$\mathcal{L}[af(t) + bg(t)] = a\mathcal{L}[f(t)] + b\mathcal{L}[g(t)] \qquad\qquad \textbf{(14-2.10)}$$

where a and b are constants. This property follows directly from the definition of the z-transform, Eq. 14-2.1.

Time Delay (Real Translation Theorem). This theorem applies to the transform of a function delayed in time by an integer multiple of sample times.

$$\mathcal{L}[f(t - kT)] = z^{-k}F(z) \qquad\qquad \textbf{(14-2.11)}$$

where k is an integer. The theorem holds only **if $f(t) = 0$** for $t < 0$, because F(z) is defined only for positive values of t. To prove it, apply Eq. 14-2.1 to the delayed function.

$$\mathcal{L}[f(t - kT)] = \sum_{n=0}^{\infty} f(nT - kT)z^{-n}$$

Substitute j = n $-$ k.

$$\mathcal{L}[f(t - kT)] = \sum_{j=-k}^{\infty} f(jT)z^{-j-k}$$

q.e.d.

$$= z^{-k}\sum_{j=0}^{\infty} f(jT)z^{-j} = z^{-k}F(z)$$

where we assumed **that** $f(jT) = 0$ for j < 0.

Final Value Theorem. This theorem lets us calculate the final or steady-state value of a function, if it exists, from the z-transform.

$$\lim_{n\to\infty} f(nT) = \lim_{z\to1}(1 - z^{-1})F(z) \qquad\qquad \textbf{(14-2.12)}$$

The proof of this theorem adds little to our understanding of it. The following three theorems are also presented without proof.

Complex Translation Theorem. This theorem is useful for obtaining z-transforms of functions that contain exponential functions of time.

$$\mathcal{L}[e^{-at}f(t)] = F(z_1)|_{z_1 = ze^{aT}} \qquad\qquad \textbf{(14-2.13)}$$

***Initial Value* Theorem.** This theorem enables us to calculate the initial value of a function from its z-transform.

$$\lim_{t \to 0} f(t) = \lim_{z \to \infty} F(z) \qquad (14\text{-}2.14)$$

The proof follows directly from inspection of the definition of the z-transform, Eq. 14-2.2.

Partial   Differentiation   Theorem. This theorem is also useful for deriving z-transforms from known ones.

$$\mathcal{L}\left[\frac{\partial}{\partial a} f(t, a)\right] = \frac{\partial}{\partial a} F(z, a) \qquad (14\text{-}2.15)$$

where a is a parameter of the function.

EXAMPLE 14-2.2

Use the linearity property and the real translation theorem to derive the z-transform of the following function, sampled at intervals $T = 0.75$.

$$f(t) = u(t - 1.5)[1 - e^{-(t-1.5)}]$$

SOLUTION

The unit step function at $t = 1.5$ shows that the function is zero for $t < 1.5$. Thus the function can be treated as a function delayed by 1.5 units of times.

$$f(t) = g(t - 1.5)$$

where

$$g(t) = u(t)[1 - e^{-t}]$$

Using the linearity property of z-transforms yields

$$G(z) = \mathcal{L}[u(t) - u(t)e^{-t}]$$

$$= \frac{1}{1 - z^{-1}} - \frac{1}{1 - e^{-T}z^{-1}} = \frac{(1 - e^{-T})z^{-1}}{(1 - z^{-1})(1 - e^{-T}z^{-1})}$$

where we have used Eqs. 14-2.3 and 14-2.4. Now, making use of the real translation theorem, Eq. 14-2.11, we find that

$$F(z) = \mathcal{Z}[g(t - 2T)] = z^{-2}G(z) = \frac{(1 - e^{-T})z^{-3}}{(1 - z^{-1})(1 - e^{-T}z^{-1})}$$

where, for $T = 0.75$, the delay of 1.5 units of time is a delay of $1.5/0.75 = 2$ samples. Substitute $T = 0.75$ to obtain

$$F(z) = \frac{0.528z^{-3}}{(1 - z^{-1})(1 - 0.472z^{-1})}$$

This numerical result shows that the z-transform of a function does not contain the value of the sampling interval in it.

EXAMPLE 14-2.3

Use the complex translation theorem to obtain the z-transform of

$$f(t) = e^{-at} \cos wt$$

SOLUTION

From the complex translation theorem, Eq. 14-2.13,

$$\mathcal{Z}[e^{-at} \cos wt] = \mathcal{Z}_1[\cos wt]_{z_1 = ze^{at}}$$

Substitute z_1 into the z-transform of the cosine function, Eq. 14-2.5, to obtain the desired transform.

$$\mathcal{Z}[e^{-at} \cos wt] = \frac{1 - z^{-1}e^{-aT} \cos wT}{1 - 2z^{-1}e^{-aT} \cos wT + z^{-2}e^{-2aT}} \qquad \textbf{(14-2.16)}$$

This result agrees with the entry in Table 14-2.1.

EXAMPLE 14-2.4

Verify the initial value theorem and the final value theorem for the function

$$f(t) = 1 - 0.8e^{-t/2} \cos 2.5t$$

sampled once a second $(T = 1)$.

SOLUTION

First obtain the z-transform by using the linearity property and Eqs. 14-2.3 and 14-2.16.

$$F(z) = \frac{1}{1 - z^{-1}} - 0.8 \frac{1 - z^{-1}e^{-1/2} \cos 2.5}{1 - 2z^{-1}e^{-1/2} \cos 2.5 + z^{-2}e^{-1}}$$

$$= \frac{1}{1 - z^{-1}} - \frac{0.8(1 + 0.486z^{-1})}{1 + 0.972z^{-1} + 0.368z^{-2}}$$

$$= \frac{0.2 + 1.383z^{-1} + 0.757z^{-2}}{(1 - z^{-1})(1 + 0.972z^{-1} + 0.368z^{-2})}$$

Then apply the initial value theorem.

$$\lim_{z \to \infty} F(z) = \frac{0.2}{(1 - 0)(1)} = 0.2$$

This checks with $f(0) = 1 - 0.8(1)(1) = 0.2$.
    Next apply the final value theorem:

$$\lim_{z \to 1} (1 - z^{-1})F(z) = \lim_{z \to 1} \frac{0.2 + 1.383z^{-1} + 0.757z^{-2}}{1 + 0.972z^{-1} + 0.368z^{-2}}$$

$$= \frac{0.2 + 1.383 + 0.757}{1 + 0.972 + 0.368} = \frac{2.340}{2.340} = 1.000$$

This checks with the final value of the original function.

$$\lim_{n \to \infty} f(nT) = 1 - 0.8e^{-\infty} \cos cc = 1.0$$

It is always good practice to check the accuracy of a z-transform by applying the initial and final value theorems before investing time and effort in any further manipulation.

EXAMPLE 14-2.5

Use the partial differentiation theorem to obtain the z-transform of the function

$$f(t) = t \sin \omega t$$

SOLUTION

We first note that

$$\frac{\partial}{\partial \omega} (\cos \omega t) = - t \sin \omega t$$

Then, applying Eq. 14-2.15, the partial differentiation theorem, yields

$$\mathfrak{Z}[t \sin \omega t] = \mathfrak{Z}\left(-\frac{\partial}{\partial \omega} \cos \omega t\right) = -\frac{\partial}{\partial \omega}\mathfrak{Z}[\cos \omega t]$$

Substituting Eq. 14-2.5, the z-transform of the cosine function, yields

$$
\begin{aligned}
F(z) &= -\frac{\partial}{\partial \omega}\left(\frac{1 - z^{-1}\cos \omega T}{1 - 2z^{-1}\cos \omega T + z^{-2}}\right)\\
&= -\frac{z^{-1}T \sin \omega T}{1 - 2z^{-1}\cos \omega T + z^{-2}} + \frac{(1 - z^{-1}\cos \omega T)(2z^{-1}T \sin \omega T)}{(1 - 2z^{-1}\cos \omega T + z^{-2})^2}\\
&= \frac{z^{-1}T \sin \omega T(1 - z^{-2})}{(1 - 2z^{-1}\cos \omega T + z^{-2})^2}
\end{aligned}
\tag{14-2.17}
$$

where we differentiated the quotient of two polynomials, combined fractions, and simplified.

14-2.4 Calculation of the Inverse z-Transform

Inverting the z-transform consists of finding the values of the sampled function from its z-transform. Because the z-transform contains only information about the function at the sampling instants, it is not possible to find the continuous function from the z-transfonn. We can, however, find a continuous function that coincides with the sampled function at the sampling instants. When the sample time is not known, the resulting function will be in terms of dimensionless time t/T.

There are several ways to calculate the inverse z-transform. We will study only the two most popular ones: partial fractions expansion and long division. The former results in a closed form of the sampled function, and the latter results in the numerical sampled values of the function.

Partial Fractions Expansion

This is the same as the method of inverting Laplace transforms presented in Chapter 2. It requires a table of z-transforms such as Table 14-2.1. The following example demonstrates the procedure.

EXAMPLE 14-2.6

Given the transform

$$F(z) = \frac{0.5z^{-1}}{1 - 1.60z^{-1} + 0.60z^{-2}}$$

find the sampled values of the function it represents.

SOLUTION

First we must find the roots of the denominator. This is easier to do if we first multiply by z^2 to get positive powers of z.

$$z^2 - 1.60z + 0.60 = 0$$

For this second-order system, we use the quadratic formula to find the roots at -1 and -0.60. Higher-order systems require a polynomial root-finding program. With the roots, we factor the denominator, back in negative powers of z.

$$(1 - 1.60z^{-1} + 0.60z^{-2}) = (1 - z^{-1})(1 - 0.60z^{-1})$$

Expand in partial fractions.

$$F(z) = \frac{0.5z^{-1}}{(1 - z^{-1})(1 - 0.60z^{-1})} = \frac{A_1}{1 - z^{-1}} + \frac{A_2}{1 - 0.60z^{-1}}$$

The coefficients A, and A, are determined by the standard procedure we learned in Chapter 2.

$$A, = \frac{0.5z^{-1}}{1 - 0.60z^{-1}} \bigg|_{z=1} = \frac{0.5}{1 - 0.60} = \frac{0.5}{0.40} = 1.25$$

$$A^2 = \frac{0.5z^{-1}}{1 - z^{-1}} \bigg|_{z=0.60} = \frac{\frac{0.5}{0.60}}{1 - \frac{1}{0.60}} = \frac{0.5}{-0.40} = -125 \quad .$$

Substitute into the expanded transfer function to obtain

$$F(z) = \frac{1.25}{1 - z^{-1}} - \frac{1.25}{1 - 0.60z^{-1}}$$

Comparing with the entries in Table 14-2.1, the first term is a step function, u(t), of magnitude 1.25. The second term is of the form

$$\frac{b}{1 - e^{-aT}z^{-1}}$$

where $e^{-aT} = 0.60$ and $b = 1.25$. The inverse is then

$$f(nT) = 1.25u(nT) - 1.25e^{-anT}$$
$$= 1.25[u(nT) \quad (e^{-aT})^n]$$
$$= 1.25[u(nT) - 0.60^n]$$

This result can now be used to compute the sampled values.

| $f(nT)$ | 0 | 0.50 1 | 0.80 2 | 0.98 3 | 1.088 4 | • • • | 1.25 ∞ |

Long Division

This method requires that the known z-transform be expressed as the ratio of two polynomials in negative powers of z.

$$F(z) = \frac{a, + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_n z^{-n}}{b_0 + b_1 z^{-1} + b_2 z^{-2} + \ldots + b_n z^{-n}}$$

The result of the polynomial division is the quotient polynomial.

$$F(z) = c_0 + c_1 z^{-1} + c_2 z^{-2} + \ldots$$

Comparing this polynomial with the definition of the z-transform, Eq. 14-2.1,

$$F(z) = f(0) + f(T)z^{-1} + f(2T)z^{-2} + \cdots$$

we conclude that

$$f(0) = c_0, \qquad f(T) = c_1, \qquad f(2T) = c_2, \qquad \text{and so on}$$

Thus the sampled values of the function are the coefficients of the quotient polynomial.

EXAMPLE 14-2.7

Let us obtain the inverse of the z-transform given in Example 14-2.6 by long division.

SOLUTION

It is easier to work with negative powers of z.

$$F(z) = \frac{0.5z^{-1}}{1 - 1.60z^{-1} + 0.60z^{-2}}$$

The long-division procedure is the division of two polynomials.

$$\begin{array}{r} 0.50z^{-1} +\ 0.80z^{-2} +\ 0.98z^{-3} +\ 1.088z^{-4} + \cdots \\ \hline 1 - 1.60z^{-1} + 0.60z^{-2}\ \overline{)0.50z^{-1}} \\ \underline{0.50z^{-1} - 0.80z^{-2} + 0.30z^{-3}} \\ 0.80z^{-2} - 0.30z^{-3} \\ \underline{0.80z^{-2} - 1.28z^{-3} + 0.480z^{-4}} \\ 0.98z^{-3} - 0.480z^{-4} \\ \underline{0.98z^{-3} - 1.568z^{-4} + 0.588z^{-5}} \\ 1.088z^{-4} - 0.588z^{-5} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \end{array}$$

The quotient polynomial is

$$F(z) = 0.50z^{-1} + 0.80z^{-2} + 0.98z^{-3} + 1.088z^{-4} + \ldots$$

Comparing this polynomial to Eq. 14-2.1, we get

n	0	1	2	3	4		∞
$f(nT)$	00.	0.50	0.80	0.98	1.088	...	1.25

where the final value theorem has been applied.

$$f(\infty) = \lim_{z\to1} (1 - z^{-1}) \frac{0.50z^{-1}}{1 - 1.60z^{-1} + 0.60z^{-2}} = \frac{0}{0}$$

Applying L'Hôpital's rule yields

$$f(\infty) = \lim_{z\to1} \frac{-0.5z^{-2} + 1.0z^{-3}}{1.60z^{-2} - 1.20z^{-3}} = \frac{0.50}{0.40} = 1.25$$

Note that partial fractions expansion should be used when a compact form of the sampled function is desired, whereas long division is a more direct way to obtain the numerical values of the sampled function. Computer programs such as **MATLAB** (1995) can be used to carry out the long division.

14-3 PULSE TRANSFER FUNCTIONS

This section presents the development of transfer functions to relate the sampled input and output signals of dynamic linear systems. Because sampled signals are basically pulses, **the** transfer functions that relate them *are* known as *pulse transfer functions.* Obviously, the z-transform is the natural method for developing pulse transfer functions.

14-3.1   Development of the Pulse Transfer Function

Consider a continuous process with a single input and a single output, as shown in the block diagram of Fig. 14-3.1. As we learned in Chapter 3, the **Laplace** transform of the output signal is

$$C(s) = G(s)M(s) \qquad\qquad (14\text{-}3.1)$$

Figure 14-3.1 Continuous block, $C(s) = G(s)M(s)$.

where

$C(s)$ = **Laplace** transform of output signal
$M(s)$ = **Laplace** transform of input signal
$G(s)$ = transfer function

Impulse Response

Suppose a unit impulse is applied to the system and its response is recorded against time. Recall from Chapter 2 that a unit impulse is a pulse of zero duration, infinite amplitude, and unity area. Its **Laplace** transform, from Table 14-2.1, is unity.

$$M(s) = 1$$

Thus, when the input is an impulse function, the output is

$$C(s) = G(s)$$

That is, the **Laplace** transform of the response to a unit impulse is equal to the system transfer function. The impulse response of the system is then the time function that coincides with the output response to the unit impulse.

$$G(t) = \mathcal{L}^{-1}[G(s)] \tag{14-3.2}$$

Note that the impulse response must be zero for negative time, because a dynamic system cannot respond to the impulse before it takes place.

Pulse Transfer Function

Because a continuous system responds to its input signal at all times, the development of a transfer function that relates the sampled values requires that the inputs to the continuous system be zero except at the sampling instants. That is, both the input and output signals must be sampled, as in Fig. 14-3.2. In the diagram, M*(s) and C*(s) are the **Laplace** transforms of the sampled input and output signals, respectively. Assuming
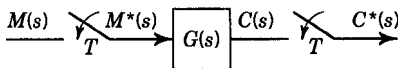
Figure 14-3.2 Continuous block with sampled input, M*(s), and output, $C^*(s)$.

the samplers are ideal, their output signals are trains of impulses, one at each sampling instant. The signal from the input sampler can be represented by

$$M''(t) = M(O) \, S(t) + M(T) \, \delta(t - T) + \ldots + M(kT) \, \delta(t - kT) + \ldots$$

Assuming the system is linear, or is a linear approximation at the operating conditions, the principle of superposition applies. Thus the continuous output of the system is the sum of the responses to all of the input impulses.

$$C(t) = M(0)G(t) + M(T)G(t - T) + M(2T)G(t - 2T)$$
$$+ \ldots + M(kT)G(t - kT) + \ldots$$
$$= \sum_{k=0}^{\infty} M(kT)G(t - kT)$$

From Eq. 14-2.2, the z-transform of the output signal is

$$C(z) = \sum_{n=0}^{\infty} C(nT)z^{-n}$$

Substitution of $C(nT)$ from the above expression gives

$$C(z) = \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} M(kT)G(nT - kT)z^{-n}$$

Let $j = n - k$.

$$C(z) = \sum_{k=0}^{\infty} \sum_{j=-k}^{\infty} M(kT)G(jT)z^{-(j+k)}$$

Because $G(jT)$ is zero for negative j, we can move the lower limit on j to zero and separate the sums to obtain

$$C(z) = \left(\sum_{j=0}^{\infty} G(jT)z^{-j} \right) \left(\sum_{k=0}^{\infty} M(kT)z^{-k} \right)$$

Applying the definition of the z-transform, Eq. 14-2.2 yields

$$C(z) = G(z)M(z) \tag{14-3.3}$$

where

$$G(z) = \sum_{j=0}^{\infty} G(jT)z^{-j} \tag{14-3.4}$$

G(z) is called the impulse transfer function, or simply the *pulse transfer function,* of the system. Comparing Eq. 14-3.3 to Eq. 14-3.1, we note that G(z) relates the sampled

input and output signals in the same way G(s) relates the continuous signals. It is understood that the samplers shown in Fig. 14-3.2 operate synchronously and at the same frequency.

EXAMPLE 14-3.1

Develop the pulse transfer function of (a) an integrator and (b) a first-order lag.

SOLUTION

(a) Integrator:

$$G(s) = \frac{1}{s}$$

From Table 14-2.1, the impulse response is

$$G(t) = \mathcal{L}^{-1}\frac{1}{0s} = u(t) \qquad \text{(unit step)}$$

$$G(z) = \sum_{n=0}^{\infty} u(nT)z^{-n} = 1 + z^{-1} + z^{-2} + \cdots$$

As before, this infinite series reduces to

$$G(z) = \frac{1}{1 - z^{-1}} \qquad (14\text{-}3.5)$$

(b) First-order lag:

$$G(s) = \frac{K}{\tau s + 1} = \frac{K}{\tau}\frac{\tau}{\tau s + 1}$$

From Table 14-2.1, the impulse response is

$$G(t) \overset{=}{\text{ }} \frac{K}{\tau}\mathcal{L}^{-1}\left[\frac{\tau}{\tau s + 1}\right] = \frac{K}{\tau}e^{-t/\tau}$$

Take the z-transform.

$$G(z) = \frac{K}{\tau}\sum_{n=0}^{\infty} e^{-nT/\tau}z^{-n}$$

This infinite series reduces to

$$G(z) = \frac{\dfrac{K}{\tau}}{1 - e^{-T/\tau}z^{-1}} \tag{14-3.6}$$

We could have obtained the same results by looking up, in Table 14-2.1, the entry corresponding to the Laplace transform of the continuous transfer function.

Pulse transfer functions can be manipulated algebraically to analyze sampled-data control systems, provided that each pulse transfer function is defined between sampled (discrete) signals. For example, for the diagram of Fig. 14-3.3a,

$$C(z) \neq G_1(z)G_2(z)M(z)$$

This is because the second block responds to the values of X between sampling instants. Such intersample contributions are not accounted for by the pulse transfer function G,(z).

In the other hand, if a sampler is inserted between the blocks, as in Fig. 14-3.3b, we can write:

$$C(z) = G_1(z)G_2(z)M(z)$$

In this case, the sampler between the blocks interrupts the continuous signal, X(s), between sampling instants, so G,(z) does represent the effect of this sampled signal on the output.

14-3.2   Steady-State Gain of a Pulse Transfer Function

The steady-state gain of a pulse transfer function can be obtained by application of the final value theorem of z-transforms, following a procedure parallel to the one we used in Chapter 3 for continuous transfer functions. Apply the final value theorem, Eq.

$$M(s) \xrightarrow{\quad} \underset{T}{\times} \xrightarrow{M(z)} \boxed{G_1(s)} \xrightarrow{X(s)} \boxed{G_2(s)} \xrightarrow{C(s)} \underset{T}{\times} \xrightarrow{C(z)}$$

(a)

$$M(s) \xrightarrow{\quad} \underset{T}{\times} \xrightarrow{M(z)} \boxed{G_1(s)} \xrightarrow{X(s)} \underset{T}{\times} \xrightarrow{X(z)} \boxed{G_2(s)} \xrightarrow{C(s)} \underset{T}{\times} \xrightarrow{C(z)}$$

(b)
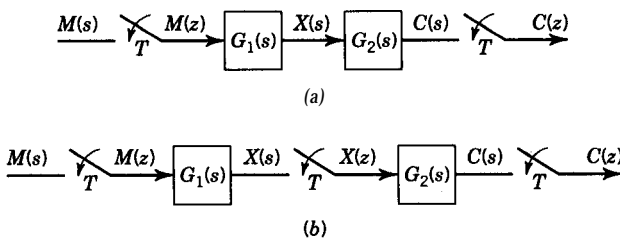
Figure 14-3.3 Pulse transfer functions require that both the input and the output signals be sampled. (a) With continuous intermediate signal, $C(z) \neq G_1(z)G_2(z)M(z)$. (b) With intermediate sampler, $C(z) = G_1(z)G_2(z)M(z)$.

14-2.12, to the output signal C(z), substitute Eq. 14-3.3, and take limits.

$$\lim_{n \to \infty} C(nT) = \lim_{z \to 1} (1 - z^{-1})C(z)$$
$$= \lim_{z \to 1} (1 - z^{-1})G(z)M(z)$$
$$= \left[\lim_{z \to 1} G(z) \right] \left[\lim_{z \to 1} (1 - z^{-1})M(z) \right]$$
$$= \left[\lim_{z \to 1} G(z) \right] \lim_{n \to \infty} M(nT)$$

Recalling that both $C(nT)$ and $M(nT)$ are deviation variables, we know the term in brackets is the gain, because it is the number that, multiplied by the steady-state change in the input M, results in the steady-state change in the output, C.

$$\boxed{\text{Gain} = \lim_{z \to 1} G(z)} \qquad \qquad \textbf{(14-3.7)}$$

Thus, to obtain the steady-state gain of a pulse transfer function, we simply set z = 1.

14-3.3 Pulse Transfer Functions of Continuous Systems

Computer process control systems, as we have noted earlier, perform their calculations at regular intervals of time. The calculated control action goes out to the process, which operates continuously, and process measurements are sampled by the computer. Pulse transfer functions are an excellent tool for representing the response of the sampled process output signals to the discretely calculated computer output signals. However, for the pulse transfer functions to represent the process properly, they must include the fact that the signals from the computer are held constant between computer updates. To represent this holding of the output signal, we use a mathematical device known as the *zero-order hold.*

To understand better the need for the hold device, consider Eq. 14-3.6. This equation was developed in Example 14-3.1, supposedly to represent a continuous first-order lag. But does it? For one thing, the steady-state gain of the pulse transfer function, Eq. 14-3.6, is, applying Eq. 14-3.7,

$$\text{Gain} = \lim_{z \to 1} \frac{\dfrac{K}{\tau}}{1 - e^{-T/\tau}z^{-1}} = \frac{\dfrac{K}{\tau}}{1 - e^{-T/\tau}}$$

This does not match the gain of the continuous transfer function, which is K, except perhaps at one specific value of the sample time. A look at the step responses of the pulse and continuous transfer functions, shown in Fig. 14-3.4, shows the reason for the discrepancy: the pulse transfer function responds to a train of impulses, not to a continuous step. To obtain the correct pulse transfer function, we must insert a zero-order hold, H(s), between the sampler and the continuous block, as shown in Fig. 14-3.5a.
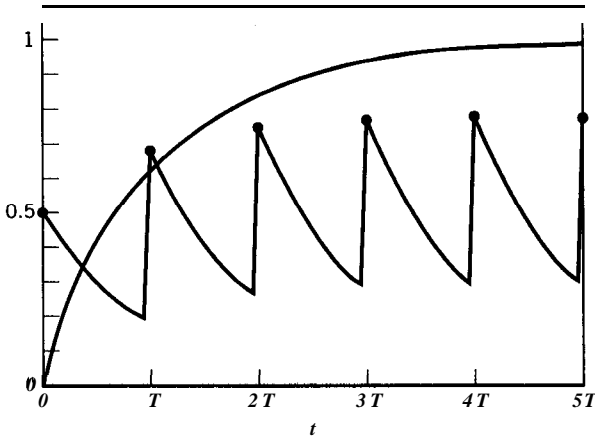
Figure 14-3.4 Step responses of continuous and pulse transfer functions of first-order lag. The pulse transfer function responds to a series of impulses, not a sustained step.

A *zero-order hold* simply holds the sampled value of the input between samples, without accumulation. Its impulse response is shown in Fig. 12-3.56 and is given by

$$H(t) = u(t) - u(t - T)$$

because it holds the area of the impulse (sampled value) for one sampling interval, T. The **Laplace** transform of the zero-order hold is then, from Table 14-2.1 and the real translation theorem of **Laplace** transforms,

$$H(s) = \frac{1}{s} - \frac{1}{s} e^{-Ts} = \frac{1 - e^{-Ts}}{s} \qquad (14\text{-}3.8)$$

Because the output of the hold is usually a continuous signal, we would not normally be interested in its pulse transfer function *per* se. However, it is essential that the transfer function of the hold device be included in the development of the pulse transfer function
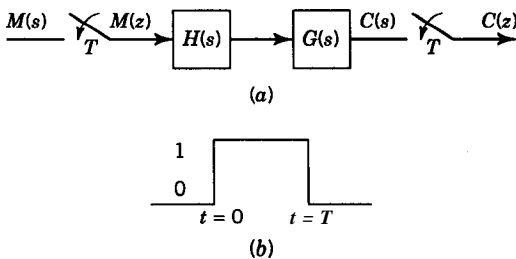
Figure 14-3.5 (a) Continuous block with zero-order hold, H(s). (b) Impulse response of zero-order hold.

of the complete system. The impulse response of the combination of the zero-order hold and a continuous transfer function is

$$C(t) = \mathcal{L}^{-1}[H(s)G(s)] = \mathcal{L}^{-1}\left[\frac{1 - e^{-Ts}}{s}G(s)\right]$$

$$= \mathcal{L}^{-1}\left[\frac{G(s)}{s}\right] - \mathcal{L}^{-1}\left[e^{-Ts}\frac{G(s)}{s}\right]$$

The second term is equal to the first term delayed by one sample time, T. Taking the z-transform yields

$$\mathcal{Z}\{\mathcal{L}^{-1}[H(s)G(s)]\} = \mathcal{Z}\left\{\mathcal{L}^{-1}\left[\frac{G(s)}{s}\right]\right\} - z^{-1}\mathcal{Z}\left\{\mathcal{L}^{-1}\left[\frac{G(s)}{s}\right]\right\}$$

$$HG(z) = (1 - z^{-1})\mathcal{Z}\left[\frac{G(s)}{s}\right] \qquad\qquad (14\text{-}3.9)$$

In this last step, we use the notation $HG(z)$ to indicate the pulse transfer function of the combination of the hold device and the process. We also dispensed with the symbol for the inverse **Laplace** transform; it is not necessary to perform the indicated inverse transform when using a table of z-transforms such as Table 14-2.1.

EXAMPLE 14-3.2

A process can be represented by a first-order lag. Its input is held by a zero-order hold. Develop the pulse transfer function of the combination.

SOLUTION

The process transfer function is

$$G(s) = \frac{K}{\tau s + 1}$$

From Eq. 14-3.9,

$$HG(z) = (1 - z^{-1})\mathcal{Z}\left[\frac{K}{s(\tau s + 1)}\right]$$

From Table 14-2.1,

$$HG(z) = (1 - z^{-1}) \frac{K(1 - e^{-T/\tau})z^{-1}}{(1 - z^{-1})(1 - e^{-T/\tau}z^{-1})}$$

$$HG(z) = \frac{K(1 - e^{-T/\tau})z^{-1}}{1 - e^{-T/\tau}z^{-1}} \qquad (14\text{-}3.10)$$

If we compare this result with the pulse transfer function of the first-order lag without the zero-order hold from Example 14-3.1, Eq. 14-3.6, we note the difference the hold device makes. Figure 14-3.6 shows the step response of the first-order lag with zero-order hold. It exactly matches the response of the continuous transfer function at the sampling instants.

The unit step response is developed as follows. From Eq. 14-3.3, the z-transform of the output is

$$C(z) = HG(z)M(z) = \frac{K(1 - a)z^{-1}}{1 - az^{-1}} M(z)$$

where we substituted Eq. 14-3.10 and, to simplify notation, let a = $e^{-T/\tau}$. Now we let the input be a unit step function, Eq. 14-2.3 and expand in partial fractions.

$$C(z) = \frac{K(1 - a)}{1 - az^{-1}} \frac{1}{1 - z^{-1}} = \frac{K}{1 - z^{-1}} - \frac{K}{1 - az^{-1}}$$
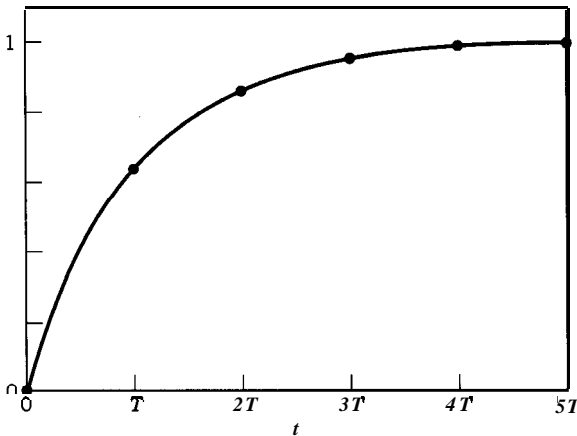
Figure 14-3.6 The step response of sampled first-order lag with zero-order hold coincides with the response of the continuous block at the sampling instants.

We invert, using Table 14-2.1, after substituting a $= e^{-T/\tau}$.

$$C(nT) = K[1 - e^{-nT/\tau}]$$

This is the response plotted in Fig. 14-3.6.

14-3.4 Transfer Functions of Discrete Blocks

As we have said, computer-based control calculations are performed at discrete instants of time. We refer to these computing devices as "discrete" blocks. Because it is not possible to define continuous transfer functions for discrete blocks, a special procedure must be devised to obtain their transfer functions.

Consider the single-input-single-output discrete block of Fig. 14-3.7. The samplers are shown for clarification, and the notation $D(z)$ indicates that the block is discrete and acts only on the sampled values. A typical computing formula recursively calculates the output signal from current and past values of the input signal and past values of the output signal. This formula is a linear difference equation of the form

$$Y_n = a_0 X_n + a_1 X_{n-1} + \cdots + a_j X_{n-j} + b_1 Y_{n-1} + b_2 Y_{n-2} + \ldots + b_k Y_{n-k} \quad (14\text{-}3.11)$$

where

$Y_n = Y(nT)$, the output at the nth sampling instant
$X_n = X(nT)$, the input at the nth sampling instant

and a and b are constant coefficients. From Eq. 14-2.2, the z-transform of the output signal is

$$Y(z) = \sum_{n=0}^{\infty} Y_n z^{-n} = Y_0 + Y_1 z^{-1} + Y_2 z^{-2} + \ldots$$

Substitute Eq. 14-3.11 for each term to obtain

$$Y(z) = \sum_{n=0}^{\infty} (a_0 X_n + a_1 X_{n-1} + \ldots + a_j X_{n-j} + b_1 Y_{n-1} + \ldots + b_k Y_{n-k}) z^{-n}$$

$$= a_0 \sum_{n=0}^{\infty} X_n z^{-n} + a_1 z^{-1} \sum_{n=0}^{\infty} X_n z^{-n} + \ldots + a_j z^{-j} \sum_{n=0}^{\infty} X_n z^{-n}$$

$$+ b_1 z^{-1} \sum_{n=0}^{\infty} Y_n z^{-n} + \ldots + b_k z^{-k} \sum_{n=0}^{\infty} Y_n z^{-n}$$
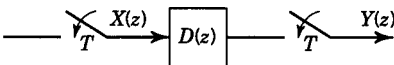
Figure 14-3.7 Block diagram of a discrete (computing) block: $Y(z) = D(z)X(z)$.

where the terms for negative n have been assumed zero. We now substitute the appropriate z-transform for each summation term.

$$Y(z) = a_0X(z) + a_1z^{-1}X(z) + \ldots + a_jz^{-j}X(z) + b_1z^{-1}Y(z) + \ldots + b_kz^{-k}Y(z)$$

Solving for Y(z) yields

$$Y(z) = \frac{a_0 + a_1z^{-1} + \cdots + a_jz^{-j}}{1 - b_1z^{-1} - \cdots - b_kz^{-k}} X(z)$$

Thus the transfer function of the discrete block is

$$D(z) = \frac{Y(z)}{X(z)} = \frac{a_0 + a_1z^{-1} + \ldots + a_jz^{-j}}{1 - b_1z^{-1} - \cdots - b_kz^{-k}}$$ (14-3.12)

Note that we can obtain this equation directly from Eq. 14-3.11 if we use the z-transform variable as the shift operator.

$$X_{n-1} = z^{-1}X_n \qquad Y_{n-1} = z^{-1}Y_n$$

Then the numerator terms are the coefficients of the input signal X, and the denominator terms are the coefficients of the output signal Y.

EXAMPLE 14-3.3   *TRANSFER FUNCTION OF EXPONENTIAL FILTER*

The formula for a digital exponential filter is

$$Y_n = (1 - a)X_n + aY_{n-1}$$ (14-3.13)

where Y_n is the filter output, X_n is the filter input, and a is the filter adjustable parameter in the range $0 \leq a \leq 1$. Determine the transfer function of the filter.

SOLUTION

Substitute the z-transform of the signals, using z as the shift operator.

$$Y(z) = (1 - a)X(z) + uz\text{-'}Y(z)$$

Solve for Y(z).

$$Y(z) = \frac{(1 - a)}{1 - az^{-1}}X(z)$$

The transfer function is

$$\mathbf{F(z)} = \frac{Y(z)}{X(z)} = \frac{1 - a}{1 - az^{-1}} \qquad \textbf{(14-3.14)}$$

Exponential filters are commonly used in digital control systems to attenuate noisy input signals such as those coming from flow measurements.

In the block diagram of the discrete block, Fig. 14-3.7, we assume, as before, that the samplers on the input and output signals are synchronous-that is, they sample at exactly the same instant. Because of the time it takes the computer to perform the calculations of the block, this is not exactly what happens in practice. However, as discussed at the beginning of this chapter, this computing time is usually a very small fraction of the sampling interval and can be neglected. Thus the samplers can be assumed to be synchronous.

An alternative to neglecting the computing time would be to have the samplers operate synchronously and to have the computing block use the previous sampled value of the input in the calculation. But doing so would introduce an unnecessary delay of one sample into the operation of the computer. This is why computers use the latest sampled values of the input signals in their calculations and update their outputs very shortly after the inputs are sampled, not one sampling interval later.

EXAMPLE 14-3.4

Determine the steady-state gains of the pulse transfer functions developed in Examples 14-3.2 and 14-3.3.

SOLUTION

First-order lag with zero-order hold, Eq. 14-3.10:

$$\lim_{z \to 1} HG(z) = \frac{K(1 - e^{-T/\tau})}{1 - e^{-T/\tau}} = K$$

Digital exponential filter, Eq. 14-3.14:

$$\lim_{z \to 1} F(z) = \frac{1 - a}{1 - a} = 1$$

14-3.5   Simulation of Continuous Systems with Discrete Blocks

To determine the recursive difference equation from the transfer function of a discrete block, we need only reverse the procedure we learned in the preceding section. We can

further extend the procedure to derive recursive formulas that simulate continuous systems by starting with the pulse transfer functions of the continuous systems. In doing this, it is essential that a *zero-order hold* be included in the determination of the pulse transfer function. The following example illustrates this procedure.

EXAMPLE 14-3.5

Determine recursive computing formulas to simulate (a) an integrator and (b) a **first**-order lag.

SOLUTION

(a) Integrator:

$$G(s) \; = \; \frac{1}{s}$$

From Eq. 14-3.9,

$$\frac{Y(z)}{X(z)} = HG(z) = (1 - z^{-1})\mathscr{Z}\left[\frac{1}{s^2}\right]$$

From Table 14-2.1,

$$\frac{Y(z)}{X(z)} = (1 - z^{-1}) \left(1 \frac{Tz^{-1}}{-z^{-1})^2}\right) = \frac{Tz^{-1}}{1 - z^{-1}} \tag{14-3.15}$$

Rearrange the equation.

$$(1 - z^{-1})Y(z) = Tz^{-1}X(z)$$

In the next step, the z variable is treated as a shift operator.

$$Y(z) - z^{-1}Y(z) = Tz^{-1}X(z)$$

$$Y_n - Y_{n-1} = TX_{n-1} \tag{14-3.16}$$

$$Y_n = Y_{n-1} + TX_{n-1}$$

This equation can be recognized as the recursive integration formula commonly used in computer programs.

(b) First-order lag:

$$G(s) = \frac{K}{\tau s + 1}$$

From Example 14-3.2,

$$\frac{Y(z)}{X(z)} = HG(z) = \frac{K(1 - e^{-T/\tau})z^{-1}}{1 - e^{-T/\tau}z^{-1}}$$

Rearrange and solve for Y(z) to obtain

$$Y(z) = K(1 - e^{-T/\tau})z^{-1}X(z) + e^{-T/\tau}z^{-1}Y(z)$$

Replace the z-transforms with the (shifted) sampled values.

$$Y_n = K(1 - e^{-T/\tau})X_{n-1} + e^{-T/\tau}Y_{n-1} \tag{14-3.17}$$

Note that for $K = 1$, $a = e^{-T/\tau}$, this formula is almost the same as the digital filter formula given in Eq. 14-3.13. What is the difference?

The recursive computing formula can be used to compute the sampled values of the response for any sequence of sampled values of the input. The calculations are best performed on a spreadsheet, because it is relatively easy to program the recursive equation in the spreadsheet. For example, the unit step response of a first-order process can be computed from Eq. 14-3.17 as follows: Let $Y_0 = 0$, let $X_n = 1$ for $n \geq 0$, and let $a = e^{-T/\tau}$. Then

$$Y_1 = K(1 - u)(1) + aY_0$$

$$Y_2 = K(1 - u)(1) + aY_1$$

and so on. Similarly, the ramp response is obtained by setting $X_n = nT$ for $n \geq 0$.

In the next chapter, we will learn simpler methods for deriving recursive computing formulas from continuous transfer functions. However, the method introduced in this section results in the most accurate formulas for simulating continuous systems.

14-4 SAMPLED-DATA FEEDBACK CONTROL SYSTEMS

A sampled-data feedback control system is one in which some of the signals in the feedback loop are sampled. Sampling is required in some composition control loops, because the analyzer must take a sample of a stream and process it for analysis, such as chromatographic analysis of hydrocarbon mixtures. In general, a computer controller is by necessity a sampled-data feedback controller.

Figure 14-4.1 contains the block diagram for a typical sampled-data control loop. In this diagram the **Laplace** transform variable, s, indicates continuous signals and transfer functions, and the z-transform variable indicates sampled signals and discrete-block transfer functions. The discrete block labeled $D(z)$ represents the digital control computation, and the transfer functions $G_P(s)$ and $G_U(s)$ combine the transfer functions of all the field elements (valve, process, and sensor/transmitter). We will next use standard block diagram analysis to develop the **closed-loop transfer function** between the sampled values of the output, C(z), and the inputs, R(z) and U(s), which are the sampled set point and the continuous disturbance, respectively.

14-4.1 Closed-Loop Transfer Function

Applying the rules of block diagram algebra that we learned in Chapter 3, we obtain

$$C(s) = H(s)G_P(s)M^*(s) + G_U(s)U(s) \tag{14-4.1}$$

where $M^*(s)$ is the **Laplace** transform of the sampled output signal (see Section 14-2), and H(s) represents the zero-order hold. Next we take the z-transform and, from the definition of pulse transfer function, obtain

$$C(z) = \mathscr{L}[H(s)G_P(s)]M(z) + \mathscr{L}[G_U(s)U(s)]$$

or, in short notation,

$$C(z) = HG_P(z)M(z) + G_UU(z) \tag{14-4.2}$$

The term $G_UU(z)$ represents the z-transform of **the signal** from the continuous block, G,(s), caused by a continuous input to the block, U(s). A pulse transfer function cannot be obtained for **the** disturbance, because it is not a sampled signal, From further application of block diagram algebra around the loop, we obtain

$$M(z) = D(z)E(z) \tag{14-4.3}$$

$$E(z) = R(z) - C(z) \tag{14-4.4}$$
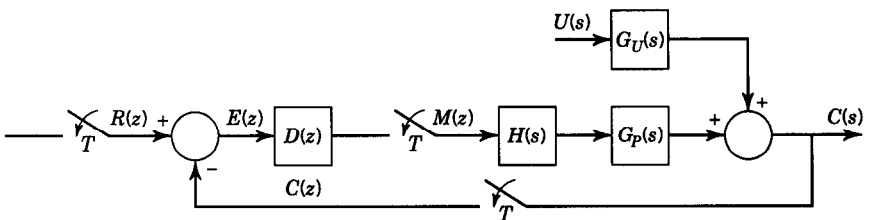
Figure **14-4.1**   Sampled-data   feedback   control   loop.

Combine the last three equations to eliminate M(z) and E(z).

$$C(z) = HG_P(z)D(z)[R(z) - C(z)] + G_U U(z)$$

Solve for the output C(z) to obtain

$$\boxed{C(z) = \frac{HG_P(z)D(z)}{1 + HG_P(z)D(z)}R(z) + \frac{G_U U(z)}{1 + HG_P(z)D(z)}} \qquad (14\text{-}4.5)$$

This shows the closed-loop transfer functions for the sampled-data feedback control loop. Note its similarity to the transfer function of a continuous unity-feedback control loop (developed in Chapter 6). Recall from there that the denominator term, common to the transfer functions for all inputs to the loop, determines the characteristic equation of the loop.

$$\boxed{1 + HG_P(z)D(z) = 0} \qquad (14\text{-}4.6)$$

The roots of this equation determine the unforced response of the closed loop and therefore its stability.

EXAMPLE 14-4.1

In the block diagram of Fig. 14-4.1, the process is first-order.

$$G_P(s) = \frac{K}{\tau s + 1}$$

For a proportional controller, $D(z) = K_c$, determine:

(a) The closed-loop transfer function, $C(z)/R(z)$
(b) The steady-state gain and the offset

SOLUTION

(a) In Example 14-3.2, we determined the pulse transfer function of a first-order process with zero-order hold, Eq. 14-3.10:

$$HG_P(z) = \frac{K(1 - e^{-T/\tau})z^{-1}}{1 - e^{-T/\tau}z^{-1}}$$

Substituting into Eq. 14-4.5 and excluding the disturbance term, we get

$$C(z) = \frac{\dfrac{K(1 - a)z^{-1}}{1 - az^{-1}}K_c}{1 + \dfrac{K(1 - a)z^{-1}}{1 - az^{-1}}K_c}R(z)$$

where a $= e^{-T/\tau}$. Simplify to obtain

$$\frac{C(z)}{R(z)} = \frac{KK_c(1-a)z^{-1}}{1 - az^{-1} + KK_c(1-a)z^{-1}} \tag{14-4.7}$$

(b) To obtain the steady-state gain, apply Eq. 14-3.7.

$$\lim_{z\to 1} \frac{C(z)}{R(z)} = \frac{KK_c(1-a)}{1 - a + KK_c(1-a)} \frac{KK_c}{1 + KK_c} \tag{14-4.8}$$

The offset is the steady-state error.

$$\text{Offset} = \lim_{n\to\infty} E(nT) = R(\infty) - C(\infty)$$

$$= R(\infty) - \frac{KK_c}{1 + KK_c} R(\infty) \tag{14-4.9}$$

$$\text{Offset} = \frac{1}{1 + KK_c} R(\infty)$$

This is exactly the same formula for the offset that we obtained in Chapter 6, Eq. 6-1.29. It shows that increasing the controller gain decreases the offset.

14-4.2   Stability of Sampled-Data Control Systems

Just like continuous systems, a discrete system is considered stable if, for a bound input, its output remains bound with time. Because most systems are open-loop stable, the stability conditions are important in analyzing feedback control loops. What is the condition for a sampled-data control system to be stable? To answer this question, let us consider the feedback control loop of Fig. 14-4.1.

The characteristic equation, Eq. 14-4.6, can be written as a polynomial of degree n in z. Thus there would be n values of z that satisfy the equation, counting repeated values. These are known as the **roots** of the characteristic equation:

$$1 + HG_P(z)D(z) = (1 - r_1 z^{-1})(1 - r_2 z^{-1}) \dots (1 - r_n z^{-1}) = 0 \tag{14-4.10}$$

where $r_i =$ root of the characteristic equation.

The denominator of the closed-loop transfer function, Eq. 14-4.5, can thus be factored in terms of these roots.

$$C(z) = \frac{b_1 + b_2 z^{-1} + b_3 z^{-2} + \dots + b_{j+1} z^{-j}}{(1 - r_1 z^{-1})(1 - r_2 z^{-1}) \dots (1 - r_n z^{-1})} \text{ (Input terms)} \tag{14-4.11}$$

where $b_i =$ numerator coefficients. Expanding in partial fractions yields

$$C(z) = \frac{A_1}{1 - r_1 z^{-1}} + \frac{A_2}{1 - r_2 z^{-1}} + \dots + \frac{A_n}{1 - r_n z^{-1}} + \text{ (Input terms)} \tag{14-4.12}$$

where A_i = constant partial fractions coefficients. The terms shown in this equation constitute the contribution of the loop transfer function to the output response, whereas the terms in parentheses represent the contribution of the set point and disturbance input signals. The stability of the loop is determined only by the transfer function terms. Inverting, with the help of Table 14-2.1, with $r_i = e^{-T/\tau_i}$, yields

$$C(kT) = A_1r_1^k + A_2r_2^k + \cdots + A_nr_n^k + \text{(Input terms)} \qquad \text{(14-4.13)}$$

If any of the roots r_i of the closed-loop characteristic equation is greater than unity in magnitude, then the corresponding term will go to infinity as k (time) increases indefinitely. This is so for **both** real and complex conjugate roots.

Figure 14-4.2 shows typical responses for the case of real roots. If the real root is positive and less than unity, then the term in the response decays monotonically to zero (Fig. 14-4.2a), but if the root is greater than unity, then the term increases at each sample and the response is unstable (Fig. 14-4.2b).

Ringing *Response.* If the real root is negative between zero and -1, then the response changes signs at each sample, decaying at each cycle (Fig. 14-4.2c). This is because the negative root will be raised alternately to even and to odd powers, resulting in positive and negative values, respectively. This oscillation, which has a period of exactly two samples and is known as *ringing,* is undesirable because it results in unnec-
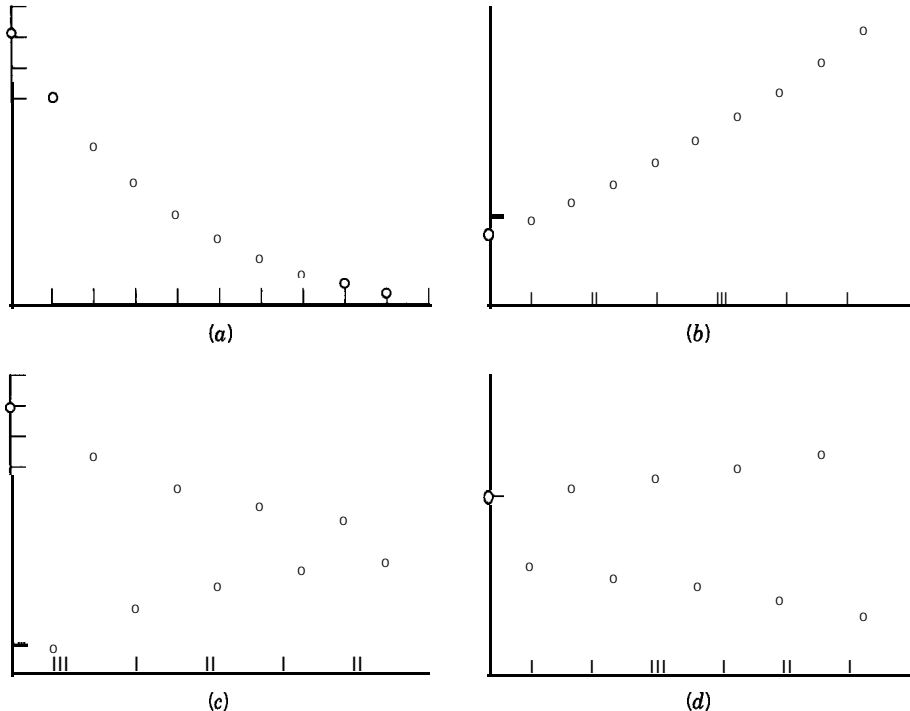
Figure 14-4.2 Responses of individual terms corresponding to real roots. (a) Stable monotonic response: $0 < r < 1$. (b) Unstable monotonic response: $1 < r$. (c) Stable ringing response, $1 < r < 0$. (d) Unstable ringing response: $r < -1$.

essary wear on the mechanical components of the control system, such as the control valve. If the real root is less than -1, then the response rings with increasing amplitude, as in Fig. 14-4.2d.

Finally, if the real root is unity in magnitude, then the response term remains constant if positive (step) or rings between $+1$ and -1 if negative.

Oscillatory Response. Figure 14-4.3 shows the stable and unstable oscillatory responses that result from complex conjugate roots:

$$r = \alpha + i\beta = |r|e^{i\theta}$$

where

α, β = real and imaginary parts, respectively
$|r| = \sqrt{\alpha^2 + \beta^2}$, magnitude or absolute value

$\theta = \tan^{-1} \dfrac{\beta}{\alpha}$, argument

$i = \sqrt{-1}$

The term in the response is

$$r^k = |r|^k e^{ik\theta} = |r|^k(\cos k\theta + i \sin k\theta)$$

$$\lim_{k \to \infty} |r|^k = \begin{cases} 0 & |r| < 1 \\ 1 & |r| = 1 \\ \infty & |r| > 1 \end{cases}$$

In other words, the oscillatory response term decays if the magnitude of the root is less than unity (Fig. 14-4.3a), maintains a constant amplitude if the magnitude is unity, and grows indefinitely if the magnitude is greater than unity (Fig. 14-4.3b).

From the preceding discussion, the criterion of stability is as follows: ***A sampled-data system is stable if and only if every one of the roots of its characteristic equation less than or equal to unity in magnitude.*** In other words, the system is unstable if

(a)   (b)
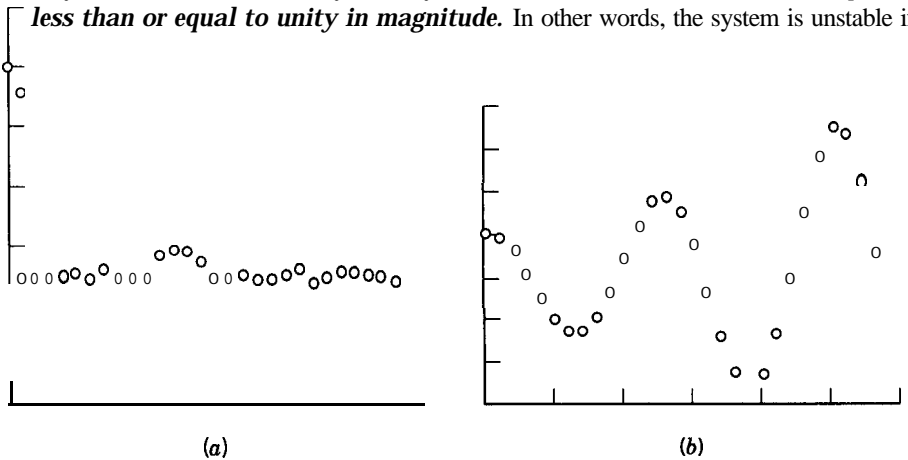
Figure 14-4.3 Responses of individual terms corresponding to complex roots. (a) Stable: $|r| < 1$. (b) Unstable: $|r| > 1$.
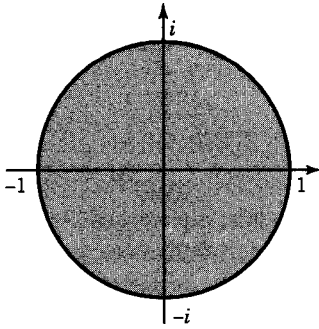
Figure 14-4.4 The stability re-
gion on the z-plane is the inside
of the unit circle.

any of the roots of the characteristic equation is greater than unity in *absolute value.*
The region of stability in the complex z-plane is the inside of the unit circle, as shown
in Fig. 14-4.4. It can easily be shown that Eq. 14-2.9 maps the left-hand side of the s-
plane into the inside of the unit circle in the z-plane.

The reader is encouraged to draw a parallel between the development of the stability
condition for sampled-data systems using z-transforms, and the analogous condition for
continuous systems using **Laplace** transforms (given in Chapter 2).

EXAMPLE 14-4.2

Determine the ultimate gain of the feedback control loop of Example 14-4.1. This is a
first-order process with a proportional controller.

SOLUTION

The characteristic equation is obtained by setting the denominator of the closed-loop
transfer function, Eq. 14-4.7, equal to zero.

$$1 - az^{-1} + KK_c(1 - a)z^{-1} = 0$$

where $a = e^{-T/\tau}$. This equation has a root at

$$r = a - KK_c(1 - a)$$

This root starts at a for $K_c = 0$ and becomes smaller as the controller gain increases.
As long as the root is positive, the response does not oscillate. This is so for loop
gains of

$$KK_c < \frac{a}{1 - a}$$

As the gain continues to increase, the response will ring with decreasing amplitude, as in Fig. 14-4.2c, until the root becomes − 1. This happens when the loop gain is

$$KK_{cu} = \frac{1 + a}{1 - a} = \frac{1 + e^{-T/\tau}}{1 - e^{-T/\tau}}$$

This is the ultimate gain. For gains greater than this, the response will ring with increasing amplitude, as in Fig. 14-4.2d. We can use this formula to see how the ultimate loop gain varies with the ratio of the sample time to the process time constant:

T/τ	10.0	1.0	0.1	0.01
KK_{cu}	1.0	2.16	20.0	200.0

Thus, as the sample time becomes smaller relative to the process time constant, the ultimate loop gain increases. When the sample time is much larger than the time constant, the ultimate loop gain is unity. Keep this result in mind when tuning flow and pressure loops that have very small time constants.

The preceding example shows that a first-order process with a sampled-data proportional controller can become oscillatory and unstable. Recall that we determined in Example 6-1.2 that a continuous proportional controller cannot make a first-order process unstable or oscillatory, no matter how high the gain. This result is also obtained in the preceding example by letting $T = 0$.

EXAMPLE 14-4.3

For the sampled-data feedback control system of Fig. 14-4.1, let R(z) = 0, $U(s)$ = l/s (unit step input), T = 1, a proportional-integral controller:

$$D(z) = K_c\left(1 + \frac{1}{5}\frac{1}{1 - z^{-1}}\right)$$

and first-order process transfer functions:

$$G_P(s) = \frac{0.9}{4s + 1} \qquad G_U(s) = \frac{1.5}{2s + 1}$$

Obtain the response of the closed loop and determine whether the system is stable for a control gain of (a) K_c = 2 %CO/%TO and (b) K_c = 9 %CO/%TO. Also plot the root locus on the z-plane.

SOLUTION

We first determine the necessary pulse transfer functions with the aid of a table of z-transforms, Table 14-2.1.

$$HG_P(z) = 0.9 \left. \frac{(1 - e^{-T/4})z^{-1}}{1 - e^{-T/4}z^{-1}} \right|_{T=1} = \frac{0.1991z^{-1}}{1 - 0.7788z^{-1}}$$

$$UG_U(z) = 1.5 \left. \frac{(1 - e^{-T/2})z^{-1}}{(1 - z^{-1})(1 - e^{-T/2}z^{-1})} \right|_{T=1} = \frac{0.5902z^{-1}}{(1 - z^{-1})(1 - 0.6065z^{-1})}$$

$$D(z) = K_c \left(\frac{1.2 - z^{-1}}{1 - z^{-1}} \right)$$

These we now combine into the closed-loop response, Eq. 14-4.5.

$$C(z) = \frac{\dfrac{0.5902z^{-1}}{(1 - z^{-1})(1 - 0.6065z^{-1})}}{1 + K_c \left(\dfrac{1.2 - z^{-1}}{1 - z^{-1}} \right) \left(\dfrac{0.1991z^{-1}}{1 - 0.7788z^{-1}} \right)}$$

$$= \left[\frac{0.5902z^{-1}(1 - 0.7788z^{-1})}{(1 - z^{-1})(1 - 0.7788z^{-1}) + 0.1991K_c z^{-1}(1.2 - z^{-1})} \right] \frac{1}{(1 - 0.6065z^{-1})}$$

The characteristic equation is

$$1 + K_c \left(\frac{1.2 - z^{-1}}{1 - z^{-1}} \right) \left(\frac{0.1991z^{-1}}{1 - 0.7788z^{-1}} \right) = 0$$

(a) For $K_c = 2$ %CO/%TO, the roots of the characteristic equation, obtained by the quadratic formula, are $r_1 = 0.8568$, $r_2 = 0.442$. **Thus** the system is **stable** for a controller gain of 2 %CO/%TO, because both roots are less than unity in magnitude.

(b) For $K_c = 9$ %CO/%TO, the roots of the characteristic equation are $r_1 = -1.2093$, $r_2 = 0.8378$. **Thus** the system is **unstable** for a control gain of 9 %CO/%TO.

Let us now compute the response for this second case by partial fractions expansion.

$$C(z) = \frac{0.5902z^{-1}(1 - 0.7788z^{-1})}{(1 + 1.209z^{-1})(1 - 0.8378z^{-1})(1 - 0.6065z^{-1})}$$

$$= \frac{-0.3156}{1 + 1.2093z^{-1}} + \frac{0.0735}{1 - 0.8378z^{-1}} + \frac{0.2421}{1 - 0.6065z^{-1}}$$

Inverting, with the help of Table 14-2.1, yields

$$C(kT) = -0.3156(-1.2093)^k + 0.0735(0.8378)^k + 0.2421(0.6065)^k$$
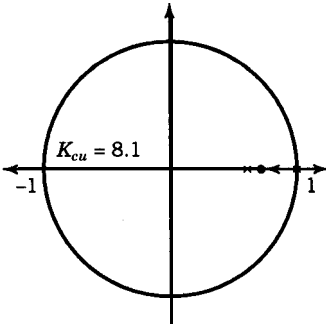
Figure 14-4.5 Root locus for Example 14-4.3.

Of the three terms shown, the last one is the contribution of the disturbance input, and the other two are contributed by the loop transfer function. The computed response is

k	C(kT)	k	C(kT)
0	0	5	0.8664
1	0.5901	6	− 0.9496
2	− 0.3209	7	1.2222
3	0.6554	8	− 1.4212
4	− 0.6060	9	1.7632

The output rings with increasing amplitude, as in Fig. $14\text{-}4.2d$. This is characteristic of a negative root of magnitude greater than unity.

To plot the root locus, we follow the procedure discussed in Chapter 8 for continuous control systems. There are two open-loop poles $(K_c = 0)$ at $z = 1$ and 0.7788, and there is an open-loop zero $(K_c = \infty)$ at $z = 1/1.2 = 0.8333$. Of the two loci, one ends at the zero, and the other follows the one asymptote at an angle of

$$\theta = \frac{180 \pm n360}{2 - 1} = 180°$$

The root locus is sketched in Fig. 14-4.5, where K_{cu} is the "ultimate" gain-that is, the gain at which the root locus crosses the unit circle. Knowing that it occurs at z = − 1, we can compute it by substituting this value into the characteristic equation.

$$K_{cu} = -\left[\frac{1 - (-1)}{1.2 - (-1)}\right]\left[\frac{1 - 0.7788(-1)}{0.1991(-1)}\right] = 8.1$$

14-5   MODIFIED z-TRANSFORM

So far, we have been able to develop pulse transfer functions using the regular z-transform by assuming that the process dead time is an integer multiple of the sample

time. To represent exactly a process in which the dead time includes a fraction of the sample time, we must **use** the **modified z-transform** to obtain the pulse transfer function. That technique is the subject of this section. Because modified z-transforms are somewhat more difficult to manipulate than regular z-transforms, it is often acceptable to avoid their use by assuming that the dead time is approximately an integer multiple of the sample time.

14-5.1 Definition and Properties of the Modified z-Transform

Consider a continuous function delayed by dead time t_0, where $t_0 = (N + \Delta)T$ (delay time) and $0 \leq A < 1$, a fraction. Figure 14-5.1 demonstrates the difficulty introduced by the fractional portion of the time delay: the sampled values of the delayed function are not the same as those of the original function. Thus the z-transform of the original function does not contain the same information as that of the delayed function. The **z**-transform of the delayed function is

$$\mathscr{Z}[f(t - t_0)] = \sum_{n=0}^{\infty} f(nT - NT - \Delta T)z^{-n}$$

Let us now define $m = 1 - A$ and $k = n - N - 1$. Substitute to obtain

$$\mathscr{Z}[f(t - t_0)] = \sum_{k=-N-1}^{\infty} f(kT + mT)z^{-k-N-1}$$

$$= z^{-N-1}\sum_{k=0}^{\infty} f(kT + mT)z^{-k} \qquad (14\text{-}5.1)$$
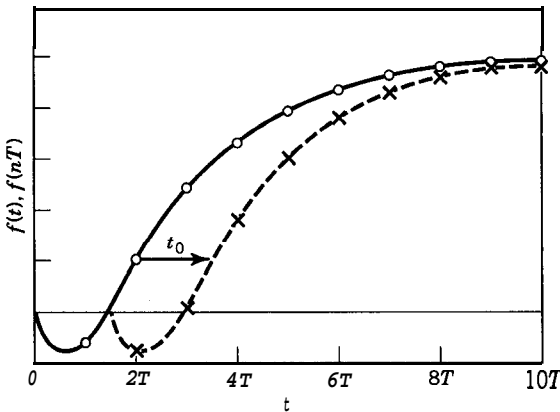
Figure 14-5.1 Continuous function delayed by a noninteger multiple of the sample time.

The modified z-transform of a continuous function of time is defined by

$$F(z, m) = \mathscr{L}_m[f(t)] = z^{-1} \sum_{n=0}^{\infty} f(nT + mT)z^{-n} \qquad \text{(14-5.2)}$$

where m is the modified z-transform variable and has the same meaning as in Eq. 14-5.1. The only difference between Eqs. 14-5.1 and 14-5.2 is the delay of N sample intervals, which is not included in the definition of the modified z-transform. We can consider Eq. 14-5.1 as the real translation theorem of the modified z-transform. The modified z-transforms of some common functions are listed in Table 14-2.1, along with the corresponding z-transforms. Unlike the z-transform, the modified transform contains information on the values of the function between samples, because any value may be obtained by varying **m** from just above zero to 1.0.

EXAMPLE 14-5.1

Derive the modified z-transform of (a) a decaying exponential and (b) a cosine wave.

SOLUTION

(a) A decaying exponential:

$$f(t) = e^{-t/\tau}$$

From Eq. 14-5.2,

$$F(z, m) = z^{-1} \sum_{n=0}^{\infty} e^{-(nT + mT)/\tau}z^{-n}$$

$$= z^{-1} \sum_{n=0}^{\infty} e^{-nT/\tau}e^{-mT/\tau}z^{-n}$$

$$= z^{-1}e^{-mT/\tau}(1 + e^{-T/\tau}z^{-1} + e^{-2T/\tau}z^{-2} + \cdots)$$

$$F(z, m) = \frac{e^{-mT/\tau}z^{-1}}{1 - e^{-T/\tau}z^{-1}} \qquad \text{(14-5.3)}$$

This result agrees with the entry in Table 14-2.1.

(b) A cosine wave:

$$f(t) = \cos \omega t$$

From Eq. 14-5.2,

$$F(z, m) = z^{-1} \sum_{n=0}^{\infty} \cos(\omega nT + \omega mT)z^{-n}$$

$$= z^{-1} \sum_{n=0}^{\infty} \frac{1}{2} (e^{i\omega nT}e^{i\omega mT} + e^{-i\omega nT}e^{-i\omega mT})z^{-n}$$

$$F(z, m) = z^{-1} \frac{1}{2} \left[\frac{e^{i\omega mT}}{1 - e^{i\omega T}z^{-1}} + \frac{e^{-i\omega mT}}{1 - e^{-i\omega T}z^{-1}} \right]$$

Combining fractions and simplifying, we get

$$F(z, m) = z^{-1} \frac{\cos \omega mT - z^{-1} \cos \omega(1 - m)T}{1 - 2z^{-1} \cos \omega T + z^{-2}} \qquad \textbf{(14-5.4)}$$

This result agrees with the entry in Table 14-2.1.

To show that Eq. 14-5.4 contains information on the values between samples, we
will now consider the case $\omega T = 2\pi$.

$$F(z, m) = z^{-1} \frac{\cos m2\pi - z^{-1} \cos(1 - m)2\pi}{(1 - z^{-1})^2}$$

and because m is a fraction, the values of the original cosine function are maintained.
We recall from Example 14-2.1 that the z-transform could not differentiate between the
$\cos \omega T$ and the unit step function when $\omega T = 2k\pi, k = 1, 2, 3, \ldots$.
The properties of the z-transform can be directly extended to the modified z-trans-
form, with the following exceptions:

Complex   Translation

$$\mathcal{L}_m[e^{-at}f(t)] = e^{aT(1-m)}F(ze^{aT}, m) \qquad \textbf{(14-5.5)}$$

Initial Valve Theorem

$$\lim_{t \to 0} f(t) = \lim_{\substack{m=0 \\ z \to \infty}} zF(z, m) \qquad \textbf{(14-5.6)}$$

EXAMPLE 14-5.2

Use the complex translation theorem to derive the modified z-transform of

$$f(t) = e^{-at} \cos \omega t$$

Apply Eq. 14-5.5 to Eq. 14-5.4.

$$F(z, m) = \frac{e^{aT(1-m)}z^{-1}e^{-aT}[\cos \omega mT - z^{-1}e^{-aT} \cos \omega(1-m)T]}{1 - 2z^{-1}e^{-aT} \cos \omega T + z^{-2}e^{-2aT}}$$

$$= \frac{e^{-amT}z^{-1}[\cos \omega mT - z^{-1}e^{-aT} \cos \omega(1-m)T]}{1 - 2z^{-1}e^{-aT} \cos \omega T + z^{-2}e^{-2aT}} \tag{14-5.7}$$

This result agrees with the entry in Table 14-2.1.

14-5.2   Inverse of the Modified z-Transform

Finding the inverse of a modified z-transform is done by the same methods as finding the inverse of the regular z-transform, but the calculations are somewhat lengthier because of the terms in m. The procedure will be demonstrated **with** an example.

EXAMPLE 14-5.3

Given the modified z-transform

$$F(z, m) = \frac{1.5z^{-1}[1 - 0.6z^{-1} - e^{-0.51m}(1 - z^{-1})]}{1 - 0.6z^{-1} - 0.4z^{-2}}$$

find the continuous function that corresponds to it.

SOLUTION

We shall calculate it by long division. A convenient approach is to separate the terms containing m.

$$F(z, m) = \frac{1.5z^{-1} - 0.9z^{-2}}{1 - 0.6z^{-1} - 0.4z^{-2}} - \frac{1.5z^{-1} - 1.5z^{-2}}{1 - 0.6z^{-1} - 0.4z^{-2}} e^{-0.51m}$$

We then perform each division separately to obtain

$$F(z, m) = (1.500z^{-1} + 0.000z^{-2} + 0.600z^{-3} + 0.360z^{-4} + \cdots)$$

$$- e^{-0.51m}(1.5z^{-1} - 0.600z^{-2} + 0.240z^{-3} - 0.096z^{-4} + \cdots)$$

Next we combine the two series to obtain

$$F(z, \text{m}) = (1.5 - 1.5e^{-0.51m})z^{-1}$$

$$+ (0.000 + 0.600e^{-0.51m})z^{-2}$$

$$+ (0.600 - 0.240e^{-0.51m})z^{-3}$$

$$+ (0.360 + 0.096e^{-0.51m})z^{-4} + \ldots$$

This solution is used to calculate the following response table.

m	z^{-1}	z^{-2}	z^{-3}	z^{-4}
0	0	0.600	0.360	0.456
0.25	0.180	0.528	0.389	0.445
0.50	0.338	0.465	0.414	0.434
0.75	0.477	0.409	0.436	0.426
1.00	0.600	0.360	0.456	0.418

The function is plotted in Fig. 14-5.2, where the crosses mark the points calculated in the table, and the circles are the sampled values, also calculated in the table. From Eq. 14-5.2, the coefficient of z^{-1} represents the function for $n = 0$, $f(mT)$, in the time period $0 < t < T$. Similarly, the coefficient of z^{-n} represents the function $f[(n - 1 + m)T]$ in the period $(n - 1)T < t < nT$, and so on.

14-5.3   Transfer Functions for Systems with Transportation Lag

Modified z-transforms offer us the means to obtain the pulse transfer function where the dead time is not an integer multiple of the sample time.

Let the process dead time be $t_0 = (N + A)T$, where N is an integer and A is a fraction, $0 \leq A < 1$. To obtain the pulse transfer function of a process with dead time, we handle
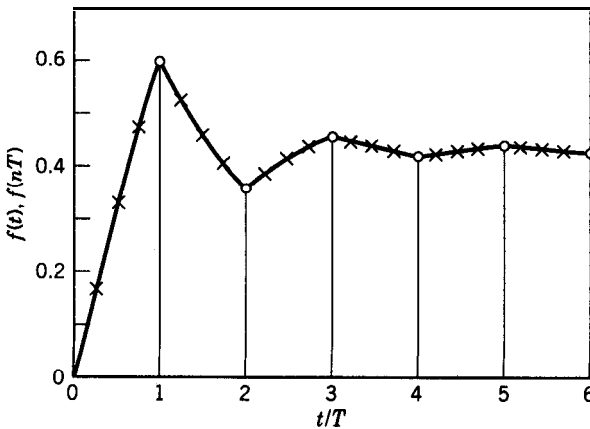
Figure 14-5.2 Response for Example 14-5.3.

the fraction A by using the modified z-transform with m = 1 ‒ A becoming just a constant. The following example illustrates the procedure.

EXAMPLE 14-5.4   *PULSE TRANSFER FUNCTION OF* **FIRST-ORDER-PLUS-DEAD-TIME (FOPDT) PROCESS**

Consider the general first-order-plus-dead-time transfer function

$$G_P(s) = \frac{Ke^{-t_0 s}}{\tau s + 1}$$

Determine the pulse transfer function $HG_P(z)$. Then evaluate it for the case **K = 1.2 %TO/%CO,** $t_0 = 1.95$ min, **$\tau = 5$** min, **T** = 1.5 min.

SOLUTION

From Equation (14-3.9),

$$HG_P(z) = (1 - z^{-1})\mathcal{L}\left[\frac{Ke^{-t_0 s}}{s(\tau s + 1)}\right]$$

Next, the dead time is expressed in terms of the sample time.

$$t_0 = (N + A)T$$

where N is an integer and A is a fraction, $0 \leq A < 1$. Substitute and apply the modified z-transform with m = 1 ‒ A.

$$HG_P(z) = K(1 - z^{-1})\mathcal{L}\left[\frac{e^{-NTs}e^{-\Delta Ts}}{s(\tau s + 1)}\right]$$

$$= K(1 - z^{-1})z^{-N}\mathcal{L}_m\left[\frac{1}{s(\tau s + 1)}\right]_{m=1-\Delta}$$

From the modified z-transform table, Table 14-2.1, after simplification,

$$HG_P(z) = Kz^{-N-1}\left[\frac{(1 - e^{-mT/\tau}) + (e^{-mT/\tau} - e^{-T/\tau})z^{-1}}{1 - e^{-T/\tau}z^{-1}}\right]_{m=1-A} \tag{14-5.8}$$

For the numerical values in the problem statement,

$$\frac{t}{T} = \frac{1.95}{1.50} = 1.3 \qquad \text{or} \qquad N = 1, \quad A = 0.3, \quad m = 1 - 0.3 = 0.7$$

Substitute into Eq. 14-5.8, along with $K = 1.2$ %TO/%CO, $\tau = 5$ min, and $T = 1.5$ min.

$$HG_P(z) = \frac{z^{-2}(0.2273 + 0.0837z^{-1})}{1 - 0.7408z^{-1}}$$

It is always a good idea to check that the steady-state gain of a pulse transfer function equals the steady-state gain of the original transfer function. This is to detect any computational or algebraic error before proceeding with further manipulation of the results. The steady-state gain of the pulse transfer function is

$$\lim_{z \to 1} HG_P(z) = \frac{0.2273 + 0.0837}{1 - 0.7408} = \frac{0.3110}{0.2592} = 1.20$$

This matches the steady-state gain of the continuous transfer function.

Modified z-transforms can also be used to determine the response between samples of the continuous blocks in a sampled-data dynamic system, but the information obtained is seldom worth the effort.

14-6 SUMMARY

This chapter presented the mathematical tools required to analyze sampled-data control systems: z-transforms, pulse transfer functions, and modified z-transforms. These techniques facilitate the design of computer control systems and analyzer controllers. We learned to develop the z-transform of a sampled signal and the pulse transfer functions of continuous systems and computer control blocks. The next chapter applies these concepts to the development and tuning of computer control algorithms.

For a more detailed presentation of z-transforms, consult a current text on sampled-data control systems, such as Deshpande (1988).

REFERENCES

1. Bristol, E. H. 1977. "Designing and Programming Control Algorithms for DDC Systems." *Control Engineering,* January: 24-26.
2. Deshpande, P. B. 1988. *Computer Process Control with Advanced Control Applications.* 2nd ed. Research Triangle Park, N.C.: Instrument Society of America.
3. MATLAB. 1995. Natick, Mass: The MathWorks, Incorporated.

PROBLEMS

14-1. Using the definition of the z-transforms, find F(z) for each of the following functions.
   (a) $f(t) = u(t)(1 - e^{-t/\tau})$
   (b) $f(t) = \sin \omega t$

(c) $f(t) = e^{-at} \sin \omega t$

where a, τ and ω are constants. Check your results against the entries in Table 14-2.1.

14-2. Use the properties of the z-transform to derive the z-transforms of the following functions.

(a) $f(t) = te^{-2t}$

(b) $f(t) = 1 - 0.7e^{-t/5} - 0.3e^{-t/8}$

(c) f(t) = $u(t - 2T)e^{-(t-2T)}(t - 2T)$

(d) $f(t) = e^{-t/5} \sin \pi t/T$

where T is the sample time. In each case, check your results against the entries in Table 14-2.1 and verify them using the initial and final value theorems.

14-3. Use the partial differentiation theorem to derive the z-transform of $f(t) = t^n$. Hint: Start with the function e^{-at} and take limits as $a \to 0$.

14-4. Obtain the inverse sampled function, $f(nt)$, in closed form, of the following z-transforms.

(a) $F(z) = \dfrac{1.2z^{-1}}{1 \quad 1.2z^{-1} + 0.2z^{-2}}$

(b) $F(z) = \dfrac{1.2z^{-1}}{1 - 2z^{-1} + z\text{-}z}$

(c) $F(z) = \dfrac{1.2z^{-1}}{1 - 2z^{-1} + 1.64z^{-2} - 0.64z^{-3}}$

Use your resulting functions to **evaluate** $f(nt)$ and sketch a graph off versus n.

14-5. Find the first four sampled values of each of the transforms of Problem 14-4 by long division.

14-6. Compute, by long division, the first four sampled values of the functions represented by the following z-transforms.

(a) $F(z) = \dfrac{0.5z^{-1}}{1 - 0.5z^{-1}}$

(b) $F(z) = \dfrac{0.5z^{-1}}{1 - 1.5z^{-1}}$

(c) $F(z) = \dfrac{0.5z^{-1}}{1 + 0.5z^{-1}}$

(d) $F(z) = \dfrac{0.5z^{-1}}{1 + 1.5z^{-1}}$

(e) $F(z) = \dfrac{0.5z^{-1}}{(1 - 0.5z^{-1})^2}$

(f) $F(z) = \dfrac{0.5z^{-1}}{1 - 0.5z^{-3}}$

(g) $F(z) = \dfrac{z\text{-}(1 + z^{-1} + z^{-2})}{1 + z^{-1} + z^{-2} + z^{-3} + z^{-4}}$

Draw a sketch of $f(nt)$ versus n for each function. Classify each function as stable or unstable, monotonic or oscillatory, and ringing or non-ringing.

14-7. A system is represented by the block diagram of Fig. **14-3.5a.** Develop the pulse

transfer function for a second-order overdamped system with a zero-order hold. The transfer function of the process is

$$G(s) = \frac{K}{(\tau_1 s + 1)(\tau_2 s + 1)}$$

where K, τ_1, and τ_2 are constants. Check that the steady-state gain of the pulse transfer function is the same as that of the continuous transfer function.

14-8. Do Problem 14-7 for a second-order underdamped system:

$$G(s) = \frac{K}{25s^2 + 7s + 1}$$

14-9. Do Problem 14-7 for the case $\tau_1 = \tau_2 = \tau$ (critically damped).

14-10. Write the pulse transfer function for each of the following discrete blocks:
   (a) A dead-time or delay block: $Y_n = X_{n-N}$
   (b) An accumulator block: $Y_n = Y_{n-1} + X_n$
   (c) A two-mode (proportional-integral) controller block:

$$M_n = M_{n-1} + K_c \left[(E_n - E_{n-1}) + \frac{T}{\tau_I} E_n \right]$$

14-11. Bristol (1977) proposes the following discrete PI control algorithm, which he attributes to F.G. Shinskey.

$$Y_n = Y_{n-1} + \left(\frac{T}{T + \tau_I} \right)(M_{n-1} - Y_{n-1})$$

$$M_n = Y_n + K_c E_n$$

Find the overall transfer function of the algorithm, $D(z) = M(z)/E(z)$.

14-12. Derive the pulse transfer function and the recursive formula for a discrete block to simulate a lead-lag unit.

$$G(s) = \frac{Y(s)}{X(s)} = \frac{\tau_{ld} s + 1}{\tau_{lg} s + 1}$$

Check that the steady-state gain of the discrete block matches that of the continuous block. Calculate also the first five sampled values of the output for a unit step in input. Use $\tau_{lg} = 1$ min and (a) $\tau_{ld} = 1.5$ min (b) $\tau_{ld} = 0.5$ min ($T = 0.1$).

14-13. Develop the recursive computing formula to simulate a continuous system with the transfer function

$$G(s) = \frac{Y(s)}{X(s)} = \frac{1.5e^{-2s}}{(5s + 1)(3s + 1)}$$

Use the recursive equation to compute the unit step and ramp responses of the process sampled every $T = 1$ min. All the time parameters are in minutes.

14-14. Consider the sampled-data feedback control loop of Fig. 14-4.1 with a zero-order hold device, an integral controller, and a first-order lag process:

$$D(z) = \frac{K_I}{1 - z^{-1}} \qquad G_P(s) = \frac{K}{\tau s + 1}$$

Obtain the expression for the closed-loop pulse transfer function, $C(z)/R(z)$; the steady-state gain; and the ultimate loop gain. Also sketch the root locus on the z-plane.

14-15. Do Problem 14-14 for a proportional-integral controller:

$$D(z) = K_c\left[1 + \frac{T}{\tau_I} \frac{1}{1 - z^{-1}} \right]$$

where K_c is the gain, τ_1 is the reset time, and T is the sample time.

14-16. Do Problem 14-14 for a first-order plant with one sample of dead time:

$$G_P(s) = \frac{Ke^{-Ts}}{\tau s + 1}$$

Such a transfer function can represent a first-order process with an analyzer controller, such as composition control on a stirred tank.

14-17. Consider the sampled-data control loop of Fig. 14-4.1, with $T = 0.5,$ a zero-order hold, a proportional controller, $D(z) = K_c$, and the following transfer function:

$$G_P(s) = \frac{1.4}{(s + 1)(0.8s + 1)}$$

Find the closed-loop transfer function $C(z)/R(z)$, the steady-state closed-loop gain, and the offset for a unit step change in set point. Write the characteristic equation of the loop, find the ultimate gain, and sketch the root locus.

14-18. Do Problem 14-17 for the process transfer function

$$G_P(s) = \frac{1}{As}$$

where A is constant. Such a transfer function can represent a level process.

14-19. Obtain the modified z-transforms of the functions of Problem 14-1.

14-20. Obtain the modified z-transforms of the functions of Problem 14-2.

14-21. Find the z-transform of the delayed function

$$f(t) = u(t - 3)[1 - e^{-(t-3)/4}]$$

sampled at intervals of $T = 2:$
(a) By applying the definition of the z-transform to the function as given.

(b) By the modified z-transform with the appropriate value of m (a constant) and the real translation theorem.

Draw a sketch of the original and delayed functions, showing the sampled values.

14-22. Develop the pulse transfer function for the general second-order-plus-dead-time (SOPDT) transfer function:

$$G_P(s) = \frac{Ke^{-t_0s}}{(\tau_1 s + 1)(\tau_2 s + 1)}$$

with zero-order hold. Assume the dead time is $t_0 = (N + \Delta)T$, where A is a fraction, $0 \leq A < 1$.

14-23. Do Problem 14-14 for the process transfer function $(T = 1)$

$$G_P(s) = \frac{0.9e^{-0.8s}}{5s + 1}$$

(a) Using the modified z-transform to obtain $HG_P(z)$ exactly.

(b) Approximating the dead time by the nearest integer multiple of the sample time. Is the approximation good enough for controller design? Discuss.

14-24. Do Problem 14-23 for the process transfer function

$$G_P(s) = \frac{0.9e^{-0.8s}}{(s + 1)(0.7s + 1)}$$

Chapter **15**

Design of Computer Control Systems

This chapter presents several special aspects of the design of computer control systems. As we noted in the preceding chapter, computers perform their control calculations recursively at uniform intervals of time. This requires sampling the continuous process signals and quickly calculating the signals to the process, then updating these output signals and holding them constant until the next update. This mode of operation is common to distributed control systems (DCS), programmable logic controllers (PLC), and computer control systems. Calculations of the output are done by modular programs, or *algorithms,* that emulate any of the control devices we studied in the first 13 chapters of this book: feedback controllers (Chapter 6), lead-lag units (Chapter 12), and computing blocks and selector switches (Chapter 11). We discussed the use of these devices in process control systems in those earlier chapters, so this chapter will concentrate on the way the computers carry out the calculations-that is, on the development of the algorithms. We will also examine tuning formulas for computer feedback controllers and look at some special computer algorithms: dead-time compensation feedback controllers, self-tuning controllers, and model-reference controllers.

15-1 DEVELOPMENT OF CONTROL ALGORITHMS

This section presents the development of the three most important computer control algorithms: filters, lead-lag units, and feedback control algorithms. We will develop these algorithms from the transfer functions of their continuous counterparts. When developing algorithms that other engineers and technicians will use, it is very important to ensure as far as possible that they are robust-that is, that they do not go unstable for certain combinations of their adjustable parameters. This is of course not possible for feedback controllers, because their stability depends on proper tuning, but we will demonstrate how different approximation formulas can lead to algorithms that are unnecessarily sensitive to their adjustable parameters.

The general formula we will use for approximating continuous algorithms is the

backward difference approximation to the derivative:

$$\frac{dC(t)}{dt} \triangleq \frac{C(nT)}{T} - \frac{C(nT - T)}{T} = \frac{C_n}{T} - \frac{C_{n-1}}{T} \qquad (15\text{-}1.1)$$

This formula is easier to use in operator form. The idea is to start with the **Laplace** transfer function and replace the s variable with the z-transform variable. To do this, take the **Laplace** transform of the left-hand side of Eq. 15-1.1 and the z-transform of the right-hand side to obtain

$$sC(s) \triangleq \frac{\mathcal{Z}[C(nT)] - \mathcal{Z}[C(nT - T)]}{T} = \frac{1 - z^{-1}}{T} C(z) \qquad (15\text{-}1.2)$$

Then treat the z variable as the shift operator to obtain

$$SC(S) \triangleq \frac{C(z) - z^{-1}C(z)}{T} = \frac{C_n - C_{n-1}}{T}$$

which is the same result as in Eq. 15-1.1. The key relationship between the operators is then, from Eq. 15-1.2,

$$\boxed{s \triangleq \frac{1 - z^{-1}}{T}} \qquad (15\text{-}1.3)$$

This is the substitution we will use to develop computer algorithms from the transfer functions of the continuous control devices.

15-1.1 Exponential Filter

The exponential filter is the standard filter used in digital control systems to attenuate noisy signals. It is an emulation of the original "R-C" or *low-pass filters*. The filter is a first-order lag with unity gain:

$$Y(s) = \frac{1}{\tau_F s + 1} X(s) \qquad (15\text{-}1.4)$$

where $Y(s)$ is the filtered output, $X(s)$ is the noisy input, and τ_F is the filter time constant, an adjustable parameter. Substitute Eq. 15-1.3 and replace the **Laplace** transforms with the z-transforms of the sampled signals to obtain

$$Y(z) = \frac{1}{\tau_F\left(\dfrac{1 - z^{-1}}{T}\right) + 1} X(z) = \frac{T}{(\tau_F + T) - \tau_F z^{-1}} X(z)$$

where we multiplied numerator and denominator by T and rearranged the denominator. The next step is to clear fractions.

$$(\tau_F + T)Y(z) - \tau_F z^{-1}Y(z) = TX(z)$$

As in Section 14-3.5, the recursive formula is obtained by replacing the z-transforms with the sampled values and treating the z variable as the shift operator.

$$(\tau_F + T)Y_n - \tau_F Y_{n-1} = TX_n$$

Finally, solve for the output,

$$\boxed{Y_n = aY_{n-1} + (1 - a)X_n}$$ **(151.5)**

where a $= \dfrac{\tau_F}{\tau_F + T}$.

This is close to the algorithm we obtained in Example **14-3.5(b),** Eq. 14-3.17. There we used the pulse transfer function with zero-order hold to obtain the recursive equation that simulates a first-order continuous process. Once we set the gain to unity, there are two basic differences between Eqs. 14-3.17 and **15-1.5:**

- The input signal is delayed by one sample in Eq. 14-3.17. This is because it is a simulation of the continuous first-order process that must always be at least one sample behind because of the order in which the computer performs its calculations. This unnecessary delay, which is not present in Eq. 15-1.5, must be avoided if the filter output is to be used in a control calculation.
- Parameter a in Eq. 14-3.17 is defined by e^{-T/τ_F}, which is a more exact relationship between the filter time constant and the parameter in the discrete formula. The value of a in Eq. 15-1.5 approaches the value given by the exponential formula when $T \ll \tau_F$, but precise matching of this relationship is not as important as making sure that a is in the range $0 \le a < 1$. It is easy to see that the formula in Eq. 15-1.5 is in this range as long as the filter time constant is positive.

The reason why it is important that parameter a be positive and less than unity in the exponential filter algorithm is that the algorithm transfer function has a pole at z = a. This means that the output would ring (switch back and forth around its **steady-**state value) if a were negative or would be unstable if a were greater than unity (see Section 14-4.2). Figure 15-1.1 shows the step and ramp responses of the filter algorithm to a noisy signal. Increasing the filter time constant increases noise attenuation at the expense of increasing the lag on the input.

To appreciate how the wrong approximation can lead to an algorithm with undesirable response, consider that if we had used the forward difference approximation for the s operator, $s = (z - 1)/T$, we would have obtained Eq. 15-1.5 but with parameter $a = 1 - (T/\tau_F)$. In this case, parameter a is negative when $\tau_F < T$; that is, the output of the filter will ring for certain values of the adjustable parameters. With the backward difference approximation, Eq. 15-1.5, the most that happens when the filter time constant is reduced all the way to zero is that the input signal is passed through to the
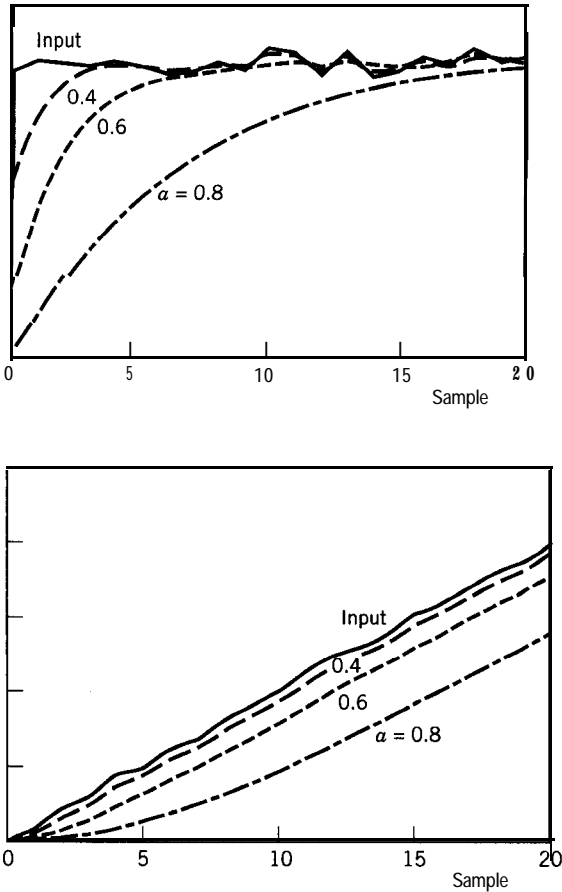
Figure   15-1.1 Response of an exponential filter to a step and to a ramp.

output without filtering. This is what anyone setting the filter time constant to zero would want to happen, which is why Eq. 151.5 is a robust algorithm for the exponential filter.

In distributed and computer control systems, the exponential filter algorithm is not a separate block but rather a filter option that can be applied to any process signal as part of the input processing. When a lag is to be used as a dynamic compensating block, the lead-lag algorithm is used with the lead set to zero. We will develop this algorithm next.

15-1.2 Lead-Lag Algorithm

The continuous lead-lag unit, which we discussed in Chapters 2 and 12, can have three adjustable parameters: the gain K, lead constant τ_{ld}, and lag constant τ_{lg}. Its transfer function is:

$$Y(s) = K \frac{\tau_{ld}s + 1}{\tau_{lg}s + 1} X(s) \tag{15-1.6}$$

where Y(s) is the output and X(s) the input. Substitute the approximation of the s operator, Eq. 15-1.3, and simplify to obtain

$$Y(z) = K \frac{\tau_{ld}(1 - z^{-1}) + T}{\tau_{lg}(1 - z^{-1}) + T} X(z)$$

Replace the transforms with the sampled values, treat the z variable as the shift operator, and solve for the output to obtain the recursive formula. It is

$$Y_n = aY_{n-1} + K\left[(1 - a)X_n + \frac{\tau_{ld}}{\tau_{la} + T} (X_n - X_{n-1}) \right] \quad (15\text{-}1.7)$$

where $a = \dfrac{\tau_{lg}}{\tau_{lg} + T}$.

Note that if we set $\tau_{ld} = 0$, $K = 1$, this becomes exactly the same as the exponential filter formula, Eq. 15-1.5. The extra term introduced by the lead is proportional to the change in input at each sample, amplified or attenuated by approximately the ratio of the lead to the lag (assuming the sample time is small relative to the lag constant).

Equation 15-1.7 is robust and can accept values of the lag constant all the way to zero, but this is not recommended, because for fast sampling, the ratio τ_{ld}/T **can** be too large and can cause undesirable pulses on the output. Robustness calls for a lower limit to be imposed on the lag constant equal to about one-tenth the lead constant; that is, $\tau_{lg} > 0.1 \ \tau_{ld}$. This is accomplished by building the limit into the lead-lag program module. Figure 15-1.2 shows the step and ramp responses of the lead-lag algorithm. They match the responses of the continuous units presented in Figs, 2-4.5 and 2-4.6.

Incremental Algorithm

The lead-lag unit is one algorithm that is often programmed in incremental or "velocity" form. Doing so facilitates combining its output with other incremental algorithms such as the feedback control algorithm. Other advantages of the incremental algorithm are that it is easy to initialize and that its output does not jump when its adjustable parameters are changed. The incremental algorithm computes the increment of the output at each sample, instead of its full value:

$$\Delta Y_n = Y_n - Y_{n-1}$$

Substitute Eq. 15-1.7 for Y_n and rearrange to obtain the incremental form of the algorithm.

$$\Delta Y_n = (1 - a)[KX_n - Y_{n-1}] + K\left(\frac{\tau_{ld}}{\tau_{lg} + T}\right)(X_n - X_{n-1}) \quad (15\text{-}1.8)$$

The actual output is obtained by the standard accumulation formula.
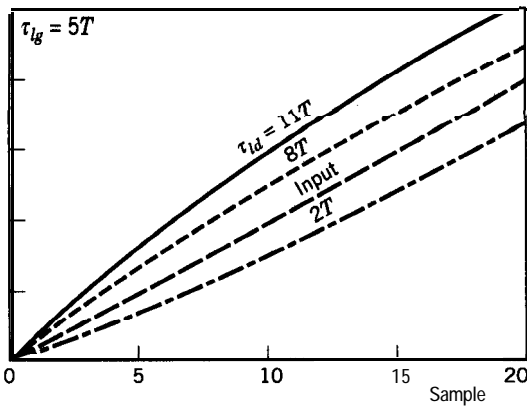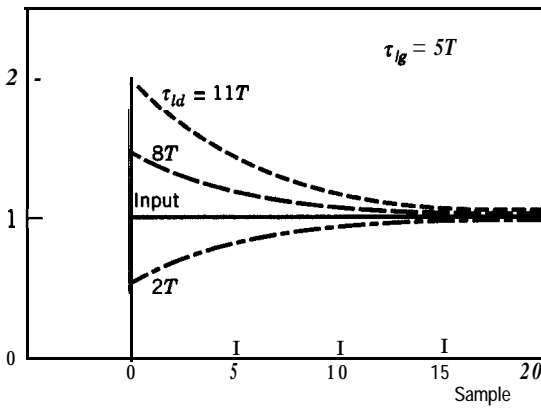
Figure **15-1.2** Response of a lead-lag algorithm to a step
and to a ramp.

$$Y_n = Y_{n-1} + \Delta Y_n \qquad\qquad (15\text{-}1.9)$$

This last formula can be extended to have more than one lead-lag unit affect the same
output. Such extension will be necessary when adjusting one output signal to compen-
sate for more than one disturbance in a feedforward control scheme. If each lead-lag
unit acts on the same output signal Y, then each can carry out the computations of Eqs.
15-1.8 and 15-1.9 independently of each other. The final signal will contain the sum of
the increments from all the lead-lag units. An incremental feedback control algorithm
could be added to the scheme just as easily. Hence the advantage of using incremental
algorithms.

15-1.3   Feedback (PID) Control Algorithms

The basic computer feedback control algorithm is an emulation of the three-mode
controller and is generically known as the *PID algorithm* (for **proportional-integral-**

derivative). Because of the flexibility of the program modules, the basic algorithm offers a number of options so that it can be applied to many different control situations. In this section, we will develop two incremental (velocity) feedback control algorithms: one corresponding to Eq. 53.17, which we will call the **parallel** version, and the other corresponding to Eq. 5-3.19, which we will call the **series** version. Both of these algorithms contain a noise-attenuating filter on the derivative term. This filter is absolutely necessary for the derivative term to work in computer control algorithms.

Parallel   Incremental   Algorithm

This algorithm is based on Eq. 5-3.17, which will be our starting point.

$$\frac{M(s)}{E(s)} = K_c\left[1 + \frac{1}{\tau_I s} + \frac{\tau_D s}{\alpha\tau_D s + 1}\right] \tag{5-3.17}$$

where $M(s)$ is the controller output and $E(s)$ is the error, calculated as the set point minus the measurement or process variable: $E(s) = R(s) - C(s)$. As before, we approximate the s variable by the backward difference approximation, Eq. 15-1.3, and rearrange the result into

$$M(z) = K_c\left[1 + \frac{T}{\tau_I(1 - z^{-1})} + \frac{\tau_D(1 - z^{-1})}{\alpha\tau_D(1 - z^{-1}) + T}\right]E(z)$$

To obtain the incremental algorithm, multiply the equation by $(1 - z^{-1})$, and realize that $AM(z) = (1 - z^{-1})M(z)$.

$$\Delta M(z) = (1 - z^{-1})M(z)$$

$$= K_c\left[(1 - z^{-1}) + \frac{T}{\tau_I} + \frac{\tau_D(1 - z^{-1})^2}{\alpha\tau_D(1 - z^{-1}) + T}\right]E(z) \tag{151.10}$$

This equation is best implemented in two steps, as follows:

$$\Delta M(z) = K_c\left[(1 - z^{-1})E(z) + \frac{T}{\tau_I}E(z) + Y(z)\right]$$

$$Y(z) = \frac{\tau_D(1 - z^{-1})^2}{\alpha\tau_D(1 - z^{-1}) + T}E(z)$$

$$= \frac{\tau_D(1 - 2z^{-1} + z^{-2})}{\alpha\tau_D(1 - z^{-1}) + T}E(z)$$

Replace the z-transforms with the sampled values and treat the z variable as the shift operator to obtain

$$Y'' = \frac{\alpha \tau_D}{\alpha \tau_D + T} Y_{n-1} + \frac{\tau_D}{\alpha \tau_D + T} (E_n - 2E_{n-1} + E_{n-2})$$

$$\Delta M_n = K_c \left[E_n - E_{n-1} + \frac{T}{\tau_I} E_n + Y_n \right]$$

(151.11)

The actual controller output is calculated by a standard accumulator:

$$M_n = M_{n-1} + \Delta M_n$$

(151.12)

In Eq. 15-1.11, the variable Y_n is the derivative term. Let us examine the effect of the filter with time constant $\alpha \tau_D$ in its calculation. If we were to eliminate the filter by setting $\alpha = 0$, the derivative term Y_n would become

$$Y_n = \frac{\tau_D}{T} (E_n - 2E_{n-1} + E_{n-2})$$

For fast sampling, $T \le 1$ s, the multiplier of the second difference of the error becomes too large for even reasonable derivative times. For example, for $\tau_D = 1$ min (60 s), $T = 1$ s, the multiplier is 60, which causes large pulses in the controller output for small changes in the error. Many original computer algorithms that did not include the filter on the derivative term made it impractical to use the derivative because of the large sensitivity of the output to changes in the error. On the other hand, for fast sampling, the derivative filter limits the multiplier on the second difference of the error to $1/\alpha$, which is the value we obtain in Eq. 15-1.11 for $T \ll \alpha \tau_D$. This is why the parameter α, which is usually fixed at about 0.1, is called the "dynamic gain limit" on the derivative. The surprising thing is that the filter does not affect the actual performance of the control algorithm, because its time constant is so small (about one-tenth the derivative time).

Another important modification of the control algorithm that prevents undesirable pulses is to remove set point changes from the derivative term. To do this, recalling that $E_n = R_n - C_n$, we replace E_n with $- C_n$ in the calculation of the derivative term.

$$Y_n = \frac{\alpha \tau_D}{\alpha \tau_D + T} Y_{n-1} - \frac{\tau_D}{\alpha \tau_D + T} (C_n - 2C_{n-1} + C_{n-2})$$

(151.13)

where C is the controlled variable, %TO. The minus sign goes with the convention introduced in Section 5-3 for calculation of the error. Recall that with this convention,

a positive controller gain results in a reverse-acting controller, and a negative gain results in a direct-acting controller.

Series Control Algorithm

The series form of the algorithm is the one used by most distributed control systems (DCS). To develop the series form of the algorithm, we start with Eq. 5-3.19.

$$\frac{M(s)}{E(s)} = K_c'\left(1 + \frac{1}{\tau_I's}\right)\left(\frac{\tau_D's + 1}{\alpha\tau_D's + 1}\right) \tag{5-3.19}$$

where the primes on the parameters indicate that they are not the same as for the parallel form of the algorithm. Equation 5-3.20 gives the relationships between the parameters of the two forms of the algorithm. The term *series* is used because the two brackets, containing a proportional-integral and a proportional-derivative term, respectively, are in series, as shown in Fig. 15-1.3a, which is a block diagram of the controller.

To avoid including the set point in the derivative term, we apply the **proportional-**derivative term to the controlled variable before the error is calculated, as shown in Fig. 15-1.3b. The equations for this modified diagram are

$$Y(s) = \frac{\tau_D's + 1}{\alpha\tau_D's + 1}C(s)$$

$$E(s) = R(s) - Y(s) \tag{15-1.14}$$

$$M(s) = K_c'\left[1 + \frac{1}{\tau_I's}\right]E(s)$$

To emulate these transfer functions on a computer, convert the **Laplace** transforms into z-transforms and use Eq. 15-1.3.
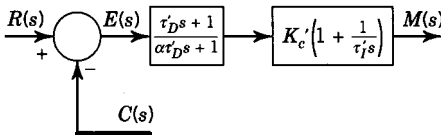
(a)
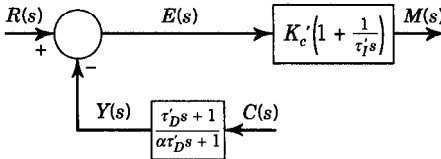
(b)

Figure 151.3 Block diagram of a series PID controller. (a) With derivative on error. (b) With derivative on measurement.

$$Y(z) = \frac{\tau_D'(1 - z^{-1}) + T}{\alpha\tau_D'(1 - z^{-1}) + T} C(z)$$

$$E(z) = R(z) - Y(z)$$

$$AM(z) = (1 - z^{-1})M(z) = K_c'\left[1 - z^{-1} + \frac{T}{\tau_I'}\right]E(z)$$

Rearrange and replace the z-transforms with the sampled values to obtain the recursive computing formula.

$$Y_n = aY_{n-1} + (1 - a)C_n + \frac{\tau_D'}{\alpha\tau_D' + T}(C_n - C_{n-1})$$

$$E_n = R_n - Y_n \qquad\qquad \textbf{(15-1.15)}$$

$$\Delta M_n = K_c'\left[E_n - E_{n-1} + \frac{T}{\tau_I'}E_n\right]_1$$

where $a = \dfrac{\alpha\tau_D'}{\alpha\tau_D' + T}$. Note that the first equation is a lead-lag algorithm, Eq. 15-1.7, with a gain of unity, a lead of τ_D', and a lag of $\alpha\tau_D'$. Because the value of α is usually set to around 0.1, the proportional-derivative term introduces a net lead into the loop of $(1 - \alpha)\tau_D' = 0.9\tau_D'$. Other than this slight (10%) decrease in the net lead of the derivative term, the filter does not affect the performance of the control algorithm. However, as with the parallel form, the filter prevents large output pulses when the sample time T is small. As before, Eq. 15-1.12 accumulates the increment AM,, to produce the controller output M_n at each execution of the algorithm.

The following example shows what happens when the filter on the derivative is not used and when the derivative acts on the error instead of on the controlled variable.

EXAMPLE 15-1.1

A temperature controller uses the parallel form of the PID algorithm, Eq. 151.11, executed once a second ($T = 1$ s). Its tuning parameters are $K_c = 4.0$ %CO/%TO, $\tau_1 = 8.0$ min, and $\tau_D = 2.0$ min. The temperature transmitter has a range of 50°C to 100°C. After a period of steady operation, the temperature changes slightly by 0.1°C and then stays constant at the new value for a few seconds (before the controller action is felt). Compute the effect on the controller output of this small change in temperature (a) when no filter is used on the algorithm ($\alpha = 0$), and (b) when the algorithm has a filter with $\alpha = 0.1$. Also calculate the effect on the controller output of a change of 1°C in set point, assuming the derivative term acts on the error, as in Eq. 15-1.11.

SOLUTION

(a) No derivative filter, $\alpha = 0$. First, substitute the values of all the parameters into Eq. 15-1.11.

$$Y_n = \frac{(2.0 \text{ min})(60 \text{ s/min})}{1 \text{ s}} (E_n - 2E_{n-1} + E_{n-2})$$

$$= 120(E_n - 2E_{n-1} + E_{n-2})$$

$$\Delta M_n = 4.0 \left[E_n - E_{n-1} + \frac{1 \text{ s}}{(8.0 \text{ min})(60 \text{ s/min})} E_n + Y_n \right]$$

$$= 4.0 \left[E_n - E_{n-1} + \frac{1}{480} E_n + Y_n \right]$$

When the temperature increases by $0.1°C$, the error becomes

$$E_n = R_n - C_n = (0 - 0.1)°C \frac{100 \text{ \%TO}}{(100 - 50)°C} = -0.2 \text{ \%TO}$$

where the set point R_n is zero because it is constant. The error then remains equal to -0.2 %TO for several samples. The system has been running steady before the change in temperature, so the past values of the error are $E_{n-1} = 0$, $E_{n-2} = 0$, and, from the foregoing equations,

$$Y_n = 120(-0.2 - 2 \cdot 0 + 0) = -24$$

$$\Delta M_n = 4.0 \left[-0.2 - 0 + \frac{-0.2}{480} + (-24) \right] = -96.8 \text{ \%CO}$$

This is quite a jump in controller output. The output is bound to saturate at zero, because its initial value is somewhere toward the middle of the scale. Note that essentially all of the jump is caused by the derivative term, Y_n. At the next execution of the algorithm, the error is still $E_n = -0.2$ %TO, but now the previous error is also $E_{n-1} = -0.2$ %TO, and $E_{n-2} = 0$, so the output values are

$$Y_n = 120[-0.2 - 2(-0.2) + 0] = 24$$

$$\Delta M_n = 4.0 \left[-0.2 - (-0.2) + \frac{-0.2}{480} + (24) \right] = 96.0 \text{ \%CO}$$

This jump, which is of about the same magnitude as the previous one but of opposite direction, is caused by the derivative term. At best, if the output does not saturate, a large pulse of one sample duration has occurred in the controller output. This pulse may not upset the process much if it lasts only one second, but look at what happens if the output saturates: the valve is slammed shut in one sample and almost completely opened in the next. Such an action greatly upsets the process, because

the initial value of the output has been lost and it will take a while for the integral mode to find it again.

At the next sample, all three values of the error, E_n, E_{n-1}, and E_{n-2}, are equal to -0.2 %TO. Thus $Y_n = 0$, and the output is determined only by the integral term, $AM_{,,} = 4.0(-0.2)/480 = -0.0017$ %CO, which is the action required to eventually bring the temperature back to its set point.

(b) Filtered derivative term with $\alpha = 0.1$. Of the foregoing formulas, only the one for Y_n changes to

$$Y_n = \frac{(0.1)(2.0 \text{ min})(60 \text{ s/min})}{(0.1)(2.0 \text{ min})(60 \text{ s/min}) + (1 \text{ s})} Y_{n-1}$$

$$+ \frac{(2.0 \text{ min})(60 \text{ s/min})}{(0.1)(2.0 \text{ min})(60 \text{ s/min}) + (1 \text{ s})} (E_n - 2E_{n-1} + E_{n-2})$$

$$= 0.923 Y_{n-1} + 9.231(E_n - 2E_{n-1} + E_{n-2})$$

The first sample after the change in temperature, the errors are $E_n = -0.2$ %TO, $E_{n-1} = 0$, and $E_{n-2} = 0$, so the output is

$$Y_n = 0.923 \cdot 0 + 9.231(-0.2 - 2 \cdot 0 + 0) = -1.846$$

$$\Delta M_n = 4.0 \left[-0.2 - 0 + \frac{-0.2}{480} + (-1.846) \right] = -8.186 \text{ %CO}$$

The calculations then proceed as follows:

E_n	E_{n-1}	E_{n-2}	Y_n	ΔM_n
-0.2	0	0	-1.846	-8.186
-0.2	-0.2	0	0.142	0.566
-0.2	-0.2	-0.2	0.131	0.523
-0.2	-0.2	-0.2	0.121	0.482

With the filter, the initial jump in the controller output is only a little over 8 %CO. This jump then slowly decays to zero, as the positive increments in output, contributed by the derivative term Y_n, eventually cancel the initial negative change. Because the time constant of the filter is $(0.1)(2.0) = 0.2$ min, it will take approximately five time constants, or 1.0 min (60 samples), for the initial change to decay to zero. In other words, instead of providing one large pulse for one sample, the filtered derivative provides a much smaller change spread over many samples.

Let us next consider the change in set point of $1.0°C$. When this is applied, assuming there is no change in the controlled temperature, the error changes to

$$E_n = R_n - C_n = 1.0°C \frac{100 \text{ %TO}}{(100 - 50)°C} - 0 = 2.0 \text{ %TO}$$

This change in error is 10 times greater than the change in temperature assumed earlier, so the changes in controller output, if the derivative term acts on the error, will be 10 times the values calculated earlier, or 960% for the unfiltered derivative and 81.86% for the filtered derivative with $\alpha = 0.1$. Both of these changes are unacceptable. On the other hand, if the derivative acts only on the controlled variable, as in Eq. 15-1.12, then there is no jump in the controller output, because the derivative term does not see the change in set point at all. Only the proportional and integral respond to the set point change, causing an increment of

$$\Delta M_n = 4.0 \left[2.0 - 0 + \frac{2.0}{480} + 0 \right] = 8.017 \ \%CO$$

This jump, most of which is due to the proportional term, occurs only at the sample after the set point is changed. Then the integral mode causes increments of 0.017 %CO at each sample until the temperature begins to increase toward the new set point.

The preceding example shows why computer control algorithms must have the filter on the derivative term and why it is important to have the derivative term act on the controlled variable, not on the error. The original computer control algorithms did not provide the filter on the derivative term, which made the derivative mode unusable when the sample time was small. They could just as well have been programmed as PI algorithms.

Having introduced the most important computer algorithms, we will next look at tuning formulas for the feedback control algorithms.

15-2 TUNING OF FEEDBACK CONTROL ALGORITHMS

The tuning procedures of Chapter 7 apply to computer control algorithms for fast sampling systems-that is, for systems in which the sample time is one-tenth or less of the dominant time constant of the process. In Chapter 7 we presented a correction for sampling for the tuning formulas based on open-loop models of the process-namely, to add half the sample time to the process dead time and use the corrected dead time in the tuning formulas. This approximation holds as long as the sample time is small relative to both the process dead time and the time constant. In this section we will develop tuning formulas for computer control algorithms using the controller synthesis procedure of Section 7-4. These formulas can be applied to any combination of process parameter values and sample times, because they are based on fundamental principles and not on empirical correlations. We will base the tuning formulas on a **second-order-** plus-dead-time (SOPDT) model of the process, because they can be easily reduced to first-order-plus-dead-time (FOPDT) models.

15-2.1 Development of the Tuning Formulas

Consider the block diagram of a computer feedback control loop shown in Fig. 15-2.1. In the diagram, the transfer function of the control algorithm is labeled D(z), and all of
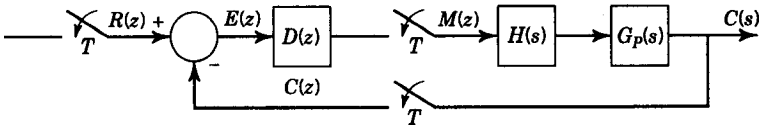
Figure 15-2.1 Block diagram of computer feedback control loop.

the field elements are grouped in the continuous transfer function $G_P(s)$. This includes the transfer functions of the control valve, process, and sensor/transmitter. The samplers are assumed to be synchronous with a sampling interval T. As we saw earlier, this assumption neglects the time it takes to perform the control calculations. That is, we assume that the controller output M(z) is computed from the value of the process variable C(s) that is sampled at the same instant the output is updated. The zero-order hold, H(s), holds the controller output constant between updates. In Section 14-3 we obtained the transfer function of the closed loop, Eq. 14-4.5, which, without the disturbance, is

$$\frac{C(z)}{R(z)} = \frac{HG_P(z)D(z)}{1 + HG_P(z)D(z)} \tag{152.1}$$

As in Section 7-4, we solve for $D(z)$ to obtain the controller synthesis formula.

$$D(z) = \frac{M(z)}{E(z)} = \frac{1}{HG_P(z)} \frac{C(z)/R(z)}{1 - [C(z)/R(z)]} \tag{152.2}$$

The synthesized controller depends on the process transfer function, $HG_P(z)$, and on the specified closed-loop response, $C(z)/R(z)$. For the closed-loop response, we will use a first-order lag with unity gain, as specified by Dahlin (1968). The pulse transfer function is that of a first-order lag with zero-order hold, developed in Example 14-3.2. **For $K = 1$:**

$$\frac{C(z)}{R(z)} = \frac{(1 - q)z^{-1}}{1 - qz^{-1}}$$

where $q = e^{-T/\tau_c}$, and τ_c is the closed-loop time constant, an adjustable parameter. However, as we noted in Section 7-4.3, to allow for process dead time, we must delay the closed loop transfer function by the process dead time. Here we will delay it by the number of samples in the dead time, N.

$$\frac{C(z)}{R(z)} = \frac{(1 - q)z^{-N-1}}{1 - qz^{-1}} \tag{152.3}$$

where $N = t_0/T$, truncated to the closest smaller integer. The algorithm based on this response was originally proposed by Dahlin (1968) and (independently) by **Higham** (1968). It is known as the **Dahlin–Higham** response or simply the *Dahlin response.*

Substitute Eq. 15-2.3 into Eq. 15-2.2.

$$D(z) = \frac{M(z)}{E(z)} = \frac{1}{HG_P(z)} \frac{(1 - q)z^{-N-1}}{1 - qz^{-1} - (1 - q)z^{-N-1}} \qquad (15\text{-}2.4)$$

For the process model, we will use the following second-order-plus-dead-time (SOPDT) transfer function:

$$G_P(s) = \frac{K(\tau_3 s + 1)e^{-t_0 s}}{(\tau_1 s + 1)(\tau_2 s + 1)} \qquad (15\text{-}2.5)$$

where

$$K = \text{gain, } \%\text{TO}/\%\text{CO}$$
$$t_0 = \text{dead time, min}$$
$$\tau_1, \tau_2, \tau_3 = \text{time constants, min}$$

The numerator term is included because it allows the model to match a wide variety of process responses, including processes with inverse response, for which $\tau_3 < 0$. To allow for the general case in which the dead time is not an exact multiple of the sample time, we will use the modified z-transform (Section 14-5) to develop the pulse transfer function. Let $t_0 = NT + AT$, where A is a fraction: $0 \leq A < 1$. Then the pulse transfer function is

$$HG_P(z) = K(1 - z^{-1})\mathscr{L}\left[\frac{(\tau_3 s + 1)e^{-NTs}e^{-\Delta Ts}}{s(\tau_1 s + 1)(\tau_2 s + 1)} \right]$$

$$= K(1 - z^{-1})z^{-N}\mathscr{L}_m\left[\frac{A_1}{s} + \frac{A_2}{\tau_1 s + 1} + \frac{A_3}{\tau_2 s + 1} \right]_{m=1-A}$$

After evaluating the coefficients by the standard procedure, taking the indicated modified z-transforms with the help of Table 14-2.1, and simplifying, we get

$$HG_P(z) = \frac{(b_0 + b_1 z^{-1} + b_2 z^{-2})z^{-N-1}}{1 - a_1 z^{-1} - a_2 z^{-2}} \qquad (15\text{-}2.6)$$

where

$$a_1 = \beta_1 + \beta_2 \qquad \beta_1 = e^{-T/\tau_1} \qquad \beta_2 = e^{-T/\tau_2}$$
$$a_2 = -\beta_1 \beta_2$$

$$b_0 = K\left[1 - \left(\frac{\tau_1 - \tau_3}{\tau_1 - \tau_2}\right)\beta_1^m + \left(\frac{\tau_2 - \tau_3}{\tau_1 - \tau_2}\right)\beta_2^m \right]$$

$$b_1 = K\left[-a_1 + \left(\frac{\tau_1 - \tau_3}{\tau_1 - \tau_2}\right)(1 + \beta_2)\beta_1^m - \left(\frac{\tau_2 - \tau_3}{\tau_1 - \tau_2}\right)(1 + \beta_1)\beta_2^m \right]$$

$$b_2 = K\left[-a_2 - \left(\frac{\tau_1 - \tau_3}{\tau_1 - \tau_2}\right)\beta_2\beta_1^m + \left(\frac{\tau_2 - \tau_3}{\tau_1 - \tau_2}\right)\beta_1\beta_2^m \right]$$

Next we substitute Eq. 15-2.6 into Eq. 15-2.4 to obtain the synthesized controller.

$$D(z) = \frac{1 - a_1 z^{-1} - a_2 z^{-2}}{b_0 + b_1 z^{-1} + b_2 z^{-2}} \frac{1 - q}{1 - q z^{-1} - (1 - q) z^{-N-1}} \tag{152.7}$$

where the term z^{-N-1} has canceled. To use this equation for tuning the PID algorithm, we must match its transfer function. For the parallel form, from Eq. 151.10, neglecting the filter $(\alpha = 0)$, the transfer function that must be matched is

$$D(z) = \frac{M(z)}{E(z)} = K_c \left[\frac{1 - z^{-1} + \dfrac{T}{T_I} + \dfrac{T_D}{T}(1 - z^{-1})^2}{1 - z - r} \right]$$

$$= K_c \left[\frac{\left(1 + \dfrac{T}{T_I} + \dfrac{T_D}{T}\right) - \left(1 + 2\dfrac{T_D}{T}\right) z^{-1} + \dfrac{T_D}{T} z^{-2}}{1 - z^{-1}} \right] \tag{152.8}$$

Comparing Eqs. 15-1.7 and 15-1.8, we note that the numerator polynomials match in order but that the denominators do not match. To proceed, we must simplify the synthesized controller without affecting its gain. The first term in the denominator is dropped as follows:

$$\text{Let} \quad b_0 + b_1 z^{-1} + b_2 z^{-2} = b_0 + b_1 + b_2$$

where, with z set equal to 1, the gain of the algorithm is not affected. This removes two poles from the algorithm transfer function. It has been found that removing these poles is beneficial in most situations, because at least one of them (and often both) is negative and causes the controller output to ring-that is, to switch back and forth above and below the steady-state value. Ringing is undesirable because it causes unnecessary wear on the mechanical components of the loop. Substitute the simplified term into Eq. 15-2.7.

$$D(z) = \left(\frac{1 - q}{b_0 + b_1 + b_2}\right) \frac{1 - a_1 z^{-1} - a_2 z^{-2}}{1 - q z^{-1} - (1 - q) z^{-N-1}} \tag{152.9}$$

We will return to this algorithm in the next section as an example of a feedback algorithm with dead-time compensation, but we must continue to simplify it to match Eq. 15-2.8. To do this, we factor the denominator and keep only the term we want.

$$1 - q z^{-1} - (1 - q) z^{-N-1} = (1 - z^{-1})[1 + (1 - q) z^{-1} + \cdots + (1 - q) z^{-N}]$$

$$= (1 - z^{-1})[1 + N(1 - q)]$$

where we have once again set $z = 1$ in the polynomial in the bracket to preserve the gain of the algorithm. Substitute into Eq. 15-2.9 to obtain an algorithm that matches the PID algorithm of Eq. 15-2.8.

Table 15-2.1 Tuning Formulas for PID Control Algorithms

For continuous transfer function:

$$G_P(s) = \frac{Ke^{-t_0 s}}{(\tau_1 s + 1)(\tau_2 s + 1)}$$

Let

$$\beta_1 = e^{-T/\tau_1} \qquad \beta_2 = e^{-T/\tau_2} \qquad N = \frac{t_0}{T}$$

$$K_c = \frac{(1 - q)(\beta_1 - 2\beta_1\beta_2 + \beta_2)}{K(1 - \beta_1)(1 - \beta_2)[1 + N(1 - q)]}$$

$$\tau_I = T\frac{\beta_1 - 2\beta_1\beta_2 + \beta_2}{(1 - \beta_1)(1 - \beta_2)}$$

$$\tau_D = T\frac{\beta_1\beta_2}{\beta_1 - 2\beta_1\beta_2 + \beta_2}$$

For discrete transfer function:

$$HG_P(z) = \frac{(b_, + b_1 z^{-1} + b_2 z^{-2})z^{-N-1}}{1 - a_1 z^{-1} - a_2 z^{-2}}$$

$$K_c = \frac{(1 - q)(a_1 + 2a_2)}{(b_0 + b_, + b_2)[1 + N(1 - q)]}$$

$$\tau_I = T\frac{a_1 + 2a_2}{1 - a_, - a_2}$$

$$\tau_D = T\frac{-a_2}{a_, + 2a_2}$$

$$D(z) = \left(\frac{1 - q}{(b_0 + b_1 + b_2)[1 + N(1 - q)]}\right)\frac{1 \quad a_1 z^{-1} \quad a_2 z^{-2}}{1 - z^{-1}} \qquad \textbf{(152.10)}$$

where the term in parentheses is a constant multiplier, Finally, the tuning formulas are
obtained by equating coefficients in the numerators of Eqs. 15-2.8 and 15-2.10. The
results are given in Table 15-2.1 as two sets of formulas. The first set gives the tuning
parameters as functions of the continuous model parameters, K, τ_1, τ_2, and N, and the
second set in terms of the discrete parameters of the model, $a_,$, a_2, $b_,$, b_1, b_2, and N.
Note that neither τ_3 nor the fraction of the sample time in the dead time, m, appears in
the tuning formulas. This is because they affect only the individual values of b_0, b_1,
and b_2, not their sum. From Eq. 15-2.6, the sum is equal to

$$b_0 + b_1 + b_2 = K(1 - a, - a_2) \tag{15-2.11}$$

Because the fraction of the sample time is not a part of the tuning formulas, the dead time appears only in parameter N, which is theoretically the integer number of samples contained in the dead time. The question arises of which value of N to use when the fraction of the sample time is greater than 0.5. For example, if the dead time is 1.9 samples, should we round to N = 2 or use N = 1 as the theory calls for? The answer is to round, because this will be more representative of the true dead time. An alternative is to use the fractional value, N = 1.9, and this is acceptable also. Note that the dead time affects only the gain of the algorithm.

When only a first-order-plus-dead-time model is available, the tuning formulas of Table 15-2.1 can be used by setting $\tau_2 = 0$ or a, = 0. This results in $\tau_D = 0$, which means the control algorithm becomes PI for first-order systems.

Adjustable Parameter q. The adjustable parameter q, like the closed-loop time constant τ_c in the synthesis formulas of Section 7-4, makes possible the adjustment of the closed-loop response to the requirements of the process. From the formulas of Table 15-2.1, we see that q affects only the controller gain, but its effect depends on the dead-time parameter N. For large values of N (fast sampling of a slow process), the value of q has little effect on the controller gain. When $q = 0$, the controller gain is the maximum predicted by the tuning formulas, and the closed-loop response is said to be a **deadbeat** response. From Eq. 15-2.3, the deadbeat response is

$$\frac{C(z)}{R(z)} = z^{-N-1} \tag{15-2.12}$$

In other words, the output is forced to the set point at the earliest possible sample instant, which is 1 more than the number of samples in the dead time, and kept there. The deadbeat response is plotted in Fig. 15-2.2a. The response with $0 \le q < 1$, shown in Fig. 15-2.2b, exponentially approaches a sustained change in set point. The figure shows that parameter q is the fraction of the current error that is left after the required minimum number of samples (N + 1). As q approaches unity, the controller gain is reduced to zero. The larger q, the smaller the gain and the more samples it takes for the error to be eliminated.

EXAMPLE 15-2.1

Determine the tuning parameters from the synthesis controller for the deadbeat response and for the response with $q = 0.60$, for a first-order process:

$$G_P(s) = \frac{K}{\tau s + 1}$$

The sample time is 0.1τ, where τ is the time constant, and K is the gain.
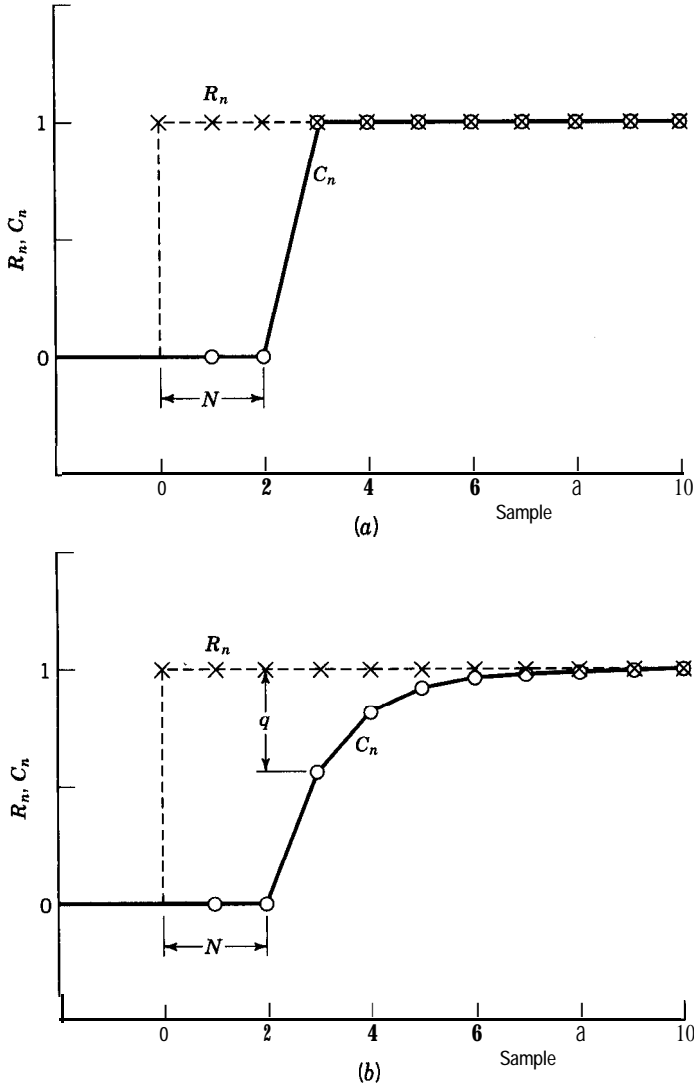
Figure 152.2 Closed-loop response specifications. (a) Deadbeat:
$q = 0$. (b) Dahlin: $0 \le q < 1$.

SOLUTION

Comparing with the model transfer function, Eq. 15-2.5, we have $\tau_1 = \tau$, $\tau_2 = 0$,
$\tau_3 = 0$, and $t_0 = 0$. Thus, from Table 15-2.1, $\beta_1 = e^{-T/\tau} = e^{-0.1} = 0.905$, $\beta_2 = 0$, $N = 0$, and the controller gain is

$$K_c = \frac{(1 - q)[0.905 - 2(0) + 0]}{K(1 - 0.905)(1 - 0)[1 + 0]} = \frac{9.5(1 - q)}{K}$$

For deadbeat response, $q = 0$, and $K_c = 9.5/K$ %CO/%TO, whereas for $q = 0.6$, $K_c = 9.5(1 - 0.6)/K = 3.8/K$ %CO/%TO.

The integral and derivative times are independent of q.

$$\tau_I = T\frac{0.905}{(1 - 0.905)(1 - 0)} = 9.5T \qquad \tau_D = T\frac{(0.905)(0)}{0.905} = 0$$

The closed-loop responses to a step change in set point are plotted in Fig. 15-2.3 for the case $K = 1$. Notice that for $q = 0$ (Fig. 15-2.3a), the controller output is switched twice: first to a large value (greater than 10 times the steady-state change) and then to the steady-state value. The controller can do this because it is designed for the exact gain and time constant of the process. It is not hard to see that such a precise change

(a)

(b)
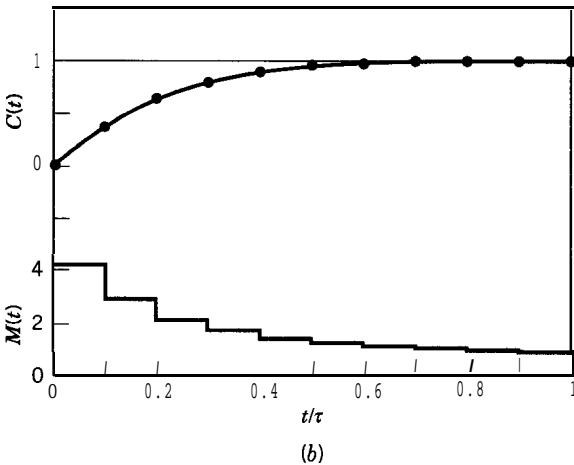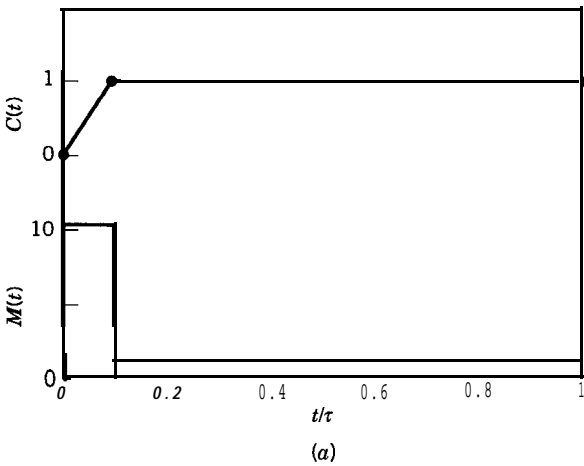
Figure 152.3 Responses to a unit step change in set point for the first-order process of Example 15-2.1. (a) Deadbeat response, $q = 0$. (b) Dahlin response, $q = 0.60$.

might cause trouble if the process gain or time constant were to change because of nonlinearity. For $q = 0.6$ (Fig. 15-2.3b), the controller gain and the initial output change are reduced to 40% of those for $q = 0$. This slows down the approach to set point but causes less upset of the process because of the smaller initial change in controller output. Because of the smaller gain, the controller is less sensitive to changes in the process gain and time constant. It will be **robust.**

The preceding example illustrates that in selecting the value of q, we must keep in mind that there is a compromise between the speed with which the error is eliminated and the movement in the controller output. The pulse in controller output observed in Fig. 15-2.3a can cause great upset to the process. Imagine, for example, that this controller output is connected to a fuel valve on a furnace. Such a large change in fuel flow can cause stack gas pollution (smoke) if it is an increase; the flame can go out if it is a decrease. The point is that in tuning the controller, we must consider both the speed of response of the controlled variable and reasonable movement of the controller output. We will discuss this in more detail in the next section.

EXAMPLE 15-2.2

The process of Example 7-4.2 is controlled with a computer with a sample time of 0.1 min. Determine the tuning parameters for the computer control algorithm. The process transfer function is

$$G_P(s) = \frac{1.0e^{-0.26s}}{s^2 + 4s + 1} = \frac{1.0e^{-0.26s}}{(3.73s + 1)(0.27s + 1)}$$

SOLUTION

The parameters are $K = 1.0$ %TO/%CO, $t_0 = 0.26$ min, $\tau_1 = 3.73$ min, $\tau_2 = 0.27$ min, and $T = 0.1$ min. The exponential terms are $\beta_1 = e^{-0.1/3.73} = 0.9736$, $\beta_2 = e^{-0.1/0.27} = 0.6905$, and $N = 0.26/0.1 = 2.6$. We will try all three possible values of N (2, 3, and 2.6) and compare the results. For the deadbeat response, $q = 0$, *the* three possible values of the gain are, by direct application of the formula in Table 15-2.1,

$$K_c = \frac{(1 - 0)[0.9736 - 2(0.9736)(0.6905) + 0.6905]}{(1.0)(1 - 0.9736)(1 - 0.6905)[1 + N]} \cdot \frac{3.03}{1 + N}$$

For the three different values of N, the gains are

N	2	2.6	3
K_c, %CO/%TO	13	11	9.8

All of these gains are high, and they should be reduced by adjusting the value of q. However, note that it all amounts to adjusting the gain to a value less than 10. Recall that in Example 7-4.2, for a PI controller, the controller gain varied between 3.6 and 6.3 %CO/%TO.

For the PID controller, the following integral and derivative times result from Table 15-2.1.

$$\tau_I = 0.1 \frac{0.9736 - 2(0.9736)(0.6905) + 0.6905}{(1 - 0.9736)(1 - 0.6905)} = 3.9 \text{ min}$$

$$\tau_D = 0.1 \frac{(0.9736)(0.6905)}{0.9736 - 2(0.9736)(0.6905) + 0.6905} = 0.21 \text{ min}$$

The derivative time is small relative to the integral time. This is because this is a very controllable process: the ratio of dead time to time constant is small (0.07). The integral time is close to the dominant time constant of the process.

EXAMPLE 15-2.3   *TUNING AN ANALYZER CONTROLLER*

The tuning formulas of Table 15-2.1 are valid for extreme ranges of the ratios of dead time to time constant and sample time to time constant. This makes them particularly useful for tuning analyzer controllers where the sample time is determined by the analysis cycle, such as in chromatographic analysis. One characteristic of sampling analyzers is that the results of the analysis are not available until the end of the cycle; this introduces a dead time of exactly one sample, $N = 1$. Determine the tuning parameters for an analyzer controller with an analysis cycle of 15 min if the process from which it takes the sample has a time constant of 3 min.

SOLUTION

The process parameters are $\beta_1 = e^{-15/3} = 0.00674$, $\beta_2 = 0$, and $N = 1$. For deadbeat response, $q = 0$, the formulas of Table 15-2.1 give

$$KK_c = \frac{0.00674}{(1 - 0.00674)(1 + 1)} = 0.00339$$

$$\tau_I = 15 \frac{0.00674}{1 - 0.00674} = 0.1018 \text{ min}$$

and $\tau_D = 0$. This is a very small loop gain and a very fast reset. To see better what kind of controller this is, let us substitute these values into the transfer function of the controller, Eq. 15-1.10.

$$\Delta M(z) = \frac{0.00339}{K}\left[1 - z^{-1} + \frac{15}{0.1018} + 0\right]E(z)$$

$$\doteq \frac{0.5}{K}E(z)$$

where we have neglected the term with multiplier 0.00339, which is small compared with the term with multiplier 0.5. This is a pure integral controller with a loop gain of 0.5 per sample. Perhaps it is easier to see the pure integral controller if we write the recursive formula:

$$M_n = M_{n-1} + \frac{0.5}{K}E_n$$

Why is the controller pure integral? Because the process is so fast that it reaches steady state between samples. Why is the controller incremented as it is? The increment in output required to eliminate the error completely is E_n/K. The increment is half of this, because the dead time of one sample requires a minimum of two samples to reach the desired steady state, so the correction is spread over two samples. This is exactly the correct strategy, and it follows directly from the tuning formulas.

152.2 Selection of the Sample Time

Selection of the sample time for computer control algorithms is an important consideration. Theoretically, the shorter the sample time, the better the performance of the loop. In practice, however, there is a point of diminishing returns that means that as the sample time is decreased, the same reduction in sample time results in smaller and smaller improvements in performance until the improvement becomes negligible. The cost of reducing the sample time is the overloading of the computer.

As we noted in Section 14-1, process control computers perform their control tasks at uniform periods of time that are multiples of their basic heartbeat, usually 1 s. There is a limit to how many tasks a computer can perform at each heartbeat. If all the control tasks are performed at the minimum sample time, one heartbeat, then the maximum number of tasks will be equal to the limit of tasks per heartbeat. But if tasks are performed at longer sample times (averaging, for example, 10 heartbeats), then the number of control tasks the computer can perform increases by a factor of 10, and so on. In other words, efficient utilization of process control computers requires that each task be performed at the longest sample time that produces acceptable performance. But how can we estimate what that sample time is?

Control performance depends on the relationship of the sample time to the other time parameters of the process-specifically, the dominant time constant and the dead time. Obviously, faster processes must be sampled more often than slow ones.

The tuning formulas of Table 15-2.1 can give us an idea how control performance is affected by the relationship among the sample time, the dominant time constant, and the dead time. Because these formulas can be extended to any ratio of these time parameters, they can tell us how the sample time affects the performance of the loop. For example, the formula for the controller gain is a function of the sample time. By

observing how the gain changes with sample time, we can infer the effect of the sample time on the performance, because the higher the controller gain, the better the controller can reject disturbances. Let us do this by assuming a first-order-plus-dead-time process, $\tau_2 = 0$, and let parameter $N = t_0/T$. Assuming the deadbeat response, $q = 0$, substitute into the formula for the gain in Table 152.1 and rearrange.

$$KK_c = \frac{\dfrac{T}{\tau} e^{-T/\tau}}{(1 - e^{-T/\tau}) \left(\dfrac{T}{\tau} + \dfrac{t_0}{\tau} \right)} \qquad (152.12)$$

We can now use this equation to do a parametric study of the deadbeat loop gain versus the two dimensionless parameters T/r and t_0/τ. The results are plotted in Fig. 15-2.4. We note that for $t_0/\tau > 0.4$, there is little increase in the loop gain when the sample time is decreased to less than one-tenth of the time constant. The higher t_0/τ, the higher the value of T/τ at which the point of diminishing returns is reached. For $t_0/\tau < 0.2$, the point of diminishing returns is reached at very small sample times (< 0.01 r), but note that the loop gain at this point is getting very high, on the order of 5. Such high gains are seldom used in practice, because the controller becomes very sensitive to variations in process parameters due to nonlinearities. As a result, the following rule of thumb gives reasonable results:

> *For good control performance and efficient use of computer resources, the sample time should be about one-tenth of the dominant process time constant.*

Longer sample times can be used when the dead time is equal to or greater than the dominant time constant, and shorter sample times are indicated when the dead time is small and large loop gains are required.
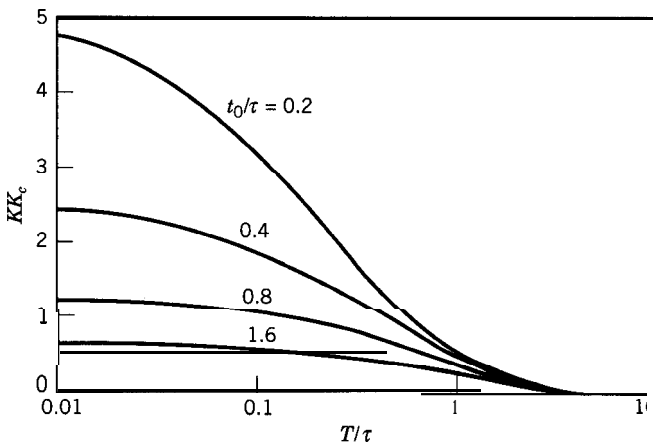
Figure 152.4 Effect of sample time on the gain of a PID controller tuned for deadbeat response.

15-3 FEEDBACK ALGORITHMS WITH DEAD-TIME COMPENSATION

Because of their ability to store and play back sampled values of the process variables, control computers are specially suited to compensating for process dead time. Feedback algorithms with dead-time compensation have been proposed by Dahlin (1968), Smith (1957), Vogel and Edgar (1980), and others. The Dahlin algorithm will be presented first, because it is the simplest and the most widely used in industry. The Smith Predictor, the first proposed dead-time compensation algorithm, will be shown to be essentially equivalent to the Dahlin algorithm. Finally, a more sophisticated algorithm, equivalent to the Vogel-Edgar algorithm, will be developed by the method of Internal Model Control (Garcia and Morari, 1982).

Dead-time compensation algorithms can greatly improve the performance of standard feedback controllers for processes with dead time equal to or greater than the dominant time constant. However, the satisfactory performance of dead-time compensation algorithms depends on having a good estimate of the process dead time. This is possible for some systems, such as paper machines and analyzers, but it is more difficult for systems in which the dead time is a strong function of throughput.

15-3.1 The Dahlin Algorithm

The first practical feedback algorithm with dead-time compensation was developed by Dahlin (1968) and independently by Higham (1968). The transfer function of the algorithm was synthesized in the preceding section on the way to developing the tuning formulas for the PID algorithm, Eq. 15-2.9.

$$D(z) = \frac{M(z)}{E(z)} = \left(\frac{1-q}{b_0 + b_1 + b_2}\right) \frac{1 - a_1 z^{-1} - a_2 z^{-2}}{1 - qz^{-1} - (1-q)z^{-N-1}} \qquad (15\text{-}2.9)$$

To obtain the recursive equation of the algorithm, replace the z-transforms with the sampled values, using the z variable as the shift operator. The result is

$$M_n = qM_{n-1} + (1-q)M_{n-N-1} + \frac{1-q}{b_0 + b_1 + b_2}(E_n - a_1 E_{n-1} - a_2 E_{n-2})$$

The incremental form of the algorithm is obtained by subtracting M_{n-1} and rearranging.

$$\Delta M_n = (1-q)(M_{n-N-1} - M_{n-1}) + \frac{1-q}{b_0 + b_1 + b_2}(E_n - a_1 E_{n-1} - a_2 E_{n-2}) \quad (15\text{-}3.1)$$

This is the Dahlin algorithm. The increment in controller output is the sum of two terms, the first of which requires the controller output M_{n-N-1} -that is, the output $(N + 1)$ samples ago. In a computer, this is done by storing at least $(N + 1)$ consecutive controller output values in a memory stack that is renewed at each execution of the algorithm so that old values are discarded as new ones are stored.

Note that the first term on the right-hand side of Eq. 15-3.1 is zero if N = O-that is, if the process dead time is less than the sample time. Evidently, it is this first term that provides the dead-time compensation.

The second term in Eq. 15-3.1 is a linear combination of the current and the previous

two values of the error, which is equivalent to the calculations of the PID algorithm, with the appropriate tuning relationships. This second term is best replaced with the full PID algorithm, as follows:

$$\Delta M_n = (1 - q)(M_{n-N-1} - M_{n-1}) + \Delta M_{\text{PID},n} \tag{153.2}$$

where $\Delta M_{\text{PID},n}$ is calculated by the standard PID algorithm, Eq. 15-1.11, tuned by the formulas of Table 15-2.1. By implementing the algorithm in this manner, we take full advantage of the features of the PID algorithm: the filter on the derivative and avoidance of taking the derivative of set point changes. When calculating the gain with the formula of Table 15-2.1, we must set $N = 0$, because the dead-time compensation term effectively eliminates the dead time of the process. Herein lies the benefit derived from the dead-time compensation term: it allows the controller gain to be greater by a factor of $[1 + N(1 - q)]$. For convenience, the entire algorithm, including the modified tuning formulas, is summarized in Table 15-3.1.

The term $\Delta M_{\text{PID},n}$ in Eq. 15-3.2 can also be calculated using the series form of the

Table 153.1 The Dahlin Dead-Time Compensation Algorithm

The algorithm, with filtered derivative on measurement:

$$Y_n = \frac{\alpha \tau_D}{\alpha \tau_D + T} Y_{n-1} - \frac{\tau_D}{\alpha \tau_D + T} (C_n - 2C_{n-1} + C_{n-2})$$

$$\text{AM}_n = (1 - q)(M_{n-N-1} - M_{n-1}) + K_c \left[E_n - E_{n-1} + \frac{T}{\tau_1} E_n + Y_n \right]$$

$$M_n = M_{n-1} + \Delta M_n$$

Tuning for the continuous transfer function:

$$G_P(s) = \frac{K e^{-t_{0}s}}{(\tau_1 s + 1)(\tau_2 s + 1)}$$

Let

$$\beta_1 = e^{-T/\tau_1} \qquad \beta_2 = e^{-T/\tau_2}$$

$$K_c = \frac{(1 - q)(\beta_1 - 2\beta_1\beta_2 + \beta_2)}{K(1 - \beta_1)(1 - \beta_2)}$$

$$\tau_I = T \frac{\beta_1 - 2\beta_1\beta_2 + \beta_2}{(1 - \beta_1)(1 - \beta_2)}$$

$$\tau_D = T \frac{\beta_1\beta_2}{\beta_1 - 2\beta_1\beta_2 + \beta_2}$$

PID algorithm, Eq. 15-1.14, provided that the tuning parameters of Table 15-2.1 are adjusted using Eq. 53.20.

EXAMPLE 15-3.1

For the analyzer controller of Example 15-2.3, use a Dahlin dead-time compensation algorithm and tune it for deadbeat response.

SOLUTION

The analyzer in Example 15-2.3 had a sample time of 15 min, a process time constant of 3 min, and one sample of dead time, $N = 1$. We found that because of the large ratio of the sample time to the time constant, the proportional term was negligible and a pure integral controller resulted.

$$M_n = M_{n-1} + \frac{K_c T}{\tau_I} E_n$$

We add the dead-time compensation term to this expression to obtain

$$M_n = M_{n-1} + \frac{K_c T}{\tau_I} E_n + (1 - q)(M_{n-N-1} - M_{n-1})$$

For the deadbeat response, $q = 0$ and $N = 1$,

$$M_n = M_{n-2} + \frac{K_c T}{\tau_I} E_n$$

Using the tuning formulas from Table 15-3.1, we get $K_c T/\tau_1 = 1/K$ and substitute.

$$M_n = M_{n-2} + \frac{1}{K} E_n$$

There are two differences between this algorithm and the result of Example 15-2.3. The increment is twice as large for this algorithm, and it is added to the output of two samples ago, not to the last output. This gives us some insight into how dead time compensation works. The output increment of E_n/K is exactly the one required to eliminate the error observed at any one sample, but the effect of this correction will not be detected until two samples later because of the one sample of dead time. If no additional disturbances upset the process, the error will be zero two samples later, and the output will not change anymore. Recall that because of the large ratio of sample time to time constant, the process reaches steady state between samples.

EXAMPLE 15-3.2

Compare the closed-loop responses of a PID algorithm and of a Dahlin dead-time compensation algorithm, with sample time $T = 1$ min, to a unit step change in set point. The transfer function of the process is

$$G_P(s) = \frac{1.2e^{-9.6s}}{(9s + 1)(5s + 1)}$$

SOLUTION

The PID algorithm is tuned with the formulas of Table 15-2.1, resulting in the following tuning parameters for $q = 0$, $N = 10$:

$$K_c = 0.99 \ \%CO/\%TO \qquad \tau_I = 13 \ \text{min} \qquad \tau_D = 3.0 \ \text{min}$$

The Dahlin algorithm is tuned with the formulas of Table 15-3.1, with $q = 0.80$. This results in $K_c = 2.2 \ \%CO/\%TO$, and the reset and derivative times are the same as for the PID controller.

The responses to a step change in set point were obtained with a computer program, DISCLOOP, available from the authors, and are plotted in Fig. 15-3.1. Note how the Dahlin algorithm makes a larger initial change in the output, which allows the controlled variable to reach the new set point faster and with a sharper response. The initial change in the output of the PID algorithm must be smaller than the one for the Dahlin algorithm because, as is evident in Fig. 15-3.1a, the integral mode continues to integrate during the dead time. This is because the error does not respond to the controller action during the period of the dead time. By using output values for up to one dead time back, the Dahlin algorithm waits for the error to respond before it takes action again. The small corrections required after the period of one dead time are caused by removal of the ringing poles from the algorithm, which was done back in the development of Eq. 15-2.9.

The Dahlin algorithm tuned for deadbeat response, $q = 0$, with gain $K_c = 11 \ \%CO/\%TO$, is shown for comparison in Fig. 15-3.2. The large upsets, particularly in the controller output, that occur at a period equal to the dead time, are due to the fact that the algorithm does not exactly match the process because of the removal of the ringing poles of Eq. 15-2.7. If these poles are not removed, the controlled variable will follow the deadbeat response exactly at the expense of ringing in the controller output. Note that during each period of one dead time, the controller output converges on the steady-state value. Surprisingly, this eventually allows the controlled variable to settle on the new set point.

153.2 The Smith Predictor

Smith (1957) proposed the idea of dead-time compensation before process control computers were available to carry it out. The scheme, known as the Smith predictor, is presented in block diagram form in Fig. 15-3.3. The block labeled $HG_m(z)$ is a model
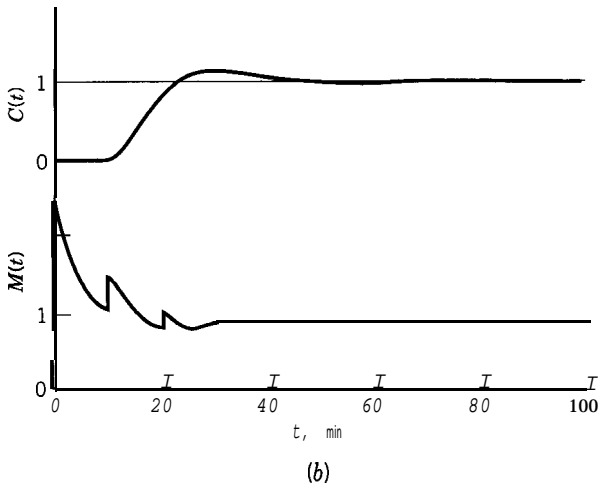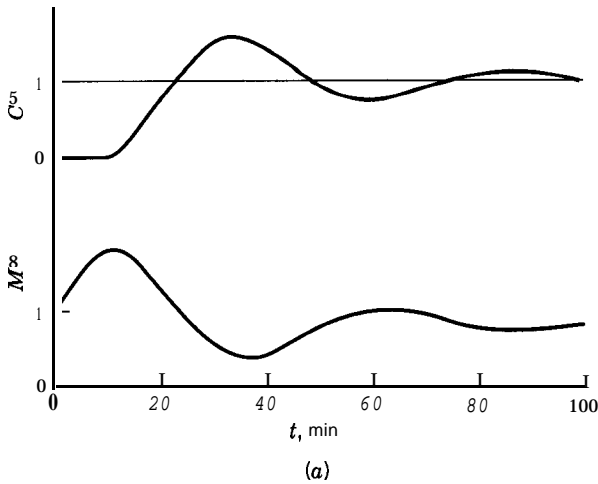
Figure 153.1 Closed-loop responses to step change in set point for Example 15-3.2. (a) PID algorithm. *(b)* Dahlin dead-time compensation algorithm.

of the process that excludes the process dead time of N samples. The output of this block is fed back to the controller, D'(z), which then would be controlling the process without dead time. To correct for model error and unmeasured disturbances, the output of the model is delayed by N samples and subtracted from the actual controlled variable, C(z), and the difference is added to the output of the model. In the absence of model error and disturbances, this difference will be zero.

The controller is made up of the three blocks: feedback control algorithm D'(z), process model $HG_m(z)$, and delay of N samples. Using block diagram algebra, we find that the controller output in Fig. 15-3.3 is

$$M(z) = D'(z)\{R(z) - [HG_m(z)M(z) + C(z) - HG_m(z)z^{-N}M(z)]\}$$
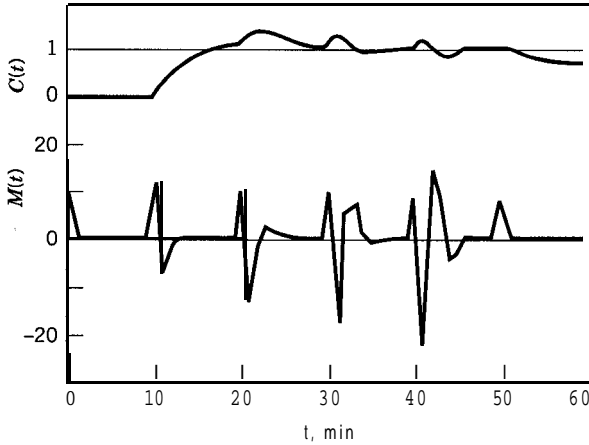
Figure 15-3.2 Deadbeat response ($q = 0$) of Dahlin dead-time compensation algorithm for the process of Example 15-3.2.

Solving for $M(z)$ yields

$$M(z) = \frac{D'(z)[R(z) - C(z)]}{1 + D'(z)HG_m(z)(1 - z^{-N})}$$

and the actual control algorithm is

$$D(z) = \frac{M(z)}{E(z)} = \frac{D'(z)}{1 + D'(z)HG_m(z)(1 - z^{-N})} \tag{15-3.3}$$

where $E(z) = R(z) - C(z)$ is the actual error. In this formula for the Smith predictor, the control algorithm $D'(z)$ is a PID algorithm that must be tuned, but the parameters of the model transfer function, $HG_m(z)$, are not used in the tuning. This makes the Smith predictor difficult to use because it requires a full model of the process, including the dead time N, as well as the tuning parameters of the controller.
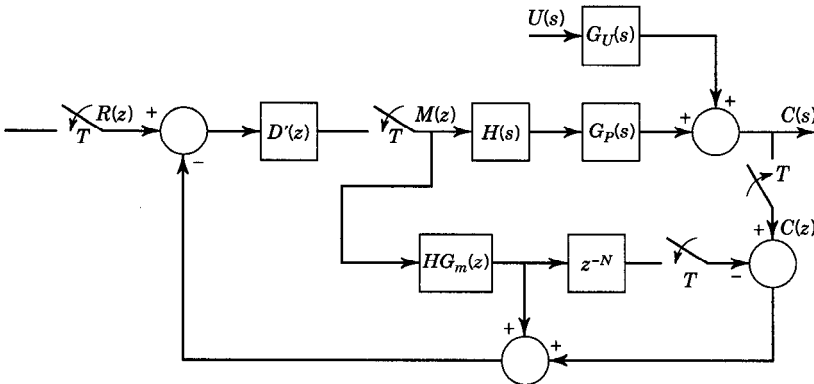
Figure 153.3 Block diagram of Smith predictor.

However, if we use the synthesis formula for the Dahlin response, Eq. 15-2.4, to develop controller D'(z), then we obtain

$$D'(z) = \frac{1}{HG_m(z)} \frac{(1 - q)z^{-1}}{1 - z^{-1}}$$

where the algorithm has been designed for the model, which does not contain dead time ($N = 0$). Substitute into Eq. 15-3.3 and simplify to obtain

$$D(z) = \frac{1}{HG_m(z)} \frac{(1 - q)z^{-1}}{1 - z^{-1} + (1 - q)z^{-1}(1 - z^{-N})}$$

$$= \frac{1}{HG_m(z)z^{-N}} \frac{(1 - q)z^{-N-1}}{1 - qz^{-1} - (1 - q)z^{-N-1}}$$

where we have multiplied numerator and denominator by z^{-N} to show explicitly that the Smith predictor can be reduced to a form identical to the Dahlin algorithm, Eq. 152.4, synthesized for the actual model of the process, including the dead time, $HG_m(z)z^{-N}$.

15-3.3   Algorithm Design by Internal Model Control

Garcia and Morari (1982) proposed the method of Internal Model Control (IMC) to synthesize and implement feedback control algorithms. In this section we will develop a dead-time compensation algorithm by the IMC technique. The algorithm coincides with the Dahlin algorithm when the process model is first-order with the dead time equal to an exact multiple of the sample time. For other cases, the IMC algorithm provides a more precise compensation for the process dynamics.

Figure 15-3.4 presents the basic block diagram of the IMC control scheme. The controller is made up of three blocks: a filter block $F(z)$, a compensator block $G_c(z)$, and a process model block $HG_m(z)$. In this scheme the model block includes the process
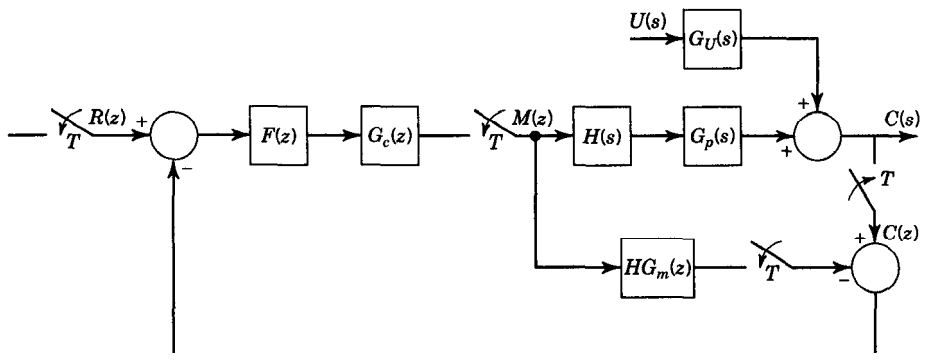
Figure 153.4 Block diagram of Internal Model Control (IMC) scheme.

dead time. Designing the IMC controller consists of selecting the filter and the compensator transfer functions. The filter is usually the simple exponential filter:

$$F(z) = \frac{1 - q}{1 - qz^{-1}} \tag{15-3.4}$$

where q is the adjustable filter constant. The compensator transfer function G,(z) is selected under the following restrictions:

1. The gain of G,(z) must be the reciprocal of the gain of the process model.

$$\lim_{z \to 1} G_c(z) = \lim_{z \to 1} \frac{1}{HG_m(z)} \tag{15-3.5}$$

2. The transfer function G,(z) is equal to the reciprocal of the invertible part of the process model $HG_m(z)$. "Invertible" means that nonminimal phase and other undesirable terms must be left out of the transfer function when it is inverted. The dead time is an example of a nonminimal phase term, and zeros with negative real parts are undesirable because they cause ringing and even instability in the controller output.

The first requirement ensures that there will be no offset. This can be shown by writing the closed-loop transfer function of the block diagram of Fig. 15-3.4.

$$C(z) = \frac{HG_v(z)G_c(z)F(z)}{1 + G_c(z)F(z)[HG_P(z) - HG_m(z)]} R(z) + \frac{UG_U(z)[\, 1 - G_c(z)F(z)HG_m(z)]}{1 + G_c(z)F(z)[HG_P(z) - HG_m(z)]}$$

To obtain the offset, substitute z = 1 and Eq. 15-3.5, and note that $F(1) = 1$.

$$\lim_{z \to 1} C(z) = \frac{HG_P(1)G_c(1)F(1)}{1 + G_c(1)F(1)HG_P(1) - \dfrac{1}{HG_m(1)}F(1)HG_m(1)} \lim_{z \to 1} R(z)$$

$$+ \frac{1 - \underset{m}{\&} \quad F(1)HG_m(1)}{1 + G_c(1)F(1)HG_P(1) - \dfrac{1}{HG_m(1)}F(1)HG_m(1)} \lim_{z \to 1} UG_U(z)$$

Apply the final value theorem of z-transforms.

$$\lim_{n \to \infty} C(nT) = \frac{HG_P(1)G_c(1)}{1 + G_c(1)HG_P(1)} \lim_{n \to \infty} R(nT)$$

$$+ \frac{1 - 1}{1 + G_c(1)HG_P(1) - 1} \lim_{n \to \infty} UG_U(nT)$$

$$= \lim_{n \to \infty} R(nT)$$

This means that the output is equal to the set point at steady state even in the presence of a disturbance $U(s)$. Thus there is no offset if condition 1, Eq. 15-3.5, is satisfied.

We will now develop the IMC control algorithm for the second-order-plus-dead-time model we used before, Eq. 15-2.6, which is now our model of the process:

$$HG_m(z) = \frac{(b, + b_1 z^{-1} + b_2 z^{-2})z^{-N-1}}{1 - a_1 z^{-1} - a_2 z^{-2}} \tag{153.6}$$

The compensator block is the reciprocal of this transfer function, after the numerator polynomial is removed because it is known to cause ringing on the controller output, and the dead-time term is removed because it is a nonminimal phase term.

$$G_c(z) = \frac{1 - a_1 z^{-1} - a_2 z^{-2}}{b_0 + b_1 + b_2} \tag{15-3.7}$$

The denominator term is included to satisfy condition 1. Equations 15-3.4, 15-3.6, and 15-3.7, when arranged as in Fig. 15-3.4, constitute the IMC algorithm. However, for comparison with the Dahlin algorithm, we will next obtain the actual transfer function of the controller. The controller output in Fig. 15-3.4 is, by block diagram algebra,

$$M(z) = G_c(z)F(z)\{R(z) - [C(z) - HG_m(z)M(z)]\}$$

$$= \frac{G_c(z)F(z)}{1 - G_c(z)F(z)HG_m(z)}[R(z) - C(z)]$$

This gives us the actual control algorithm.

$$D(z) = \frac{M(z)}{E(z)} = \frac{G_c(z)F(z)}{1 - G_c(z)F(z)HG_m(z)} \tag{153.8}$$

where $E(z) = R(z) - C(z)$ is the actual error. Substitute Eqs. 15-3.4, 15-3.6 and 15-3.7 and simplify to obtain

$$D(z) = \frac{M(z)}{E(z)} = \frac{1 - q}{b_0 + b, + b_2} \frac{1 - a_1 z^{-1} - a_2 z^{-2}}{1 - qz^{-1} - (1 - q)\frac{(b, + b_1 z^{-1} + b_2 z^{-2})z^{-N-1}}{b_0 + b, + b_2}}$$

To obtain the recursive formula for the algorithm, substitute the sampled values for the z-transforms, treating the z variable as the shift operator.

$$M_n = qM_{n-1} + (1 - q)\frac{b_0 M_{n-N-1} + b_1 M_{n-N-2} + b_2 M_{n-N-3}}{b_0 + b_1 + b_2}$$

$$+ \frac{1 - q}{b_0 + b_1 + b_2}(E_n - a_1 E_{n-1} - a_2 E_{n-2})$$

Subtract $M_n -$, and rearrange for the incremental form.

$$\Delta M_n = (1 - q) \left[\frac{b_0 M_{n-N-1} + \mathbf{ML\text{-}2} + b_2 M_{n-N-3}}{b_0 + b_1 + b_2} - M_1 \right]$$

$$+ \frac{1 - q}{b_0 + b, + b_2} (E_n - a_1 E_{n-1} - a_2 E_{n-2}) \quad \textbf{(153.9)}$$

This is the IMC algorithm for a second-order-plus-dead-time process. Comparing it with the Dahlin algorithm, Eq. 15-3.1, we note that the last terms are identical. As we noted when discussing the Dahlin algorithm, this term is best calculated by the PID algorithm so that we can make use of its desirable features. The tuning relationships are exactly those of Table 15-3.1. The difference between the IMC and Dahlin algorithms is in the compensation term. In the IMC algorithm, the compensation term keeps the parameters b_0, b_1, and b_2 separate. Recall, from their definition in Eq. 15-2.6, that these parameters contain the effect of the numerator time constant τ_3 and of the fraction m of the sample time in the dead time. Therefore, when good estimates of these parameters are available, the IMC algorithm offers more precise compensation for the process dynamics. Unfortunately, as a result of process nonlinearities, the actual process dynamic parameters vary. Because of this, we cannot depend on the more precise compensation provided by the IMC algorithm to tighten the controller tuning. Once the tuning is loosened, by increasing the filter parameter q, the advantage of the more precise compensator is lost. The Dahlin algorithm provides a coarse but simple compensation for dead time, which is why it is more widely used in industry.

EXAMPLE 15-3.3

Obtain the response to a set point change of the process of Example 15-3.2, with an IMC control algorithm tuned for deadbeat response and for $q = 0.95$.

SOLUTION

The dead time is 9.6 min and the sample time 1 min, so N = 9 and m $= 1 - 0.6 =$ 0.4. From Eq. 15-2.6, for $\tau_1 = 9$ min, $\tau_2 = 5$ min, and $\tau_3 = 0$, the parameter values are

$$b_0 = 0.00205 \qquad b_1 = 0.01708 \qquad b_2 = 0.00374$$

and the algorithm has the form, from Eq. 15-3.9:

$$AM,, = (1 - q)[0.089 M_{n-N-1} + 0.747 M_{n-N-2} + 0.164 M_{n-N-3} - M_{n-1}]$$

$$+ \frac{1 - q}{0.0229} (E_n - 1.7136 E_{n-1} + 0.7326 E_{n-2})$$

The response to a set point change is shown in Fig. 15-3.5a for $q = 0$. Because the model matches the process perfectly, the controlled variable gets to the set point in the minimum number of samples, which for a second-order system is N + 2 (11 in this case), and stays there. This takes three switches of the controller output. In general, the
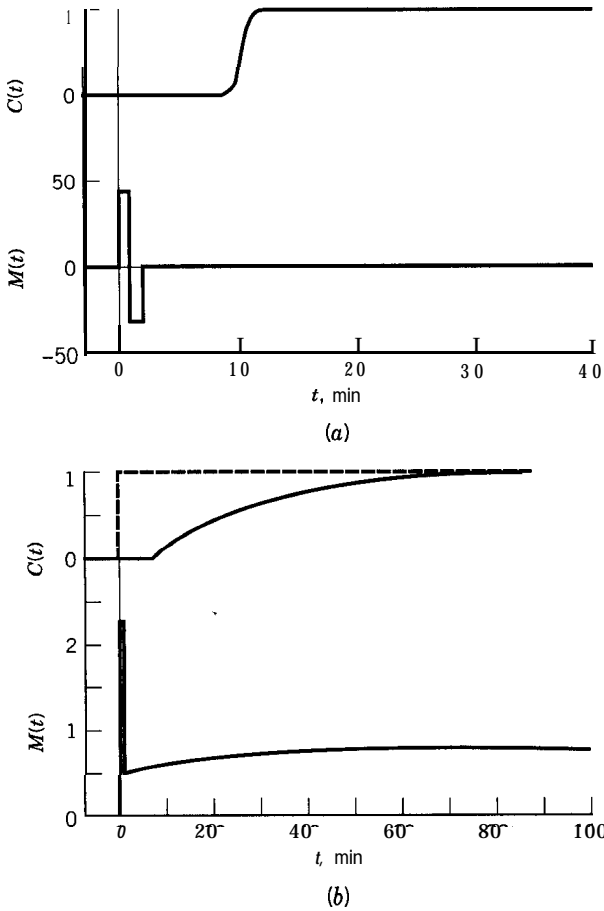
Figure 153.5 Response of IMC algorithm to step change in set point for Example 15-3.3. (a) With $q = 0$. (b) With $q = 0.95$.

minimum required number of switches of the controller output is 1 more than the order of the process. Note the large initial change in controller output-to 43%!—for the deadbeat response. This undesirable initial kick can be reduced by increasing the value of q to 0.95, where the value of q is selected to obtain an initial change in controller output similar to that obtained with the Dahlin algorithm in Example 15-3.2, Fig. 15-3.1b. The response of the de-tuned IMC algorithm is plotted in Fig. 15-3.5b. Comparing these two responses reveals that the Dahlin response reaches the set point faster. The reason is the initial pulse in the controller output for the IMC algorithm. This pulse is caused by the derivative action on the set point change. If the IMC algorithm were put in the form of the PID algorithm with the derivative acting on the measurement instead of on the error, then parameter q could be reduced to 0.80, and its response would be very similar to the response of the Dahlin algorithm.

The preceding example shows that the more precise dynamic compensation provided by the IMC algorithm can be advantageous if the large initial change in controller output

for set point changes can be tolerated. The size of this initial change can be reduced by increasing the ratio of the sample time to the process time constant and is not a problem for disturbance inputs. However, the tight tuning shown for the case $q = 0$ can be used only when the process model exactly matches the actual process dynamics, as in the preceding example. This, of course, is seldom possible because the process parameters change with operating conditions and time.

15-3.4 Selection of the Adjustable Parameter

Although the adjustable parameter q in the Dahlin and IMC algorithms is bracketed in the range $0 \leq q < 1$, its selection is difficult because the ratio of the sample time to the process time constant greatly affects the value of q that is appropriate for a given loop. A formula has been proposed to relate the value of parameter q to the sensitivity of the closed loop to set point changes (Hunter *et al.*, 1994). The idea is to specify the magnitude of the initial change in controller output on a set point change and calculate the value of q that corresponds to that ratio. To eliminate the process gain as a parameter, the initial change in controller output is ratioed to the steady-state change that is required by the change in set point.

From the formulas of Table 15-3.1, assuming that the derivative mode acts on the measurement and not on the error, the initial change in controller output on a set point change is caused by the proportional and integral terms.

$$\Delta M_0 = K_c \left[\Delta R - 0 + \frac{T}{\tau_I} \Delta R \right]$$

$$= \frac{(1 - q)(\beta_1 - 2\beta_1\beta_2 + \beta_2)}{K(1 - \beta_1)(1 - \beta_2)} \left[1 + \frac{(1 - \beta_1)(1 - \beta_2)}{\beta_1 \quad 2\beta_1\beta_2 + \beta_2} \right] \Delta R$$

$$= \frac{(1 - q)(1 - \beta_1\beta_2)}{K(1 - \beta_1)(1 - \beta_2)} \Delta R$$

where the β parameters are defined in Table 15-3.1. The final steady-state change in output caused by the set point change is

$$\Delta M_\infty = \frac{\Delta R}{K}$$

The ratio of these two changes is

$$\gamma = \frac{\Delta M_0}{\Delta M_\infty} = \frac{(1 - q)(1 - \beta_1\beta_2)}{(1 - \beta_1)(1 - \beta_2)} \tag{153.10}$$

We call γ the *set point sensitivity ratio*. Solve for q to obtain

$$q = 1 - \gamma \frac{(1 - \beta_1)(1 - \beta_2)}{1 \quad \beta_1\beta_2} \tag{153.11}$$

The usefulness of this formula derives from the fact that the value of γ is directly related to the tightness of the tuning. For example, $\gamma = 10$ obviously results in much tighter tuning than $\gamma = 1.5$. The value of γ must be less than a maximum value, because q cannot be negative. This maximum value is obtained by setting $q = 0$ in Eq. 15-3.10.

EXAMPLE 15-3.4

Calculate the values of the set point sensitivity ratio for the various values of q used in Examples 15-3.2 and 15-3.3.

SOLUTION

The values of the parameters for these examples are

$$\beta_1 = e^{-1/9} = 0.895 \qquad \beta_2 = e^{-1/5} = 0.819$$

In Example 15-3.2, the Dahlin algorithm was run with $q = 0.8$, for which the set point sensitivity ratio is

$$= \frac{(1 - 0.8)(1 - 0.895 \cdot 0.819)}{(1 - 0.895)(1 - 0.819)} = 2.81$$

The final steady-state change in output is $1/K = 1/1.2 = 0.833$ %CO. The predicted initial output change is then $(2.81)(0.833) = 2.3$ %CO. This agrees with the response of Fig. 15-3.1b.

In Example 15-3.2 we also ran the deadbeat response, $q = 0$, for which the set point sensitivity ratio is five times that for $q = 0.8$, or $\gamma = 14.0$. This agrees with the initial output change of $(14.0)(0.833) = 11.7$ %CO in Fig. 15-3.2.

The PID controller also used $q = 0$, but its gain was reduced by the factor $(1 + N)$, so its set point sensitivity ratio is $\gamma = 14.0/(1 + 10) = 1.3$. This causes an initial output change of a little over 1.0, as verified in Fig. 15-3.1a.

The IMC algorithm has a higher set point sensitivity ratio because the derivative mode also acts on the set point change. Because of this, the formula must be modified as follows (Hunter *et al.*, 1994):

$$\gamma = \frac{1 - q}{1 - a_1 - a_2} = \frac{1 - q}{(1 - \beta_1)(1 - \beta_2)}$$

In Example 15-3.3, the IMC algorithm was run with $q = 0$ and $q = 0.95$. From the foregoing formula, the corresponding sensitivity ratios are 52.5 and 2.6, respectively. This produces initial output changes of 43.7 and 2.2 %CO, which agree with the responses of Fig. 15-3.5.

This section has presented two popular dead-time compensation feedback control algorithms, the Dahlin and IMC algorithms. We also showed that the Smith predictor

can be reduced to the Dahlin algorithm when the feedback controller in the scheme is obtained by the Dahlin synthesis procedure.

15-4   AUTOMATIC   CONTROLLER   TUNING

One of the advantages of computer process control is that the computer has the ability to collect dynamic information from the process and compute the controller tuning parameters from these data. In this section we will outline the procedure for carrying out automatic tuning calculations. The calculations can be carried out in two different modes: auto-tuning and adaptive control. The auto-tuning mode consists of collecting the dynamic process data and running a program that calculates the controller tuning parameters. This operation is initiated by the control engineer, technician, or plant operator whenever it is desired to check whether process conditions have changed sufficiently to warrant a re-tuning of the controller. In the *adaptive* mode, the controller parameters are re-computed and re-adjusted at regular intervals of time without human intervention. In this manner the controller parameters are continuously adapted to the changes in the process dynamics. The motivation for either mode is that the process dynamics change with operating conditions and/or time because of the nonlinear nature of the process (see Chapters 3 and 4).

Automatic tuning requires identification of the process dynamics and adjustment of the tuning parameters to the identified process. Process identification consists of

1. Postulating a dynamic model of the process
2. Estimating the model parameters
3. Verifying that the model fits the process behavior

In selecting the process model for automatic tuning, we must ensure that the model can be fitted to the process response and that its parameters can be used to compute the tuning parameters of the controller we select. Assuming that the controller we want to auto-tune is the PID algorithm, we must select a model that is related to that algorithm. In Section 15-2 we presented the second-order-plus-dead-time (SOPDT) model, Eq. 15-2.5, as one that can fit a wide variety of process responses. Then we developed the tuning formulas of Table 15-2.1, which enable us to compute the PID algorithm parameters from the model parameters. The SOPDT model is therefore a very good candidate for auto-tuning the PID algorithm.

Because the computer collects data in the form of sampled values of the process input and output signals, the model must be put in a form that relates these data. The pulse transfer function of the SOPDT model with zero-order hold, Eq. 15-2.6, is therefore our starting point.

$$HG_P(z) = \frac{C(z)}{M(z)} = \frac{(b, + \ b_1 z^{-1} + \ b_2 z^{-2}) z^{-N-1}}{1 - a_1 z^{-1} - a_2 z^{-2}} \qquad \textbf{(152.6)}$$

To obtain an equation relating the sampled values, replace the z-transforms with the sampled values, using the z variable as the shift operator. After rearrangement,

$$C_n = \ a_1 C_{n-1} + \ a_2 C_{n-2} + \ b_0 M_{n-N-1} + \ b_1 M_{n-N-2} + \ b_2 M_{n-N-3} \qquad \textbf{(154.1)}$$

This is our process model. Note that the second set of tuning formulas of Table 15-2.1 can be used to calculate the PID tuning parameters directly from the parameters of this model.

The next step in the identification procedure is estimation of the parameters of the model. Equation 15-4.1 is linear in the parameters, and the coefficients of the parameters are the known sampled values of the input and output variables. This makes multiple linear regression an excellent candidate for the parameter estimation procedure. Computer programs such as the **MATLAB** System Identification Toolbox (Ljung, 1991) provide routines specifically designed for process identification. These routines, which are based on a model of the form of Eq. 15-4.1, perform sophisticated regression calculations and provide means for checking the validity of the model. The input to these routines consists of dynamic data in the form of the time series of sampled values of the pairs C_n-M_n, and the outputs are the values of the parameters, a,, a_2, b_0, b,, and b_2. Some routines also estimate the dead time N. For the routines that do not automatically estimate the dead time N, a value must be estimated and given to the routine. The model parameters are then used with the formulas of Table 15-2.1 to obtain the PID algorithm parameters, or with those of Table 15-3.1 for the parameters of the Dahlin dead-time compensation   algorithm.

For details on the theoretical background on process identification, see Box and Jenkins (1994) and Ljung (1987). For guidelines on the practical application of auto-tuning and adaptive techniques, see Åström and Wittenmark (1990) and Corripio (1990).

15-5 MODEL-REFERENCE CONTROL

The flexibility of process control computers has led to the development of controllers known generically as model-reference controllers. The idea is to use an on-line model of the process that can predict the future values of the process variable. These predicted values are then used in controlling the process. We discussed two such schemes in Section 15-3: the Smith predictor, and the Internal Model Control (IMC) algorithm, There we transformed the control calculations into a single feedback control algorithm that was found to be the PID algorithm with a dead-time compensation term. However, one of the main advantages of model-reference control schemes is that they can be extended to the control of multivariable systems. Several commercial schemes that fall into this category are Dynamic Matrix Control (DMC), Identification Command (IDCOM), RMPC, STAR, and Process Insights. For DMC, see Cutler and Ramaker, (1979). For the other schemes, see the References for this chapter. This section outlines the first of these schemes to provide an idea of what is involved in applying them.

Figure 15-5.1 presents a block diagram of the DMC control scheme for a single loop. In this diagram, the double lines represent vectors and the single lines represent scalar signals. Comparing the DMC algorithm with the Smith predictor of Fig. 15-3.3 and the IMC scheme of Fig. 15-3.4, we note that it is more like the former than the latter. This is because a model error is computed by subtracting the predicted model output, $C^*(z)$, from the measured variable $C(z)$. A correction based on this model error is then added to the model output vector $c^*(z)$ to obtain the corrected output vector $c+(z)$. By contrast, the IMC scheme simply computes the difference between the model output and the measured variable and adds that difference to the set point.

The DMC scheme models the process by a vector of sampled values of the unit step
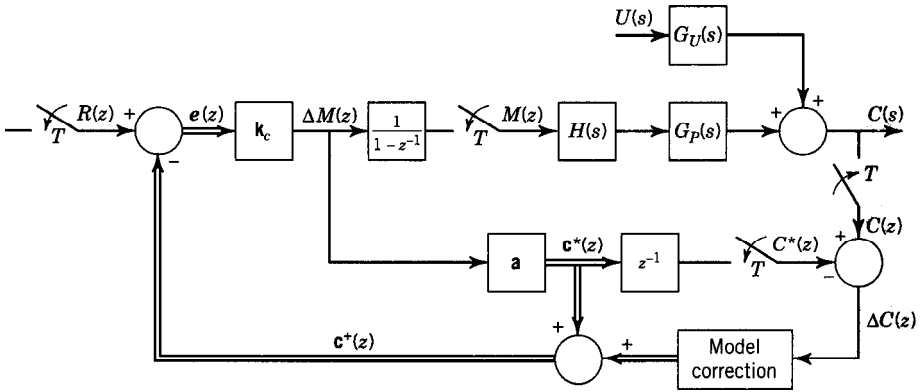
Figure **15-5.1** Block diagram for Dynamic Matrix Control (DMC).

response of the controlled variable. The size of this vector, n, is the response horizon. Typically it has a value of 30, although values of 60, 90, and 120 are also used. The elements of the response vector are usually obtained by regression of data obtained from open-loop process tests. Figure 15-5.2 shows a typical plot of the unit step response and the elements of the response vector.

If a pulse transfer function of the process is available, then the elements of the response vector can be computed by inverting the transform.

$$C(z) = HG_P(z) \frac{1}{1 - z^{-1}} = a_1 z^{-1} + a_2 z^{-2} + \cdots + a_n z^{-n}$$

where $HG_P(z)$ is the pulse transfer function, and a,, $a_{,,}$ . . . , a,, are the elements of the response vector. The inverse is calculated by long division, or with a spreadsheet, as discussed in Section 14-2. However, the main advantage of modeling the process as
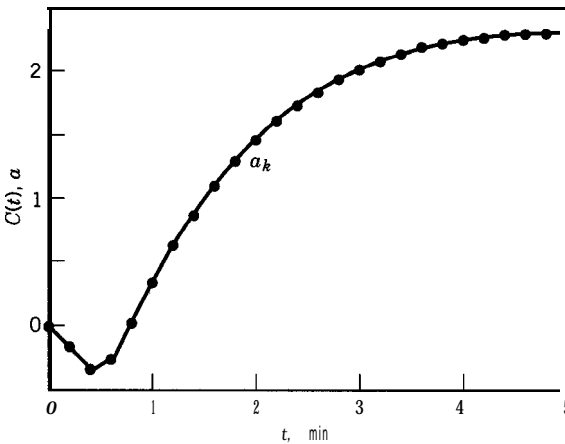
Figure 15-5.2 Unit step response showing the elements of the DMC response vector.

a vector of sampled values of the response is that it is not necessary to fit any particular pulse transfer function to the response.

The DMC scheme consists of using the response vector to estimate future values of the controlled variable, $c*(z)$ in Fig. 15-5.1. At each execution of the algorithm, the predicted value for the current sample is compared with the sampled value of the controlled variable, $C(z)$, and the resulting difference, which accounts for model error and unmeasured disturbances, is used to correct the future values predicted by the model. These corrected values, $c+(z)$, are then subtracted from the set point, $R(z)$, and the resulting vector of future errors, $e(z)$, is multiplied by the algorithm gain vector, $k,$, to compute the current increment in the manipulated variable, $\Delta M(z)$.

The theoretical background of the algorithm will be presented in the derivation of the formula to calculate the algorithm gain vector.

At any sampling instant, future values of the controlled variable are assumed to respond to the current and future increments in the output by following the principle of superposition. That is, each sampled output is the sum of the responses to each increment in the manipulated variable, each of which is a step starting at corresponding later samples, as in Fig. 15-5.3, where the individual responses to each increment in the controller output are shown along with their sum, which is the response. Let us call the current sample the zeroth sample.

$$C_1 = C_1^0 + a_1 \, \Delta M_0$$

$$C_2 = C_2^0 + a_2 \, \Delta M_0 + a, \; AM,$$

$$\textbf{(15-5.1)}$$

$$. \; . \; . \; .$$

$$C_n = C_n^0 + a_n \, \Delta M_0 + a_{n-1} \, \Delta M_1 + \cdots + a_{n-k+1} \, \Delta M_{k-1}$$

where C^0 is what the controlled variable would be if the current and future moves, and all disturbances, were zero. This assumes that the controlled variable continues to re-
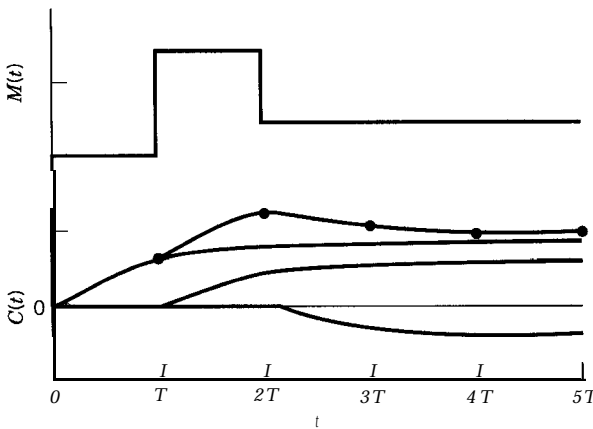
Figure 15-5.3 By the principle of superposition, the response is the sum of the step responses caused by every increment in controller output.

spond to past increments in the output. We consider only k future increments in the manipulated variable, where k is the **output** horizon. In the interest of compactness, let **us** express Eq. 15-5.1 in **matrix** notation.

$$\mathbf{c} = \mathbf{c}^0 + \mathbf{A}\,\Delta\mathbf{m} \qquad (15\text{-}5.2)$$

where **A** is the dynamic matrix and has the form

$$\mathbf{A} \doteq \begin{bmatrix} a_1 & 0 & 0 & . & . & . & 0 \\ a_2 & a, & \mathbf{0} & . & . & . & \mathbf{0} \\ . & & & & & & \\ a_n & a_{n-1} & a_{n-2} & \cdots & a_{n-k+1} \end{bmatrix}$$

The vector of predicted errors is obtained next.

$$\mathbf{e} = \mathbf{r} - \mathbf{c} = \mathbf{r} - \mathbf{c}^0 - \mathbf{A}\,\Delta\mathbf{m} = \mathbf{e}^0 - \mathbf{A}\,\Delta\mathbf{m} \qquad (15\text{-}5.3)$$

where $\mathbf{e}^0$ is the vector of future errors if the current and future increments in the manipulated variable are zero.

The control algorithm is designed so as to minimize the sum of the square of the error, plus a penalty term to prevent excessive increments in the manipulated variable. In vector form, the quadratic minimization problem is stated thus:

$$\min_{\Delta\mathbf{m}}(\mathbf{e}^T\mathbf{e} + \Delta\mathbf{m}^T\Lambda\,\mathbf{Am})$$

$$= \min_{\Delta\mathbf{m}}[(\mathbf{e}^0 - \mathbf{A}\,\Delta\mathbf{m})^T(\mathbf{e}^0 - \mathbf{A}\,\mathbf{Am}) + \Delta\mathbf{m}^T\Lambda\,\mathbf{Am}] \qquad (15\text{-}5.4)$$

where Λ is a diagonal matrix containing an adjustable scalar parameter A, the **move suppression parameter,** on each diagonal element:

$$A = \begin{bmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ . & . & . & . \\ 0 & 0 & \cdots & \lambda \end{bmatrix} \qquad (15\text{-}5.5)$$

The solution to Eq. 15-5.4 is

$$\mathbf{Am} = (\mathbf{A}^T\mathbf{A} + \Lambda)^{-1}\mathbf{A}^T\mathbf{e}^0 = \mathbf{K}_c\mathbf{e}^0 \qquad (15\text{-}5.6)$$

where $\mathbf{K}_c$ is the controller gain matrix, defined by

$$\boxed{\mathbf{K}_c = (\mathbf{A}^T\mathbf{A} + \Lambda)^{-1}\mathbf{A}^T} \qquad (15\text{-}5.7)$$

Equation 15-5.6 provides for calculation of the current and $(k - 1)$ future increments of the manipulated variable at each execution of the algorithm, but only the current

increment needs to be calculated. Therefore, only the first row of the gain matrix is used in the execution of the DMC scheme.

$$\Delta M_0 = K_{c_{11}} E_1^0 + K_{c_{12}} E_2^0 + \ldots + K_{c_{1n}} E_n^0$$

(15-5.8)

For a single loop, the tuning of the algorithm requires only adjustment of the move suppression parameter A. The larger λ, the smaller the gain, and therefore the longer it takes to eliminate the error.

EXAMPLE 15-5.1

Design and analyze the operation of a DMC control algorithm for the analyzer controller of Example 15-2.3.

SOLUTION

In Example 15-2.3, we found that the sample time of 15 min is so long relative to the process time constant of 3 min that the process reaches steady state between samples. Because the analyzer introduces one sample of dead time, the pulse transfer function of the process is

$$HG_P(z) = \frac{C(z)}{M(z)} = Kz^{-2}$$

where K is the process gain. The unit step response is

$$C(z) = Kz^{-2} \frac{1}{1 - z^{-1}} = Kz^{-2} + Kz^{-3} + Kz^{-4} + \cdots$$

The response vector need have only two elements $(n = 2)$, and we will use an output horizon of $k = 1$. The dynamic matrix is then equal to the response vector.

$$A = a = \begin{bmatrix} 0 \\ K \end{bmatrix}$$

Substitute into Eq. 15-5.7 to obtain the controller gain matrix.

$$K_c = (K^2 + \lambda)^{-1}[0 \ K] = \begin{bmatrix} 0 & \dfrac{K}{K^2 + \lambda} \end{bmatrix}$$

For a move suppression of $\lambda = 0$, the gain matrix is $[0 \ \ 1/K]$. When we use this gain vector to calculate the current increment in output, the result, from Eq. 15-5.6, is

$$\Delta M_0 = K_c e^0 = \begin{bmatrix} 0 & \dfrac{1}{K} \end{bmatrix} \begin{bmatrix} E_0^0 \\ E_1^0 \end{bmatrix} = 0 + \frac{E_1^0}{K}$$

As in Example 15-3.1, the increment in output is E/K, which is the value required to remove the error in one sample when the system reaches steady state between samples. However, the error that is used is not the current error E_0^0, which is multiplied by zero, but the predicted error one sample from now, E_1^0. In this way, the algorithm takes into account the dead time of one sample introduced by the analyzer.

Let us follow the calculations of the algorithm for a few samples when the set point is changed by 5 %TO with the error initially zero. At each sample, the error predictions are calculated using Eq. 15-5.1.

$$C_1 = C_1^0 + a_1 \, \Delta M_0 = C_1^0$$

$$C_2 = C_2^0 + a_2 \, \Delta M_0 + a_1 \, \Delta M_1 = C_2^0 + K \, \Delta M_0$$

where we have substituted $a_1 = 0$, $a_2 = K$. At time zero, the set point becomes $R = 5 \, \%TO$, and at each sample, $C_0^0 \leftarrow C_1$ and $C_1^0 \leftarrow C_2$ before the error is calculated. Then, the calculations are summarized in the following table.

R	C_1	C_2	E_0	E_1	ΔM_0
0	0	0	0	0	0
5	0	0	5	5	5/K
5	0	5	5	0	0
5	5	5	0	0	0

This assumes that the model is perfect so that, two samples after the change in set point, the feedback sample of the measured variable is indeed 5 %TO above its initial value. If this were not the case, the feedback error correction would modify the value of C_1 on the last line, creating an error that would then be corrected by the controller.

Extension to Multivariable Systems. The matrix operations on which the DMC scheme is based can easily be extended to any number of manipulated and controlled variables. For each manipulated variable-controlled-variable pair, a unit step response vector a is required. Each of these vectors is used to form a dynamic matrix A. The individual dynamic matrices then become submatrices of the dynamic matrix for the system. The vectors of future errors and manipulated variable increments are stacked in order, and the basic calculations of the gain matrix are essentially the same. A different move suppression parameter is required for each manipulated variable. Also, because the different controlled variables may not be of equal importance, the error of each variable is weighted by a different scaling factor.

EXAMPLE 15-5.2 *DMC ALGORITHM FOR TWO-POINT COMPOSITION CONTROL OF A DISTILLATION COLUMN*

In Example 13-2.4 we looked at the control of the product compositions of a benzene–toluene column by manipulating the reflux flow, w_R, and the heat rate to the reboiler, Q. Calculate the gain of a DMC control algorithm, assuming that the analyzers are so

slow that the composition reaches steady state between samples. Assume also that the analyzers introduce one sample of dead time on each composition measurement.

SOLUTION

The open-loop gains for the column were determined in Example 13-2.4. Using them, and representing each transfer function by one sample of dead time and no lag, we find that the transfer functions for the column are

$$Y_D(z) = -0.97z^{-2}W_R(z) + 5.88z^{-2}Q(z)$$

$$X_B(z) = 1.22z^{-2}W_R(z) - 7.73z^{-2}Q(z)$$

where Y_D is the distillate mole % toluene and X_B is the bottoms mole % benzene. As in Example 15-5.1, each response vector need have only two terms $(n = 2)$, the first of which is zero and the other equal to the corresponding gain. Here we will also use an output horizon of one move $(k = 1)$. The dynamic matrix and the move suppression matrix then look like this:

$$\mathbf{A} = \begin{bmatrix} 0 & 0 \\ -0.97 & 5.88 \\ \hline 0 & 0 \\ 1.22 & -7.73 \end{bmatrix} \qquad \Lambda = \begin{bmatrix} \lambda_1 & \\ & \lambda_1 \end{bmatrix}$$

To compute the gain matrix, Eq. 15-5.7 was programmed in MathCad (1994), with the following results. For no move suppression, $\lambda_1 = 0$, $\lambda_2 = 0$,

$$\mathbf{K}_c = \begin{bmatrix} 0 & -23.8 & 0 & -18.1 \\ \hline 0 & -3.76 & 0 & -2.99 \end{bmatrix}$$

These are very high gains, considering that the numbers have units of (klb/h)/mole % in the first row and (MBtu/h)/mole % in the second row. With move suppression parameters of $\lambda_1 = 0.01$, $\lambda_2 = 0.01$, the gain matrix is

$$\mathbf{K}_c = \begin{bmatrix} 0 & -2.35 & 0 & -1.77 \\ \hline 0 & -0.314 & 0 & -0.365 \end{bmatrix}$$

The gain matrix is multiplied by the vector of predicted errors, Eq. 15-5.6, which results in the following output increments:

$$\Delta w_R = -2.35E_{y_1}^0 - 1.77E_{x_1}^0$$

$$\Delta Q = -0.314E_{y_1}^0 - 0.365E_{x_1}^0$$

where the subscript y denotes the error in the distillate composition, x denotes the error in the bottoms composition, and the 1 denotes the predicted error one sample from now. As in Example 15-5.1, the zeros in the gain matrix make the algorithm ignore the current errors. This is to account for the one sample of dead time introduced by the analyzers. Note that each manipulated variable responds to both errors, which means that the DMC controller combines the decoupling and feedback control functions.

The DMC scheme has been presented here as an example of computer control techniques that are radically different from traditional control schemes. Many similar algorithms, listed at the beginning of this section, are also available. One of them, the Process Insights "Process Perfecter" (see the end of the References), is radically different from the others in that it uses nonlinear models developed with artificial neural networks.

15-6 SUMMARY

This chapter has presented control techniques that are commonly implemented on process control computers. A method of deriving recursive computing formulas from continuous transfer functions was used to derive filter, lead-lag, and feedback control algorithms. A second-order-plus-dead-time model was used to develop tuning relationships for the PID algorithm, two dead-time compensation algorithms, and a basis for automatic tuning of the PID algorithm. The techniques of Internal Model Control and Dynamic Matrix Control were also introduced.

REFERENCES

1. Åström, K. J., and B. Wittenmark. 1990. *Computer Controlled Systems-Theory and Design.* 2nd. ed. Englewood Cliffs, N.J.: Prentice-Hall.
2. Box, G. E. P., and G. M. Jenkins. 1994. *Time Series Analysis.* 3rd ed. San Francisco, Holden-Day.
3. Corripio, A. B. 1990. *Tuning of Industrial Control Systems.* Research Triangle Park, N.C.: Instrument Society of America.
4. Cutler, C. R., and B. L. Ramaker. 1979. "DMC-A Computer Control Algorithm." AIChE 1979 Houston Meeting, Paper No. 516, AIChE, New York, 1979. DMC, Inc., Houston, Tex.
5. Dahlin, E. B. 1968. "Designing and Tuning Digital Controllers." *Instrument & Control Systems* 41 (June):77.
6. Garcia, C. E., and M. Morari. 1982. "Internal Model Control. 1. A Unifying Review and Some Results." *Z&EC Process Des. & Dev.* 21:308–323.
7. Higham, J. D. 1968. "Single-Term Control of First- and Second-Order Processes with Dead Time." *Control,* February.
8. Hunter, S. R., S. M. Peebles, and A. B. Corripio. 1994. "Implementation of a Dynamically-Compensated PID Control Algorithm." *Computers & Chemical Engineering* 18(10):995-1000.
9. IDCOM (IDentification COMmand), SetPoint, Inc., Houston, Tex.
10. Ljung, L. 1987. *System Identification: Theory for the User.* Englewood Cliffs, N.J.: Prentice-Hall.
11. Ljung, L. 1991. *MATLAB System Identification Toolbox.* The MathWorks, Inc., Natick, Mass.

12. MathCad. 1994. Cambridge, Mass.: MathSoft, Inc.

13. Process Insights Process Perfecter, Pavilion Technologies, Austin, Tex.

14. RMPC, Honeywell, Inc., Phoenix, Ariz.

15. Smith, 0. J. M. 1957. "Closer Control of Loops with Dead Time." *Chemical Engineering Progress* 53 (May):217–219.

16. STAR, Dot Enterprises, Houston, Tex.

17. Vogel, E. F., and T. F. Edgar. 1980. "A New Dead Time Compensator for Digital Control." *Proceedings ISA/80,* ISA. Research Triangle Park, N.C.: Instrument Society of America.

PROBLEMS

15-1. A second-order filter has the continuous transfer function

$$\frac{Y(s)}{X(s)} = \frac{1}{\tau_F^2 s^2 + 2\tau_F s + 1}$$

Derive the recursive computing equation for the equivalent filter algorithm.

15-2. Do Problem 15-1 for the filtered derivative transfer function

$$\frac{Y(s)}{X(s)} = \frac{s}{\tau_F s + 1}$$

15-3. Show that if the forward difference approximation is used to approximate the s variable in the Laplace transforms of devices,

$$s = \frac{z - 1}{T}$$

the parameter of the exponential filter, Eq. 15-1.5, becomes

$$a = 1 - \frac{T}{\tau_F}$$

Find the range of values of the ratio τ_F/T for which the filter will have a negative pole, which will cause ringing. For what range of values of the ratio will the filter be unstable (pole < -1)?

15-4. Do Example 15-1.1 for a sample time of 0.5 min.

15-5. Compute and sketch a plot of the controller output M_n in Example 15-1.1 for at least 10 samples after the temperature changes, assuming the temperature remains constant after the change. Assume that the initial value of the controller output is 50% and that it is limited in the range 0 to 100%. This means that if the calculated value is less than zero it is set to zero, and if greater than 100%, it is set to 100%. Perform the calculation for both $\alpha = 0$ and $\alpha = 0.1$. *Hint:* The calculations can be easily programmed and plotted on a spreadsheet.

15-6. Derive Eq. 15-2.6 for the case in which the dead time is an integer number of sample times, $t_0 = NT$ (therefore A = 0). Obtain the formulas for the parameters a,, **a,**, b_0, b_1, and **b,**, and check that they agree with the formulas given in Eq. 15-2.6 when $m = 1$.

15-7. Derive Eq. 15-2.6 and the formulas for the parameters a,, a_2, b_0, b_1, and b_2 for an integrating process that has the transfer function

$$G_P(s) = \frac{K_I}{s(\tau s + 1)}$$

where K_I is the process gain in (% TO/min)/%CO.

15-8. Derive Eq. 15-2.6 and the formulas for the parameters a,, a_2, **b,**, **b,**, and b_2 for the case in which $\tau_1 = \tau_2 = \tau$.

15-9. Derive the tuning formulas of Table 15-2.1 by matching the coefficients of Eq. 15-2.10 with those of Eq. 15-2.8.

15-10. Tune a PID control algorithm using the formulas of Table 15-2.1 for the process of Problem 7-3 and a sample time of 0.1 min. Use the deadbeat response, $q = 0$.

15-11. Do Problem 15-10 for the moisture controller of the vacuum filter of Problem 7-15 with a sample time of 1.0 min.

15-12. Do Problem 15-10 for the concentration controller on the absorber of Problem 7-16. Assume that an analyzer with a sample time of 1 min is used to measure the composition and that the analyzer adds a dead time of one sample time to the process dead time.

15-13. Do Problem 15-10 for the furnace outlet temperature controller of Problem 7-17 with a sample time of 0.25 min.

15-14. Do Problem 15-10 for the temperature controller on the oil heater of Problem 6-24 if the sample time is one-tenth of the dominant time constant of the process.

15-15. Our friend Rogelio Piiieiro makes his living installing computer control systems in Louisiana sugar mills. He controls, among other things, the multiple-effect evaporators. Do Problem 15-10 for the double-effect evaporator of Problem 7-20 with a sample time of 1 s.

15-16. Do Problem 15-10 for the phosphate pebbles moisture controller of Problem 7-22 and a sample time of 1 s.

15-17. Study the sensitivity of the gain in Problem 15-10 to the sample time, using $T = 0.01$, 0.05, 0.1, 0.5, and 1 min.

15-18. Show that if we take limits as $T \rightarrow 0$ in the tuning formulas of Table 15-2.1, then the resulting tuning parameters are

$$K_c = \frac{\tau_1 + \tau_2}{Kt_0} \qquad \tau_I = \tau_1 + \tau_2 \qquad \tau_D = \frac{\tau_1 \tau_2}{\tau_1 + \tau_2}$$

Hint: First approximate the exponential terms by $e^{-T/\tau} \approx 1 - T/\tau$, which is valid when $T \ll \tau$. Then simplify as much as possible before taking limits.

15-19. Vogel and Edgar (1980) proposed the following modification of the Dahlin response, Eq. 15-2.3:

$$\frac{C(z)}{R(z)} = \frac{(1 - q)(b_0 + b_1 z^{-1} + b_2 z^{-2})z^{-N-1}}{(b_0 + b_1 + b_2)(1 - qz^{-1})}$$

Show that when this response and the second-order model, Eq. 15-2.6, are substituted into the synthesis formula, Eq. 15-2.2, the algorithm that results is identical to the IMC algorithm, Eq. 15-3.9.

15-20. Do Problem 15-10 assuming that a Dahlin dead-time compensation algorithm, Table 15-3.1, is to be used. Study the sensitivity of the gain to the sample time, using $T = 0.01$, 0.05, 0.1, 0.5, and 1 min.

15-21. Do Example 15-5.1 using an output horizon $k = 2$.

15-22. Do Example 15-5.1 if the process time constant is 15 min instead of 3 min. Use a response horizon of $n = 8$ and an output horizon of $k = 3$. *Note:* A computer program should be used to carry out the matrix inversion calculation.

15-23. The analyzer of Example 15-5.1 is installed on the blending tank of Fig. 13-1.1a, in which both the outlet composition and flow must be controlled by manipulating the two inlet flows. Because of the slow sampling of the analyzer, both the analyzer and the flow reach steady state between samples, but there is no dead time in the flow transmitter, only on the analyzer. Thus the transfer functions of the tank are

$$W(z) = 1.0z^{-1}M_1(z) + 1.0z^{-1}M_2(z)$$

$$X(z) = -0.5z^{-2}M_1(z) + 0.5z^{-2}M_2(z)$$

where W is the product flow, X is the product weight % solute, and the gains are from Example 13-5.3, in %TO/%CO. Design a DMC control algorithm and determine the control gain vectors.

Appendix A

Instrumentation Symbols and Labels

This appendix presents the symbols and labels used in this book for the instrumentation diagrams. Most companies have their own symbols and labels, and even though most of them are similar, they are not all identical. The symbols and labels used in this book follow closely the standard published by the Instrument Society of America (ISA); see the References. The appendix presents just the information needed for this book. For more information, see the ISA standard.

In general, the instrument identification, also referred to as tag number, is of the following form:

Typical Tag Number

LRC 101	Instrument identification or tag number
L 101	Loop identification
101	Loop number
LRC	Functional identification
L	First letter
RC	Succeeding letters

Expanded Tag Number

20-TAH-6A		Tag number
20		Optional prefix
	A	Optional suffix

Note: Hyphens are optional as separators.

The meanings of some identification letters are given in Table A-l.

Some symbols used in this book to designate the functions of computing blocks, or software, are presented in Table A-2. Table A-3 presents some instrument symbols, and Table A-4 presents some instrument line (signal) symbols.

Table A-l Meanings of Identification Letters (Courtesy of the Instrument Society of America)

	First Letter		Succeeding Letters		
	Measured or Initiating Variable	Modifier	Readout or Passive Function	output Function	Modifier
A	Analysis		Alarm		
B	Burner, combustion		User's choice	User's choice	User's choice
C	User's choice			Control	
D	User's choice	Differential			
E	Voltage		Sensor (primary element)		
F	Flow rate	Ratio (fraction)			
G	User's choice		Glass, viewing, device		
H	Hand				High
I	Current (electrical)		Indicate		
J	Power	scan			
K	Time, time schedule	Time rate of change		Control station	
L	Level		Light		Low

Letter	First-letter variable	Modifier	Readout/passive function	Output function	Modifier
M	User's choice	Momentary			Middle, intermediate
N	User's choice		User's choice	User's choice	User's choice
0	User's choice		Orifice, restriction		
P	Pressure, vacuum		Point (test) connection		
Q	Quantity	Integrate, totalize			
R	Radiation		Record		
S	Speed, frequency	Safety		Switch	
T	Temperature			Transmit	
U	Multivariable		Multifunction	Multifunction	Multifunction
V	Vibration, mechanical analysis			Valve, damper, louver	
W	Weight, force		Well		
X	Unclassified	X axis	Unclassified	Unclassified	Unclassified
Y	Event, state, or presence	Y axis		Relay, compute convert	
Z	Position, dimension	Z axis		Driver, actuator, unclassified final control element	

Table A-l Meanings of Identification Letters (Continued)

First Letters	Initiating or Measured Variable	Controllers			Self-Actuated Control Valves	Readout Devices		Switches and Alarm Devices			Transmitters			Solenoids, Relays, Computing Devices	Primary Element	Test Point	Well or Probe	Viewing Device, Glass	Safety Device	Final Element
		Recording	Indicating	Blind		Recording	Indicating	High	LOW	Comb.	Recording	Indicating	Blind							
A	Analysis	ARC	AIC	A C		A R	A I	ASH	ASL	ASHL	ART	AIT	A T	A Y	A E	AP	AW			A V
B	Burner combustion	BRC	BIC	B C		B R	BI	BSH	BSL	BSHL	BRT	BIT	B T	B Y	B E		BW	BG		B Z
C	User's choice																			
D	User's choice																			
E	Voltage	ERC	EIC	E C		E R	EI	ESH	ESL	ESHL	ERT	EIT	E T	E Y	E E					E z
F	Flow rate	FRC	FIC	FC	FCV, FICV	F R	FI	FSH	FSL	FSHL	FRT	FIT	FT	N	FE	FP		F G		Fv
FQ	Flow quantity	FQRC	FQIC			FQR	FQI	FQSH	FQSL			FQIT	FQT	FQY	FQE					FQV
FF	Flow ratio	FFRC	FFIC	FFC		FFR	FFI	FFSH	FFSL						FE					FFV
G	User's choice																			
H	Hand		HIC	H C						H S										H V
	Current	IRC	IIC			I R	II	ISH	ISL	ISHL	IRT	IIT	IT	I Y	I E					Iz
J	Power	JRC	IIC			JR	JI	JSH	JSL	JSHL	JRT	JIT	JT	J Y	JE					J V
K	Time	KRC	KIC	K C	KCV	K R	KI	KSH	KSL	KSHL	KRT	KIT	K T	K Y	K E					K V
L	Level	LRC	LIC	LC	LCV	L R	LI	LSH	LSL	LSHL	LRT	LIT	L T	L Y	L E		LW	LG		LV
M	User's choice																			
N	User's choice																			

O	User's choice																		
P	Pressure vacuum	PRC	PIC	PC	PCV	PR	PI	PSH	PSL	PSHL	PRT	PIT	IT	PY	PE	PP		PSV, PSE	PV
PD	Pressure differential	PDRC	PDIC	PDC	PDCV	PDR	PDI	PDSH	PDSL		PDRT	PDIT	PDT	PDY	PE	PP			PDV
Q	Quantity	QRC	QIC			QR	QI	QSH	QSL	QSHL	QRT	QIT	QT	QY	QE				QZ
R	Radiation	RRC	RIC	RC		RR	RI	RSH	RSL	RSHL	RRT	RIT	RT	RY	RE		RW		Rz
S	Speed frequency	SRC	SIC	SC	SCV	SR	SI	SSH	SSL	SSHL	SRT	SIT	ST	SY	SE				sv
T	Temperature	TRC	TIC	TC	TCV	TR	TI	TSH	TSL	TSHL	TRT	TIT	l-r	TY	TE	TP	TW	TSE	TV
TD	Temperature differential	TDRC	TDIC	TDC	TDCV	TDR	TDI	TDSH	TDSL		TDRT	TDIT	TDT	TDY	T E	TP	TW		TDV
u	Multivariable					UR	UI							UY					UV
v	Vibration machinery analysis					VR	VI	VSH	VSL	VSHL	VRT	VIT	VT	VY	VE				VZ
W	Weight force	WRC	WIC	WC	wcv	WR	WI	WSH	WSL	WSHL	WRT	WIT	WT	WY	WE				WZ
WD	Weight force, differential	WDRC	WDIC	WDC	WDCV	WDR	WDI	WDSH	WDSL		WDRT	WDIT	WDT	WDY	WE				WDZ
X	Unclassified																		
Y	Event state presence		YIC	YC		YR	YI	YSH	YSL				YT	YY	YE				YZ
Z	Position dimension	ZRC	ZIC	zc	zcv	ZR	ZI	ZSH	ZSL	ZSHL	ZRT	ZIT	ZT	ZY	ZE				ZV
ZD	Gauging deviation	ZDRC	ZDIC	ZDC	ZDCV	ZDR	ZDI	ZDSH	ZDSL		ZDRT	ZDIT	ZDT	ZDY	ZDE				ZDV

Table A-2 Function and Symbols of Computing Blocks or Software

Function	Symbol	Function	Symbol
Summation	Σ	Integral	$\int$
Multiplication	X or *	Division	$\div$
Square root	$\sqrt{\ }$	Function	f(x)
High selector	> or HS	Low selector	< or LS
High limiter	> or HL	Low limiter	< or LL
Bias	B_0	Lead-Lag	L/L

Figure A-l shows different ways to draw a control system, particularly a flow control loop. Figure A-la shows a flow element, FE-10, which is an orifice plate with flange taps, connected to an electronic flow transmitter, FT-10. The output of the transmitter goes to a square root extractor, FY-10A, and from here the signal goes to a flow-indicating controller, FIC-10. The output from the controller goes to an I/P transducer,

Computer-based algorithm

Analog instrument, accessible board-mounted

Analog instrument, mounted behind board

Pneumatic-operated globe valve

Pneumatic-operated butterfly valve, damper, or louver

Hand-actuated control valve

Control valve with positioner

Motor

Solenoid

Single-acting cylinder

Double-acting cylinder

Pressure-reducing regulator, self-contained

Back-pressure-reducing regulator, self-contained

Pressure relief or safety valve, angle pattern

Pressure-relief or safety valve, straight through pattern

Temperature regulator, filled-system type

Table A-3. General instrument symbols.

Three-way valve
FO to path A-C

Orifice plate with
flange or corner taps

Orifice plate with vena
contracta, radius, or
pipe taps

Orifice plate with vena
contracta, radius, or
pipe taps connected to
differential pressure
transmitter

Venturi tube or
flow nozzle

Turbine flowmeter

Magnetic flowmeter

Level transmitter, external
float or external type
displacer element

Level transmitter,
differential pressure
type element

Temperature element
without well

Temperature element with well

Table A-3. *(Continued)*

FY-10B, to convert the electrical signal to a pneumatic signal. The signal from the transducer then goes to a flow valve, FV-10. Often the labels for the flow element and valves are omitted for the sake of simplicity; the resulting diagram is shown in Fig. A-lb. The signals drawn in Fig. A-lb indicate that the control system used is electrical. Figure A-lc shows the control system when a computer control system is used; note the difference in signals. Figure A-ld shows the symbols used in this book. The figure shows the control concept without concern for specific hardware.

Pneumatic Signal		Electrical Signal		Generic Signal	
Software or Data Link		Mechanical Data Link		Electrical Binary Signal	

Table A-4. Instrument line (signal) symbols.
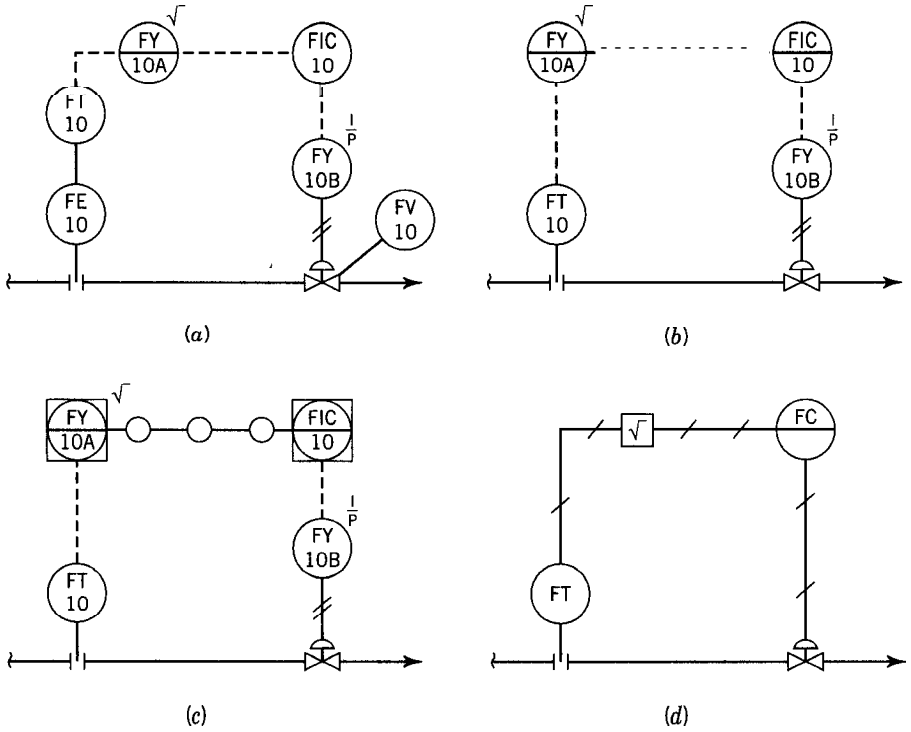
Figure A-l. Flow control system.

REFERENCES

1. "Instrumentation Symbols and Identification." Standard ISA-S5.1-1984. Research Triangle Park, N.C.: Instrument Society of America.

Appendix B

Case Studies

This appendix presents a series of design case studies that provide the reader with an opportunity to design process control systems from scratch. The first step in designing control systems for process plants is deciding which process variables must be controlled. This decision should be made by the process engineer who designed the process, the instrument or control engineer who will design the control system and specify the instrumentation, and the operating personnel who will run the process. This is certainly very challenging and requires team effort. The second step is the actual design of the control system. It is the second step that is the subject of these case studies; the first step has been done. Please note that, like any design problem, these problems are open-ended. That is, there are multiple correct answers.

Case 1. Ammonium Nitrate Prilling Plant Control System

Ammonium nitrate is a major fertilizer. The flow sheet in Fig. B-l shows the process for its manufacture. A weak solution of ammonium nitrate (NH_4NO_3) is pumped from a feed tank to an evaporator. At the top of the evaporator there is a steam ejector vacuum system. The air fed to the system controls the vacuum drawn. The concentrated solution is pumped to a surge tank and then fed into the top of a prilling tower. The development of this tower is one of the major postwar milestones in the fertilizer industry. In this tower the concentrated solution of NH_4NO_3 is dropped from the top against a strong updraft of air. The air is supplied by a blower at the bottom of the tower. The air chills the droplets in spherical form and removes part of the moisture, leaving damp pellets, or prills. The pellets are then conveyed to a rotary dryer where they are dried. They are then cooled, conveyed to a mixer for the addition of an antisticking agent (clay or diatomaceous earth), and bagged for shipping.

   A. Design the system to implement the following:
      1. Control the level in the evaporator.
      2. Control the pressure in the evaporator. This can be accomplished by manipulating the flow of air to the exit pipe of the evaporator.
      3. Control the level in the surge tank.
      4. Control the temperature of the dried pellets leaving the dryer.

707
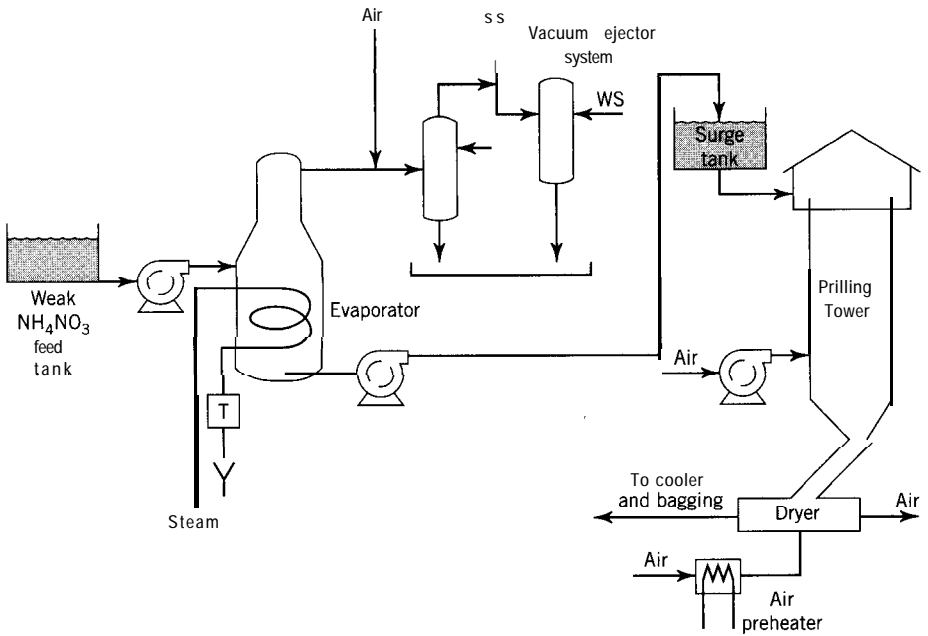
Figure B-l Ammonium nitrate prilling plant.

5. Control the density of the strong solution leaving the evaporator.
Be sure to specify the action of valves and controllers.
B. How would you control the production rate of this unit?
C. If the flow to the prilling tower varies often, it may also be desired to vary the air flow through the tower. How would you implement this?
D. A difficult loop to tune is the temperature loop of the dried pellets. Therefore, the following data were obtained by changing the temperature controller output by + 10%.

Time (min)	Temperature ("F)
0	200
1	200
3	202
5	208
6.5	214
8.5	218
11.0	220
13.0	222
15.0	221.8
17.0	223.0
∞	223.0

The temperature transmitter for this loop has a range of 100°F to 300°F. Tune a PI controller by the controller synthesis method and a PID controller by the minimum IAE method.

REFERENCES

1. The Foxboro Co. 1972. Application Engineering Data AED 288-3, January.

Case 2. Natural Gas Dehydration Control System

Consider the process shown in Fig. B-2. The process is used to dehydrate the natural gas entering the absorber by using a liquid dehydrant (glycol). The glycol enters the top of the absorber and flows down the tower countercurrent to the gas, picking up the moisture in the gas. From the absorber, the glycol flows through a cross-heat-exchanger into the stripper. In the reboiler, at the base of the stripper, the glycol is stripped of its moisture, which is boiled off as steam. This steam leaves the top of the stripper and is condensed and used for the water reflux. This water reflux is used to condense the glycol vapors, which might otherwise be exhausted along with the steam.

The process engineer who designed the process has decided that the following must be controlled:
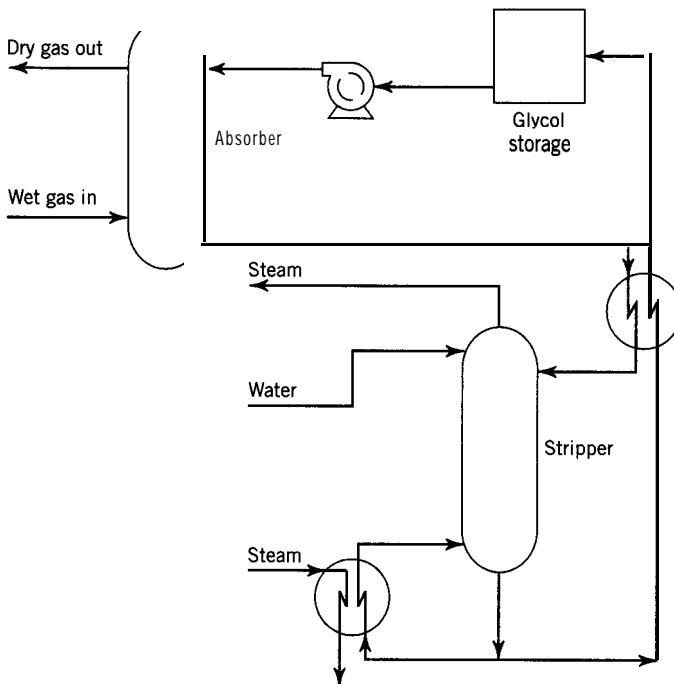
Figure B-2 Natural gas dehydration process.

1. The liquid level at the bottom of the absorber.
2. The water reflux into the stripper.
3. The pressure in the stripper.
4. The temperature in the top third of the stripper.
5. The liquid level at the bottom of the stripper.
6. Efficient absorber operation at various throughputs.

Design the control system to accomplish the desired control.

Case 3.   Sodium Hypochlorite Bleach Preparation Control System

Sodium hypochlorite ($NaOCl$) is formed by the following reaction:

$$2NaOH + Cl_2 \rightarrow NaOCl + H_2O + NaCl$$

The flow sheet in Fig. B-3 shows the process for its manufacture.

Dilute caustic ($NaOH$) is continuously prepared, to a set concentration (15% solution), by water dilution of a 50% caustic solution and stored in an intermediate tank. From this tank, the solution is pumped to the hypochlorite reactor. Chlorine gas is introduced into the reactor for the reaction.

A. Design the control system to accomplish the following:
   1. Control the level in the dilution tank.
   2. Control the dilution of the 50% caustic solution. The concentration of this stream is to be measured by a conductivity cell. When the dilution of this stream decreases, the output from this cell increases.
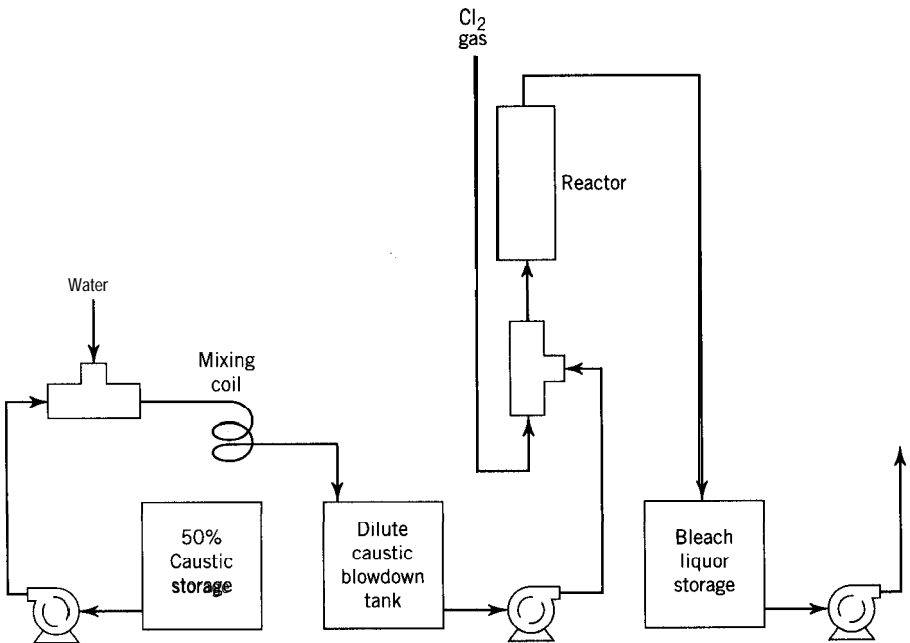
Figure B-3 Sodium hypochlorite bleach process.

3. Control the level in the bleach liquor storage tank.
4. Control the ratio of excess NaOH to available Cl, in the outlet stream from the hypochlorite reactor. This ratio is measured by an ORP (oxidation–reduction potential) technique. As the ratio increases, the ORP signal also increases.

Specify the action of valves and controllers. Briefly discuss your design.

B. How would you set the production rate from this unit?

C. For safety reasons, when the flow of caustic solution from the dilute caustic tank to the reactor fails, the flow of chlorine must be stopped immediately. Design this scheme and explain it.

REFERENCES

1. The Foxboro *Co.* **1972.** *Application Engineering Data, January.*

Case 4. Control Systems in the Sugar Refining Process

The process units shown in Fig. B-4 form part of a process to refine sugar. Raw sugar is fed to the process through a screw conveyor. Water is sprayed over it to form a sugar syrup. The syrup is heated in the dilution tank. From the dilution tank the syrup flows to the preparation tank, where more heating and mixing are accomplished. From the
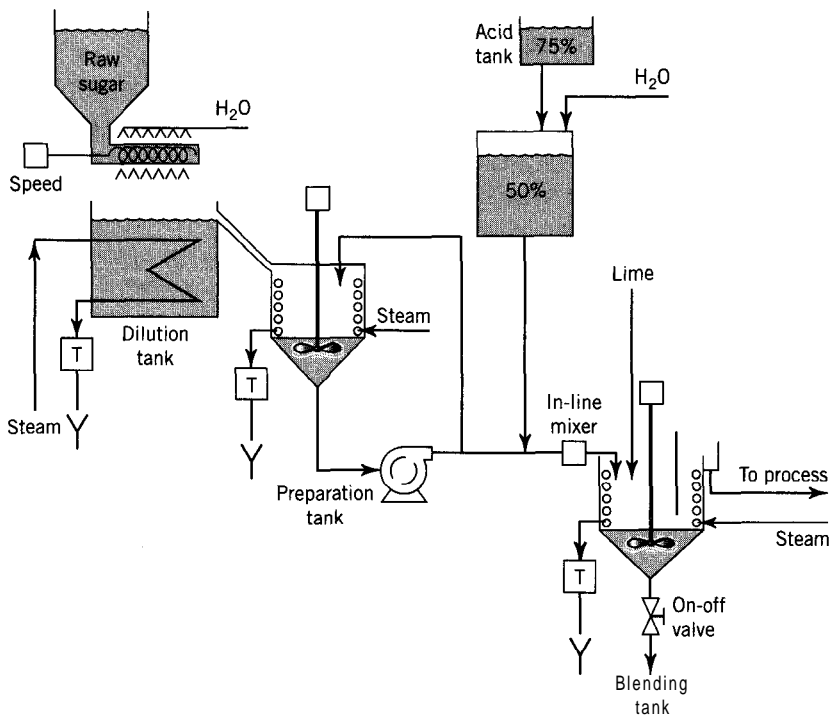
Figure B-4 Sugar refining process.

preparation tank the syrup flows to the blending tank. Phosphoric acid is added to the syrup as it flows to the blending tank. In the blending tank lime is added. This treatment with acid, lime, and heat serves two purposes. The first is that of clarification; that is, the treatment causes the coagulation and precipitation of the no-sugar organics. The second purpose is to eliminate the coloration of the raw sugar. From the blending tank the syrup continues to the process.

   A. The following variables are thought to be important to control:
      1. Temperature in the dilution tank.
      2. Temperature in the preparation tank.
      3. Density of the syrup leaving the preparation tank.
      4. Level in the preparation tank.
      5. Level in the 50% acid tank. The level in the 75% acid tank can be assumed constant.
      6. The strength of the 50% acid. The strength of the 75% acid can be assumed constant.
      7. The flow of syrup and 50% acid to the blending tank.
      8. The pH of the solution in the blending tank.
      9. Temperature in the blending tank.
      10. The blending tank requires only a high-level alarm.

      The flowmeters used in this process are magnetic flowmeters. The density unit used in the sugar industry is "Brix, which is roughly equivalent to the percentage of sugar solids in the solution by weight.

      Design the control systems necessary to control all of these variables. Show the action of control valves and controllers.
   B. How would you control the production rate?

Case 5. CO, Removal from Synthesis Gas

Consider the process shown in Fig. B-5 for removing CO, from synthesis gas. The plant treats 1646.12 MSCFH of feed gas. The feed gas will be supplied at 1526°F and 223 psig. The products from this plant will be synthesis gas at 115°F and 600 psig with a maximum of 50 volumetric ppm CO,, and CO, gas at 115°F and 325 psig.

   The process is as follows. The feed gas enters the plant at 1526°F and 223 psig. The gas must be cooled to 105°F before entering the absorber for CO_2 removal. This cooling is done in four stages. First, the feed gas passes through a superheater (E-15) and a boiler (E-14). The heat removed produces 27,320 lb/h medium-pressure steam. Second, the feed gas passes through an economizer (E-13), heating the demineralized water before deaeration. Third, the feed gas passes through a reboiler (E-l 1) where the feed gas provides 84% of the reboiler duty under full operation. Finally, the feed gas is cooled in the feed gas heat exchanger (E-12) by plant cooling water. By means of these four stages, 75% of the heat removed from the feed gas is recovered for process heating requirements.

   The cooled gases enter the absorber (C-6) in countercurrent fashion, where monoethanolamine (MEA) strips the CO, from the gases. The remaining gases are compressed in a two-stage centrifugal compression (B-1A and B). The compressor is driven by a steam turbine (M-l). Interstage cooling and exit cooling of the gases are provided (E-7A and B) with "knockout" drums (C-12A and B) to separate out any condensation.

The turbine operates on medium-pressure steam of which 68% is provided by the feed gas steam boiler. Ninety-three percent of the outlet low-pressure steam is available for plant use. The remaining 7% provides heat to one of the reboilers (E-10) of the regenerator column (C-7).

The CO, is carried with the MEA to the regenerator (C-7), where it is separated from the MEA. The regenerator is operated at low pressure and high temperature. This causes the CO, to be released with water vapor out the top of the tower while the lean MEA is recirculated (P-l) to the absorber. Before this MEA enters the top of the absorber, it passes through four heat exchangers (E-l, E-2, E-3, and E-4), where it is cross-exchanged with the bottoms of the absorber. These four exchangers recover 8.8 MM Btu/h. From these exchangers, the MEA passes through another cooler (E-5) and finally enters the absorber.

The CO, gases from the regenerator are compressed in a two-stage compressor (B-2 and 3) powered by electric motors. Interstage and exit coolers (E-9A and B) are provided, along with knockout drums (C-13A and B).

Table B-l gives the conditions of the streams numbered in the flow diagram. The process engineer believes that the following variables must be controlled:

1. Temperature of superheated steam leaving E-15.
2. Pressure of superheated steam produced in E-14/E-15.
3. Level in boiler drum.
4. Pressure in deaerator.
5. Level in deaerator.
6. Flow of low-pressure makeup steam to deaerator.
7. Temperature of feed gas leaving E-13 going to reboiler E-l 1.
8. Temperature of feed gas into absorber C-6.
9. Flow of MEA to absorber C-6.
10. Temperature in the bottom third of C-6.
11. Level in regenerator C-7.
12. Temperature in bottom of regenerator C-7.
13. Temperature of CO, gases leaving E-8.
14. Pressure in the regenerator area.
15. Temperature of synthesis gas between the two stages of compressor and exit temperature.
16. Pressure of the synthesis gas leaving the compressor B-l.
17. Interstage and exit temperature of CO_2 gases through compressor B-2.

These control loops may not all be necessary for smooth operation; however, they are the first ones proposed by the process engineer. You may propose your own. Design the control systems for the variables mentioned. Specify the fail-safe action of valves and the action of controllers.

Case 6. Sulfuric Acid Process

Figure B-6 shows a simplified flow diagram for the manufacture of sulfuric acid (H_2SO_4).

Sulfur is loaded into a melting tank, where it is kept in the liquid state. From this tank the sulfur goes to a burner, where it is reacted with the oxygen in the air to produce

Figure B-5 Process for CO, removal from synthesis gas.

E-13 Economizer
E-15 Superheater
E-3 Boiler feedwater pump
E-14 Boiler
C-8 Deareator

C-I Reformer gas knockout drum
C-6 MEA absorber

E-5 MEA solution cooler
C-2 Stripped gas knockout drum
E-7 A & B Synthesis gas cooler
E-I, E-2, E-3, E-4 MEA heat exchanger
B-1A & B Synthesis gas compressor

SO, by the reaction

$$S_{(l)} + O_{2(g)} \rightarrow SO_{2(g)}$$
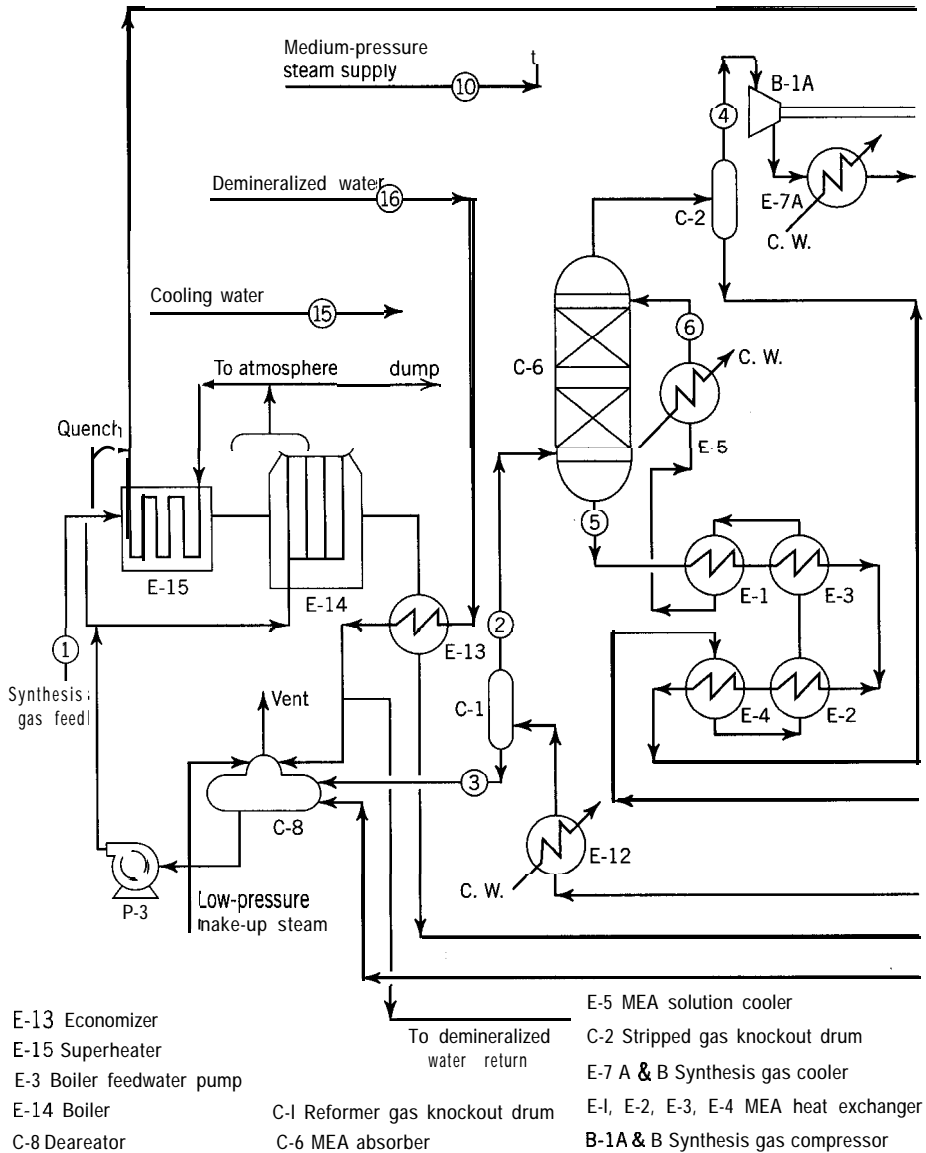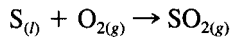
From the burner, the gases are passed through a waste-heat boiler, where the heat of reaction of the foregoing reaction is recovered by producing steam. From the boiler, the gases are then passed through a four-stage catalytic converter (reactor). In this converter the following reaction takes place:
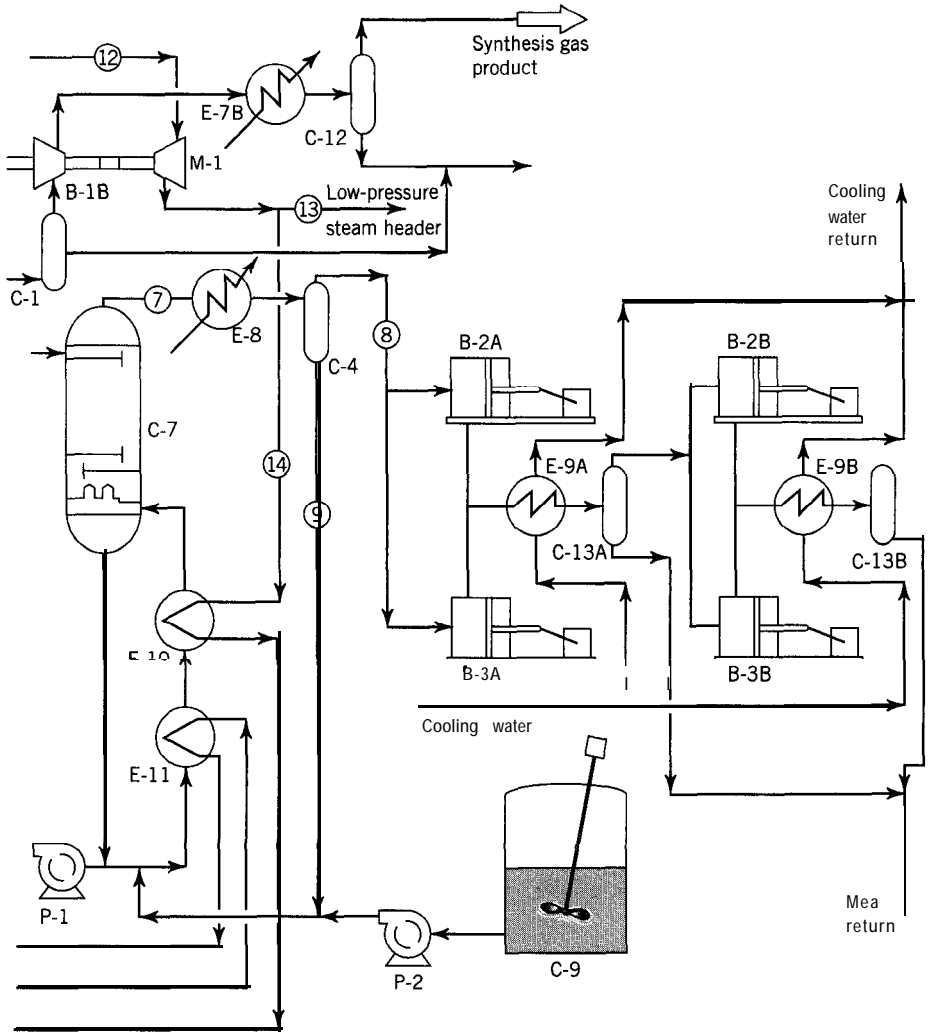
N-l Turbine	C-4 Acid gas knockout drum	E-9A & B CO, gas cooler
P-l MEA Recirculating pump	C-12 Synthesis gas knockout drum	C-13A & B CO, gas knockout drum
C-7 MEA Regenerator	B-2A & B CO, compressor	
E-10 MEA Reboiler	B-3A & B CO, spare compressor	
E-11 MEA Reboiler	C-9 Amine injection tank	
E-8 Acid gas cooler	P-2 Amine pump	

Figure B-5 *(Continued)*

$$SO_{2(g)} + \frac{1}{2} O_{2(g)} \rightleftarrows SO_{3(g)}$$
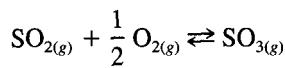
From the converter, the gases are sent to an absorber column, where the SO, gases are
absorbed by dilute H_2SO_4 (93%). The water in the dilute H_2SO_4 reacts with the SO,
gas, producing H_2SO_4

$$H_2O_{(l)} + SO_{3(g)} \rightarrow H_2SO_{4(l)}$$
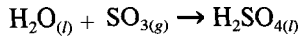
Table B-l Steady-State Conditions of the Streams in Fig. B-5

	Stream							
	1	2	3	4	5	6	7	8
		Absorber	Gas Knockout	Synthesis Gas to	Absorber		Stripper	co, to
Component	Feed Gas	Feed	Drum	Compressor	Bottoms	MEA Return	Overhead	Compressor
H$_2$, lbmol/h	1782.24	1782.08	0.16	1782.08				
N$_2$, lbmol/h	5.84	5.84	0.00	5.75	0.09		0.09	0.09
CO,, lbmol/h	198.42	197.90	0.52	0.12	217.24	19.45	200.78	197.78
CO, lbmol/h	352.28	352.25	0.03	346.39	5.86		5.86	5.86
CH,, lbmol/h	93.02	93.00	0.02	89.60	3.41		3.41	3.41
H$_2$O, lbmol/h	1112.40	13.08	1099.38	0.58	22.86	10.36	22.86	12.50
MEA solution, lbmol/h				0.21	2262.82	2263.03	77.90	0.19
Total mass flow, lbmol/h	3599.21	2444.16	1100.04	2224.74	2512.28	2292.84	310.91	219.83
Temperature, °F	1526	105	105	100	107	100	300	101
Pressure, psia	238	222	222	220	222	221	26	20

	Stream							
	9	10	11	12	13	14	15	16
	Stripper	Medium-	Medium-Pressure	Medium-Pressure	Low-	Low-		
	Knockout	Pressure	Steam from	Steam to	Pressure	Pressure	Cooling	Demineralized
Component	Drum Bottoms	Steam	Boiler	Turbine	Steam	Steam	Water	Water
H,, lbmol/h								
N,, lbmol/h								
CO,, lbmol/h	3.00							
CO, lbmol/h								
CH$_4$, lbmol/h								
H$_2$O, lbmol/h	10.36	706.02	1557.40	2363.42	1961.64	401.78	1252.57	1546.93
MEA solution, lbmol/h	77.71							
Total mass flow, lbmol/h	91.07	706.02	1557.40	2363.42	1961.64	401.78	1252.57	1546.93
Temperature, °F	101	600	600	600	435	435	86	60
Pressure, psia	20	415	415	400	201	201	43	115
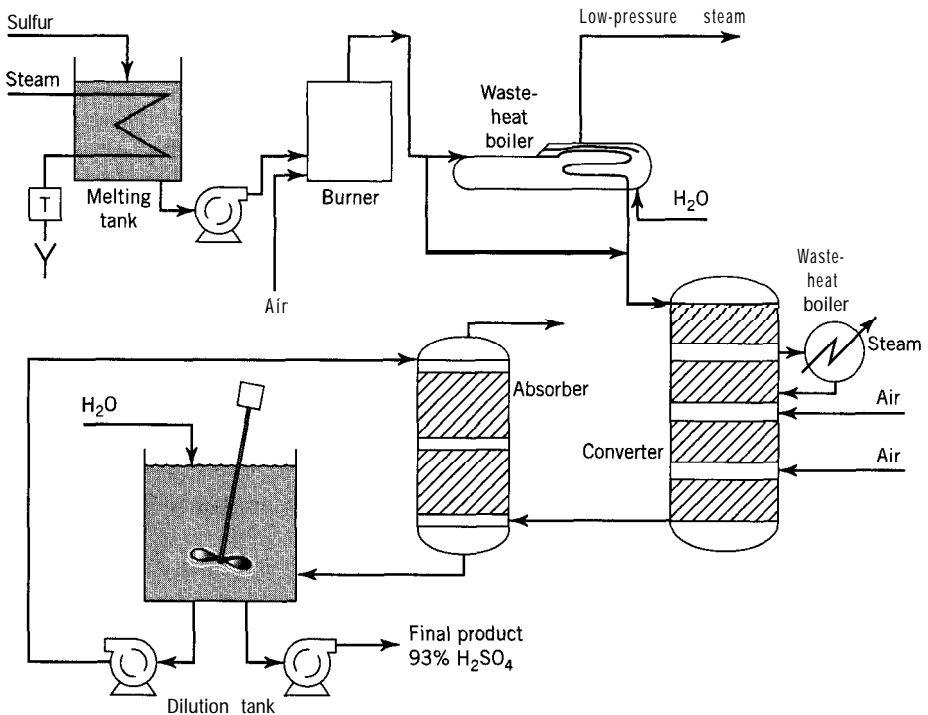
Figure B-6 Sulfuric acid process.

The liquid leaving the absorber, concentrated H_2SO_4 (98%), goes to a circulation tank where it is diluted back to 93% with H_2O. Part of the liquid from this tank is then used as the absorbing medium in the absorber.

 A. The following variables are thought to be important to control:
  1. Level in the melting tank.
  2. Temperature of sulfur in the melting tank.
  3. Air to the burner.
  4. Level of water in the waste-heat boiler.
  5. Concentration of SO_3 in the gas leaving the absorber.
  6. Concentration of H_2SO_4 in the dilution tank.
  7. Level in the dilution tank.
  8. Temperature of the gases entering the first stage of the converter.
  Design the necessary control systems. Be sure to specify the action of valves and controllers. Briefly discuss your design.
 B. How would you set the production rate for this plant?

Case 7. Fatty Acid Process

Consider the process shown in Fig. B-7. The process hydrolyzes crude fats into crude fatty acids (CFA) and dilute glycerine by using a continuous high-pressure fat splitter
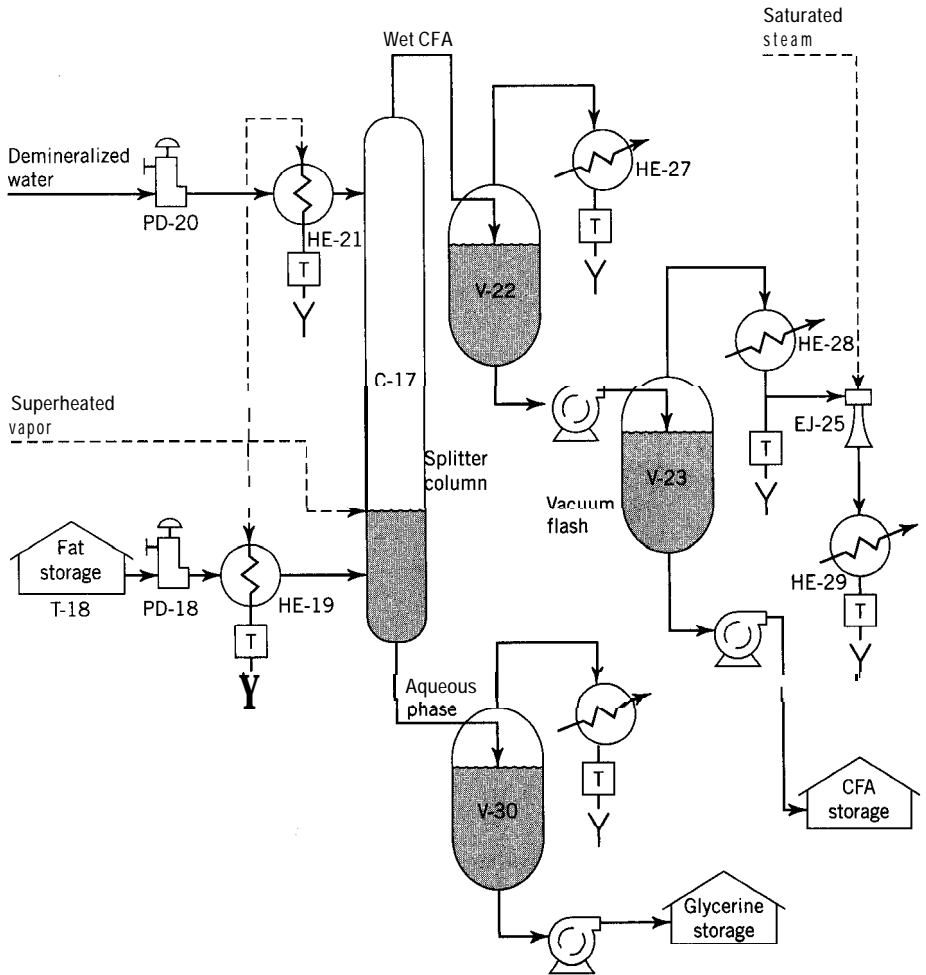
Figure B-7 Fatty acid process.

column (C-17). The main product is high-quality CFA. The CFA quality is primarily a function of the acid value. In the column the following reaction takes place:

$$
\begin{array}{lll}
CH_2OCOR & & CH_2OH \quad RCOOH \\
CHOCOR' \; + \; 3H_2O \xrightarrow[\text{and pressure}]{\text{High temperature}} & CHOH \; + \; R'COOH \\
CH_2OCOR'' & & CH_2OH \quad R''COOH \\
\text{TRIGLYCERIDE} \quad \text{WATER} & & \text{GLYCERINE} \quad \text{MIXED FATTY ACIDS}
\end{array}
$$
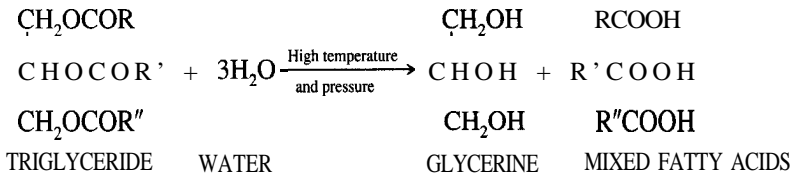
Fat is stored (T-18) at 120°F and pumped into the column by means of a positive displacement pump (PD-18). The fat is preheated (HE-19) to 400°F with superheated steam before it enters the column. The column operates continuously at **700** psig and 500°F with a crude fat feed rate of 25,000 lb/h.

Demineralized water is pumped into the column by means of a positive displacement pump (PD-20). The water is preheated (HE-21) to 500°F. Excess water is required to ensure complete hydrolysis of the crude fat.

Superheated steam at 800 psig and 700°F is sparged directly into the column. The steam provides heat and mixing to break up the fat.

The splitter is basically a countercurrent contractor. Water feed at the top has a higher specific gravity than the CFA. Crude fat feed at the bottom is insoluble in water and rises as the water migrates down the column. The glycerine produced by the reaction is soluble in water and increases the specific gravity of the aqueous phase.

An interface forms in the column. Above the interface the material is mostly fat and CFA. Below the interface it is mostly aqueous phase of water and glycerine. The best operation of the column is achieved when this interface is located near the steam sparger. If the interface level is low, then the amount of CFA in the aqueous phase increases. If the level is too high, fat dispersion into the water is lost and incomplete hydrolysis results. High temperature is required to produce the hydrolysis reaction, but boiling must be avoided, because this condition causes the aqueous phase to rise and upset the column.

The material removed overhead contains CFA and a small amount of water. This wet CFA is a light brown, milky material. The overhead product is dried by a two-step flash process. The sensible heat of the material is enough to dry the material without heat. The material is sprayed into the first vessel (V-22), and most of the water evaporates. The overhead water is condensed (HE-27). The resulting CFA is then sent to a vacuum flash (V-23) to dry the material fully. A steam jet ejector (EJ-25) is used to draw vacuum. The overhead water in the vacuum flash is condensed in the precondenser (HE-28) and sent to the sewer. The noncondensables from the precondenser are pulled through the steam jet ejector, and the motive steam is condensed in the barometric condenser (HE-29). The vacuum flash tank should be operated at 100 mm Hg. The ejector is significantly oversized for normal duty and consumes 2500 lb/h of 150-psig saturated steam. Very low pressure will cause low-molecular-weight elements of the CFA to vaporize and foul the precondenser. Loss of vacuum allows wet CFA to remain in the tank, which will cause problems in downstream processes.

The aqueous phase is removed from the bottom of the column and should be 20 weight % glycerine dissolved in water. Like the CFA, the aqueous phase is flashed at atmospheric pressure (V-30). Any fatty material in the aqueous phase makes purification of the glycerine very difficult. Excess water in the aqueous phase requires additional energy in the glycerine purification. Glycerine is a clear, colorless liquid.

   A. Prepare a detailed instrument diagram to control:
     1. The level in the splitter column.
     2. The level in all flash tanks.
     3. The pressure in the column.
     4. The pressure in the vacuum flash.
     5. The temperature in the splitter column.
     6. The temperature in the heaters.
    All instruments shown should be tagged, and the normal operating value and proposed range of the instrument should be provided.
   B. Describe methods to determine by inference process parameters that cannot easily be directly measured.

C. Besides control, you should select other process parameters to monitor to help the operator detect and diagnose problems.

D. Recommended alarm settings should be provided to alert the operator to abnormal operation. Alarms should be provided only for conditions that significantly affect the unit operation. Also provide a recommended list of process interlocks that describe automatic control actions to stop pumps or close valves on the basis of monitored process conditions.

E. How would you set the production rate for this plant?

Appendix C

Sensors, Transmitters, and Control Valves

This appendix presents some of the hardware necessary to implement control systems and is closely related to Chapter 5. Some of the most common sensors-pressure, flow, level, and temperature-are presented, as well as two different types of transmitters, one pneumatic and the other electronic. The appendix ends with a presentation of the different types of control valves and of additional considerations in the sizing of these valves.

C-l PRESSURE SENSORS

The most common pressure sensor (see References **1, 2, 3,** and **4)** is the **Bourdon tube,** developed by the French engineer Eugene Bourdon. The Bourdon tube, shown in Fig. C-l. 1, is basically a piece of tubing in the form of a horseshoe with one end sealed and the other end connected to the pressure source. The cross section of the tube is elliptical or flat, so the tubing tends to straighten as pressure is applied, and when the pressure is released, the tubing returns to its original form so long as the elastic limit of the material of the tubing was not exceeded. The amount of straightening that the tubing undergoes is proportional to the applied pressure. Thus if the open end of the tubing is fixed, then the closed end can be connected to a pointer to indicate pressure or to a transmitter to generate a signal.

The pressure range that can be measured by the Bourdon tube depends on the wall thickness and on the material of the tubing. An extended Bourdon tube in the form of a helical spiral was developed to permit additional motion of the sealed end. This element, called the helix, is shown in Fig. C-1.2. The helix can handle pressure ranges of about 10:1 with an accuracy of $\pm$ 1% of the calibrated span (Ryan, 1975). Another common type of Bourdon tube is the *spiral* element, shown in Fig. C-1.2*d.*

Another type of pressure sensor is the **bellows,** shown in Figs. C-1.2*c*, which looks like a corrugated capsule made up of a somewhat elastic material such as stainless steel or brass. Upon increasing pressure the bellows expands, and upon decreasing pressure
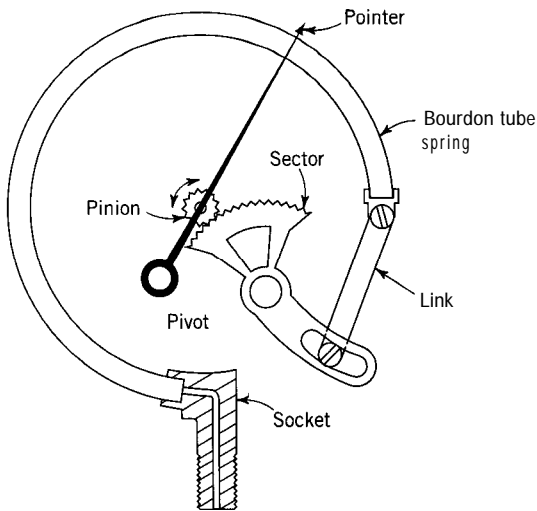
Figure **C-l.1** Simple Bourdon tube. (Courtesy of the
Instrument Society of America.)
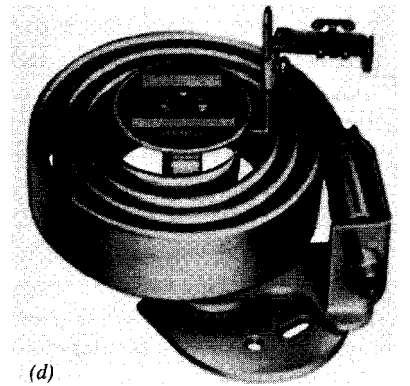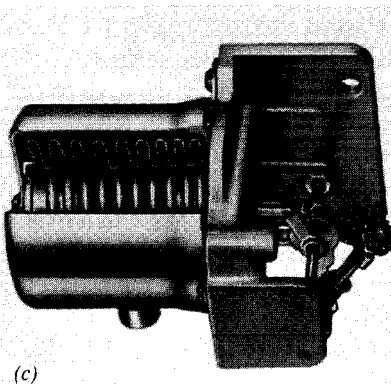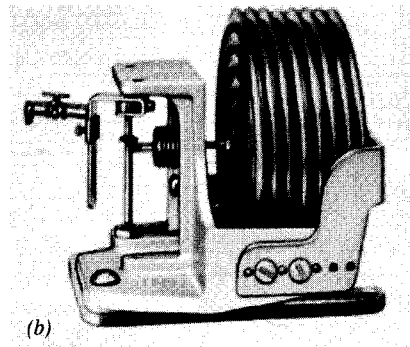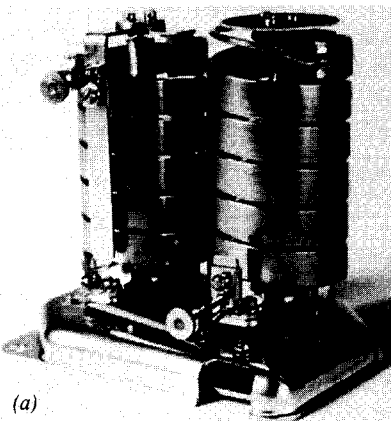
Figure C-l.2 Types of Bourdon tubes. (Courtesy of the Foxboro Co.) (a) Helical.
(b) Diaphragm. (c) Bellows. (d) Spiral.

it contracts. The amount of expansion or contraction is proportional to the applied pressure. Similar to the bellows is the **diaphragm** sensor, shown in Fig. C-1.2*b*. As the process pressure increases, the center of the diaphragm moves away from the pressure. The amount of motion is proportional to the applied pressure.

C-2 FLOW SENSORS

Flow is one of the two most commonly sensed process variables, the other being temperature; consequently, many different types of flow sensors have been developed (see References 2, 4, 5, 6, and 7). This section describes the most often used ones and mentions some others. Table C-2.1 (Zientara, 1972) shows several characteristics of some common sensors.

A common **flow** sensor is the **orifice meter,** which is a flat disk with a machined hole; see Fig. C-2.1. The disk is inserted in the process line perpendicular to the fluid motion with the intention of producing a pressure drop, Ap,. This pressure drop across the orifice is a nonlinear function of the volumetric flow rate through the orifice. Accurate orifice meter flow equations are complex and are presented in many fine books (see References 8, 9, and 10). However, most installations probably use the following simple equation:

$$f = C_o A_o \sqrt{\frac{\Delta p_o}{\rho(1 \quad \beta^4)}} \qquad \text{(C-2.1)}$$

where

f = volumetric flow rate
Ap, = pressure drop across orifice
A, = area of orifice
C_o = orifice coefficient
p = fluid density
β = dimensionless ratio of the diameter of the orifice, *d,* to the diameter of the pipe, *D*

Equation C-2.1 derives from the application of a mass balance and a mechanical energy balance (Bernoulli). The references cited also show how to size the required orifice diameter. Most orifice diameters vary between 10% and 75% of the pipe diameter, $0.1 < \beta < 0.75$.

The pressure drop across the orifice is measured with taps. Flange taps, shown in Fig. C-2.2, are the most common. They measure the pressure drop across the flanges holding the orifice in the process line. Other types include vena contracta taps, radius taps, corner taps, and line taps. These are not as popular as flange taps.

The tap upstream from the orifice is called the high-pressure tap, and the one downstream from the orifice is called the low-pressure tap. Most tap diameters vary between 1/4 and 3/4 in. The pressure drop sensed will be a function of tap location as well as flow rate. A differential pressure sensor (Fig. C-2.3) is used to measure the pressure drop across the orifice.

Several things must be stressed about the use of orifice meters to measure flows. The first is that the output signal from the orifice/transmitter combination is the pressure

Table C-2.1 Characteristics of Typical Flow Sensors

Primary Element	Type of Fluid	Pressure Loss*	Flow Rang.	Error	Upstream Piping[b]	Viscosity Effect	Readout
Concentric orifice	Liquids, gases, & steam	50–90%	3:1	314%	10–30D	High	Square root
Segmental orifice	Liquid slurries	60–100%	3:1	2.5%	10–30D	High	Square root
Eccentric orifice	Liquid-gas comb.	60–100%	3:1	2%	10–30D	High	Square root
Quadrant edged orifice	Viscous liquids	45-85%	3:1	1%	20–50D	Low	Square root
Segmental wedge	Slurries & viscous liquids	30–80%	3:1	1%	10–30D	Low	Square root
Venturi tube	Liquids & gases	10–20%	3:1	1%	5–10D	Very high	Square root
Dall tube	Liquids	5–10%	3:1	1%	5–10D	High	Square root
Flow nozzle	Liquids & gases & steam	30–70%	3:1	1.5%	10–30D	High	Square root
Elbow meter	Liquid	None	3:1	1%	30D	Negligible	Square root

Rotameter	All fluids	1 – 200" WG	10:1	2%	None	Medium	Linear
V-notch weir	Liquids	None	30:1	4%	None	Negligible	5/2
Trapezoidal weir	Liquids	None	10:1	4%	None	Negligible	3/2
Parshall flume	Liquid slurries	None	10:1	3%	None	Negligible	3/2
Magnetic flowmeter	Liquid slurries	None	30:1	1%	None	None	Linear
Turbine meter	Clean liquids	0-7 psi	14:1	0.5%	5–10D	High	Linear
Pitot tube	Liquids	None	3:1	1%	20-30D	Low	Square root
Pitot venturi	Liquids & gases	None	3:1	1%	20-30D	High	Square root
Positive displace-ment	Liquids	0- 15 psi	10:1	OS-2%	None	None	Linear totalization
Swirlmeter	Gases	0-2 psi	10:1-100:1	1%	10D	None	Linear
Vortex shedding	Liquids & gases	0-6 psi 0–5" WG	30:1-100:1	0.25%	15–30D	Minimum Reynolds No. 10,000	Linear
Ultrasonic	Liquids	None			None	None	Linear

[a] Pressure loss percentages are stated as percentages of differential pressure produced.

[b] Upstream piping is stated in the number of straight pipe diameters required before the primary element.

A-703-47

(a)

(b)

(c)

(d)
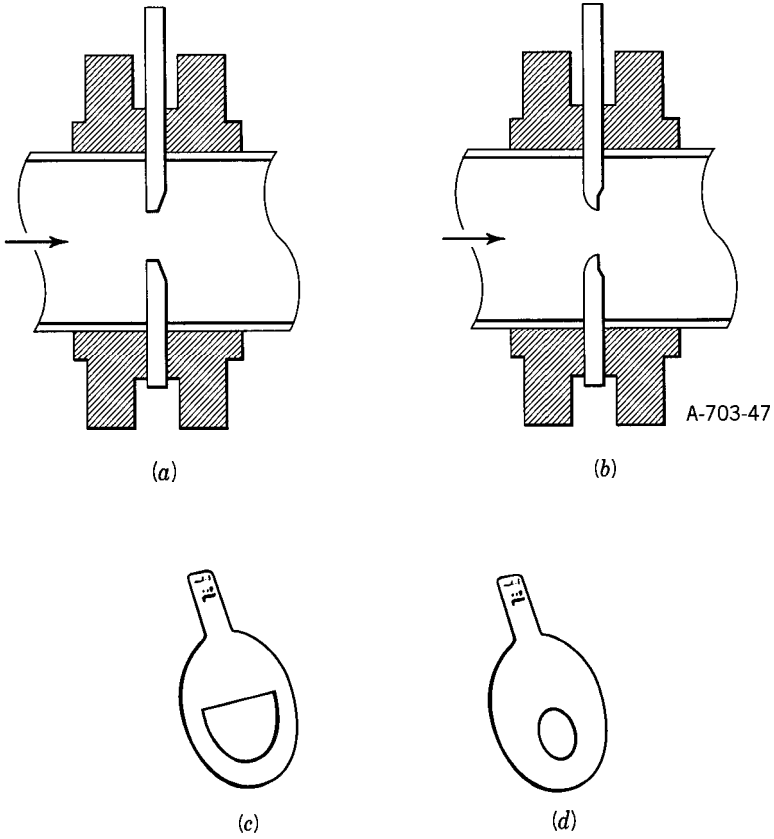
Figure C-2.1 Schematic of orifice meters. (a) Sharp edge. (b) Quadrant edge. (Courtesy of ABB Kent-Taylor.) (c) Segmental edge. (d) Eccentric edge. (Courtesy of Foxboro Co.)

Weld

Weld

(a)

(b)

(c)
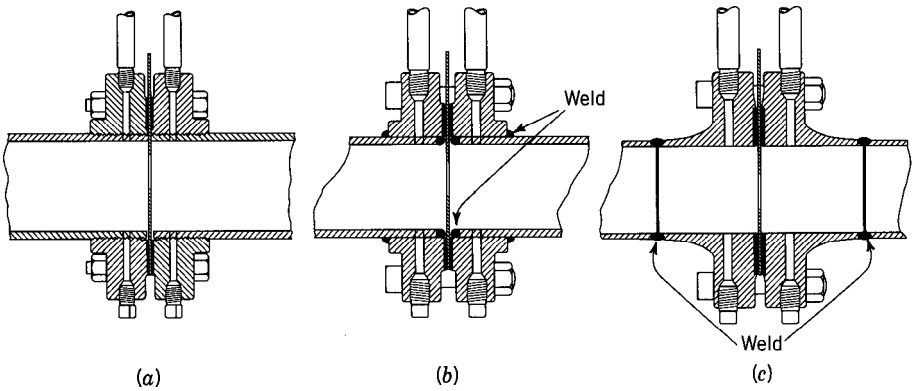
Figure C-2.2 Flange taps. (a) Threaded union. (b) Slip-on union. (c) Welding-neck union. (Courtesy of the Foxboro Co.)

Pivot

Low pressure

Low

High

High pressure

Diaphragm

(a)

(b)
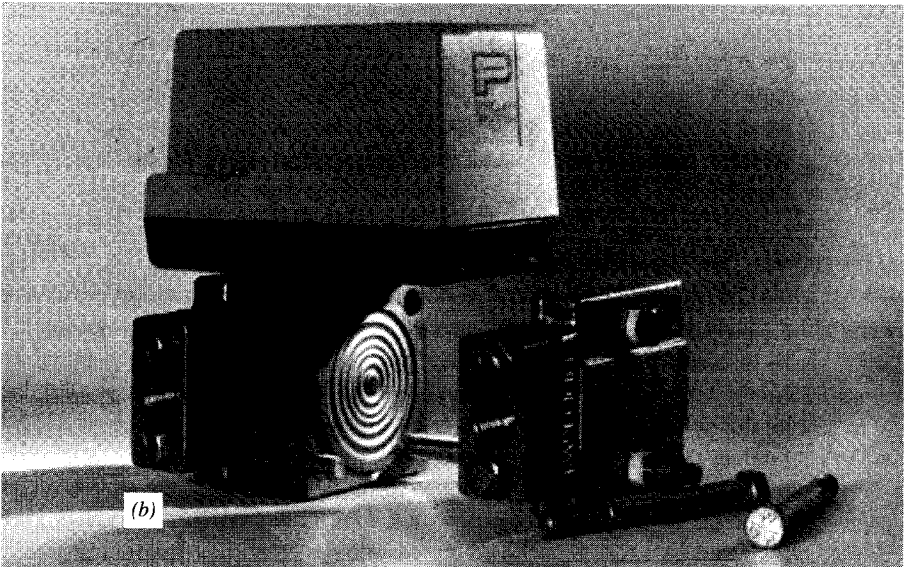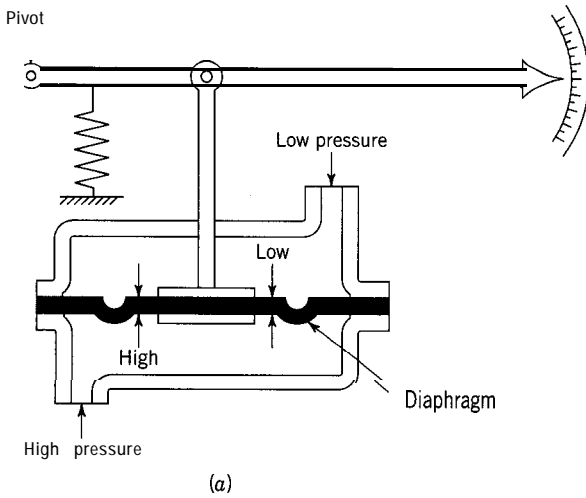
Figure C-Z.3 Differential pressure sensor and transmitter. (a) (Courtesy of the Instrument Society of America.) (b) (Courtesy of Fischer & Porter.)

drop across the orifice, not the flow. Equation C-2.1 shows that this pressure drop is related to the square of the volumetric flow rate, or

$$\Delta p_o \propto \rho f^2 \qquad\qquad \text{(C-2.2)}$$

Consequently, if the flow is desired, then the square root of the pressure drop must be obtained; Chapter 11 presents square root extractors. Most manufacturers offer the option of installing a square root extraction unit within the transmitter. In this case, the output signal from the transmitter is linearly related to the volumetric flow. More im-
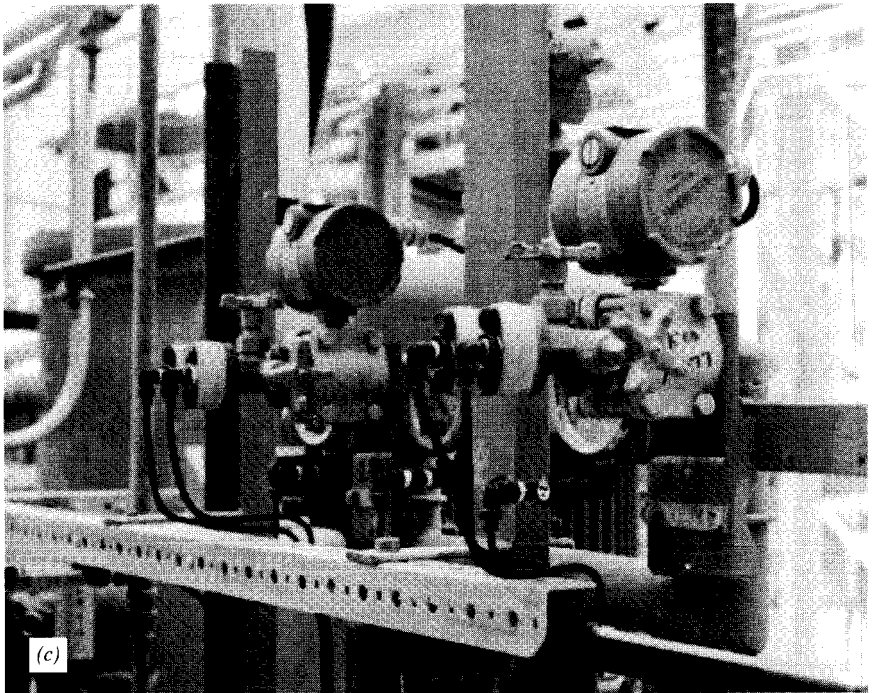
Figure C-2.3 (Continued) (c) (Courtesy of Rosemount Inc.)

portant, in distributed control systems (DCSs) and other microprocessor-based systems, a square root extractor is not needed because the square root is an input option. That is, the control system can be configured such that when it reads a signal, it extracts the square root automatically and keeps that result in memory. The second thing that must be stressed is that not all of the pressure drop measured by the taps is lost by the process fluid. A certain amount is recovered by the fluid, in the next few pipe diameters, as it reestablishes its flow regime. Finally, the rangeability of the orifice meter-the ratio of the maximum measurable flow to the minimum measurable flow-is about 3 : 1, as indicated in Table C-2.1. This rangeability is important to know; it indicates the expected accuracy when running the process at low or high loads.

Several conditions may prevent the use of orifice sensors. Among such causes are available pressure not being enough to provide pressure drop, as in the case of gravity flow; the flow of corrosive fluids; fluids with suspended solids that may plug the orifice; and fluids close to their saturated vapor pressure that may flash when subjected to a drop in pressure. These cases require the use of other sensors to measure flow.

Another common type of sensor is the *magnetic flowmeter*, shown in Fig. C-2.4. The operating principle of this element is Faraday's law; that is, as a conductive material (a fluid) moves at right angles through a magnetic field, it induces a voltage. The voltage created is proportional to the intensity of the magnetic field and to the velocity of the fluid. If the intensity of the magnetic field is constant, then the voltage is proportional only to the velocity of the fluid. Furthermore, the velocity measured is the average velocity, so this sensor can be used for both regimes, laminar and turbulent. During calibration of this flowmeter, the cross-sectional area of the pipe is taken into consid-
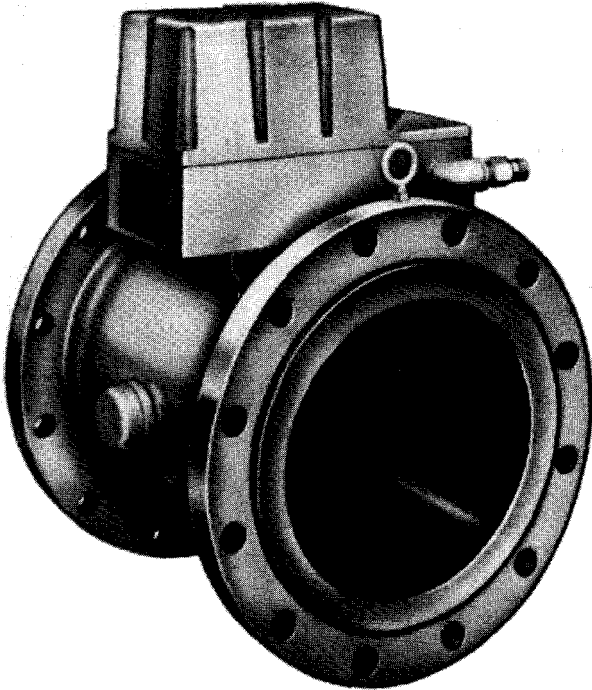
Figure C-2.4 Magnetic flowmeter. (Courtesy of Fischer &
Porter.)

eration so that the electronics associated with the meter can calculate the volumetric
flow. Thus the output is linearly related to the volumetric flow rate.

Because the magnetic flowmeter does not restrict flow, it is a zero-pressure-drop
device suitable for measuring gravity flow, slurry flows, and flow of fluids close to their
vapor pressure. However, the fluid must have a minimum required conductivity of about
10 μohm/cm^2, which makes the meter unsuitable for the measurement of both gases
and hydrocarbon liquids.

Table C-2.1 shows that the rangeability of magnetic flowmeters is 30: 1, which is
significantly greater than that of orifice meters; however, their cost is also greater. The
cost differential increases as the size of the process pipe increases.

An important consideration in the application and maintenance of magnetic flow-
meters is coating of the electrodes. This coating represents another electrical resistance
that results in erroneous readings. Manufacturers offer techniques such as ultrasonic
cleaners for maintaining clean electrodes.

Another important flowmeter is the turbine *meter* shown in Fig. C-2.5. This meter
is one of the most accurate of the commercially available flowmeters. Its working
principle consists of a rotor that the fluid velocity causes to spin. The rotation of the
blades is detected by a magnetic pickup coil that emits pulses the frequency of which
is proportional to the volumetric flow rate; this pulse is equally converted to a
4 to 20-mA signal. The problems most commonly associated with turbine meters arise
with the bearings, which require clean fluids that have some lubricating properties.

The mass flow rate measurement of liquids using the *Coriolis effect* is common in
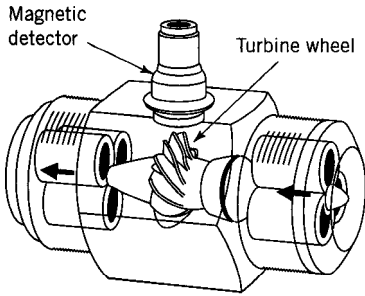
(a)
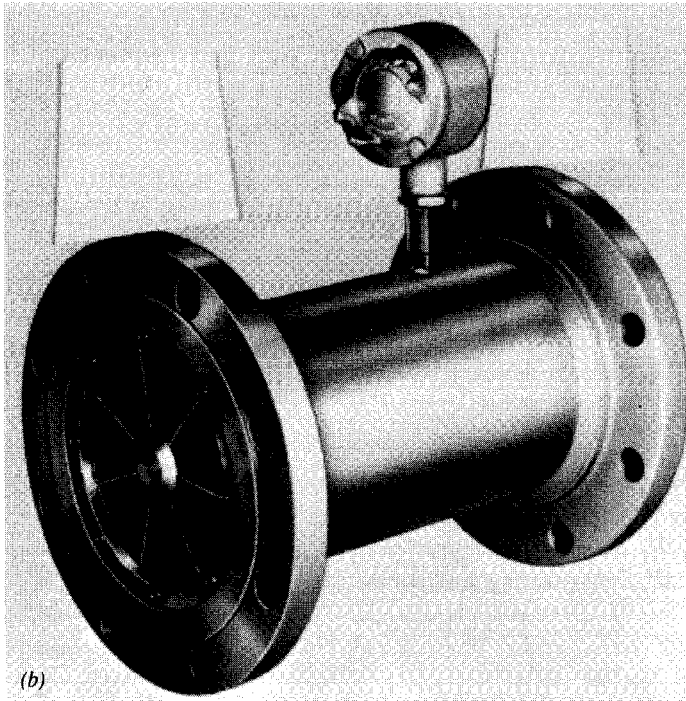
(b)

Figure C-2.5 Turbine flowmeter. (a) (Courtesy of the Instrument Society of America.) (b) (Courtesy of Fischer & Porter.)

the process industries. The measurement accuracy achieved with this effect is unaffected by changes in the fluid's temperature, density, pressure, or viscosity or by changes in velocity profile. Because the measurement principle is based only on mass flow, once the calibration of the flowmeter has been established with a conventional fluid such as water, it applies equally well to other fluids. The meter has a maximum error of ± 0.15% of reading over a dynamic range of 10:1 and is suitable for most fluids over a range of 100:1 with an accuracy of ± 1.5% of reading.

It can be shown that when a liquid is transported through a conduit fixed at its ends and vibrating at its center in a periodic fashion, as shown in Fig. C-2.6, it experiences a force perpendicular to the direction of flow. This force, referred to as the Coriolis
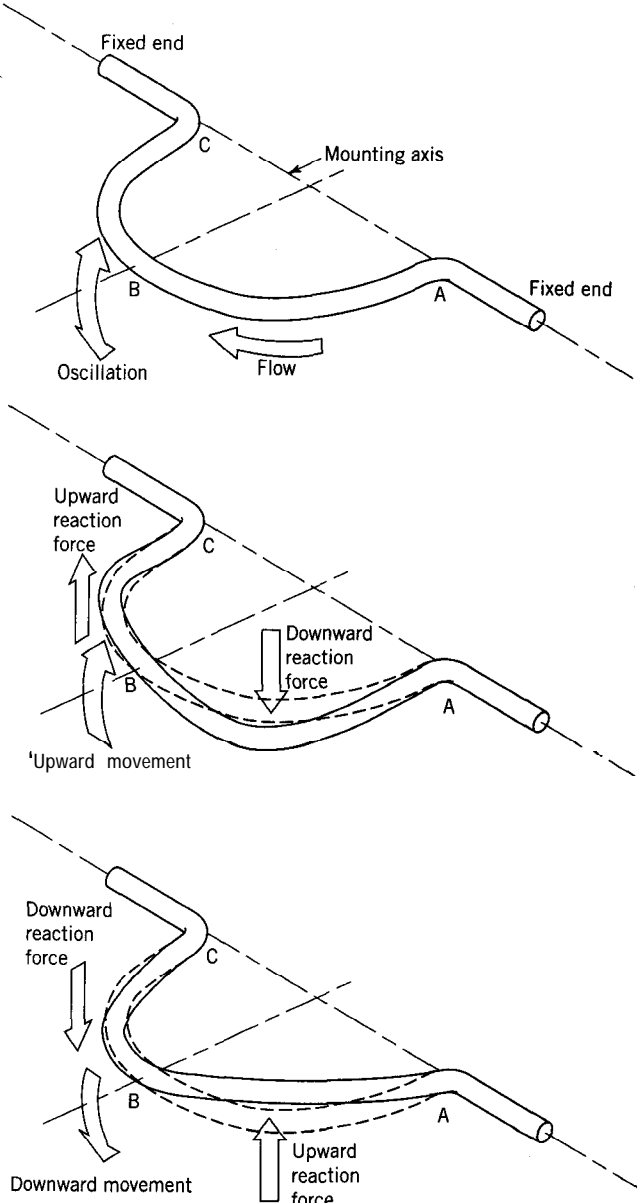
Figure C-2.6 Coriolis effect. (Drawing courtesy of **ABB**
K-FLOW   Inc.)

force, is exerted on the inner walls of the conduit. If the conduit has sufficient me-
chanical elasticity, then the Coriolis force produces small, elastic deformation in the
conduit. One can determine the mass flow rate by measuring these small deformations.

Because the deformations induced by the Coriolis force are small, accurately mea-
suring the effects of the force presents a challenge. Several ways have been found to
increase the meter's sensing capabilities and to increase the effect of the Coriolis force.

Young (1985) and Dahlin and Franci (1985) describe in more detail this important, accurate, and popular flow sensor.

We have briefly discussed four of the most common flowmeters in use in the process industries. There are many other types. They range from rotameters, flow nozzles, venturi tubes, pitot tubes, and annubars, which have been used for many years, to more recent developments such as vortex-shedding meters, ultrasonic meters, thermal conductivity mass meters, and swirlmeters. Limited space prevents our discussing these meters. The reader is directed to the many fine references cited at the beginning of this section for discussion of these meters.
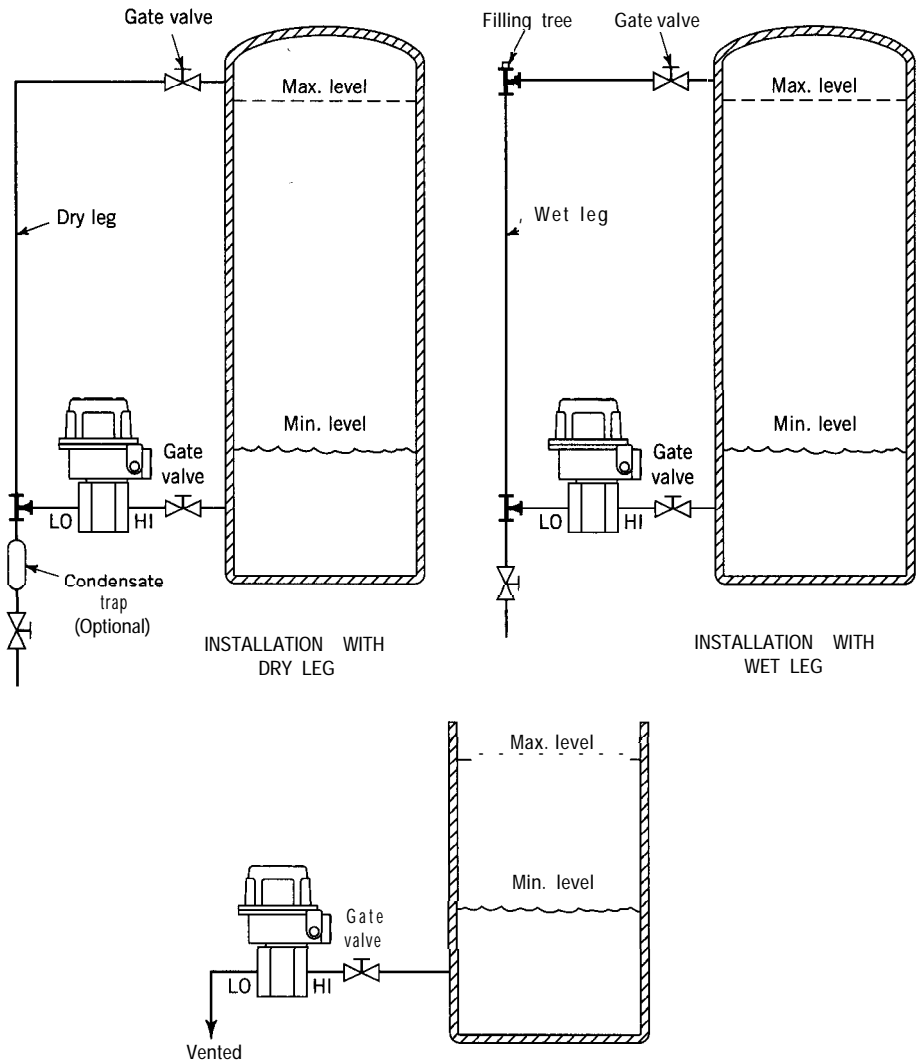
Figure C-3.1 Differential pressure transmitters installed in closed and in open vessels. (Courtesy of ABB Kent-Taylor.)

C-3   LEVEL   SENSORS

The three most important level sensors (see References 2, 5, 6, and 11) are the differential pressure, float, and air bubbler sensors. The *differential pressure* method consists of sensing the difference in pressure between the pressure at the bottom of a liquid and that above the liquid level, as shown in Fig. C-3.1. This differential pressure is caused by the hydrostatic head developed by the liquid level. The side that senses the pressure at the bottom of the liquid is referred to as the high-pressure side, and the one that senses the pressure above the liquid level is referred to as the low-pressure side. Knowing the differential pressure and the density of the liquid makes it possible to obtain the level. Figure C-3.1 shows the installation of the differential pressure sensor in open and closed vessels. If the vapors above the liquid level are noncondensable, then the low-pressure piping, also known as the dry leg, can be empty. However, if the vapors are likely to condense, then the wet leg must be filled with a suitable seal liquid. If the density of the liquid varies, then some compensation technique must be employed.

The *float sensor* detects the change in buoyant force on a body immersed in the liquid. This sensor is generally installed in an assembly mounted externally to the vessel,

(a)

Float

(b)
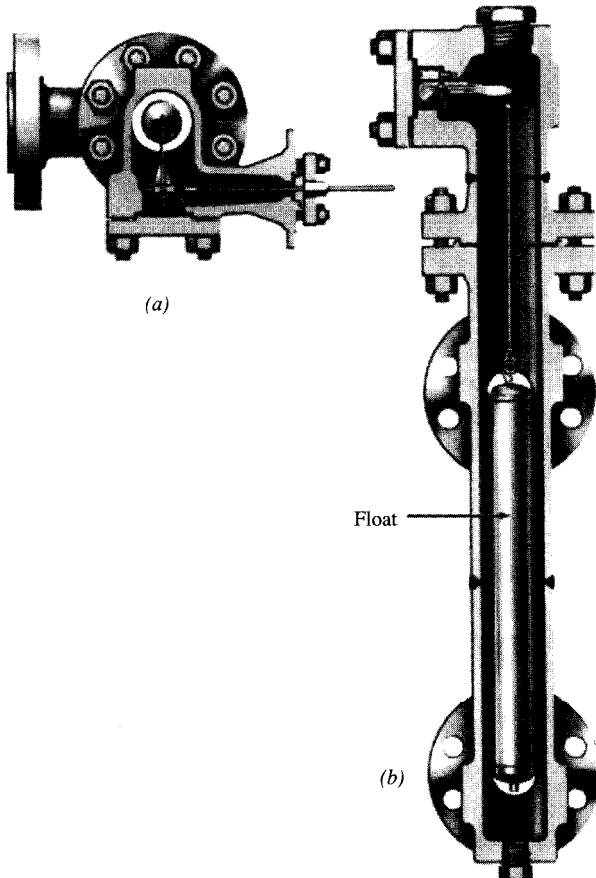
Figure C-3.2 Level float sensor. (a) Top view. (b) Front view. (Courtesy of ABB Kent-Taylor.)
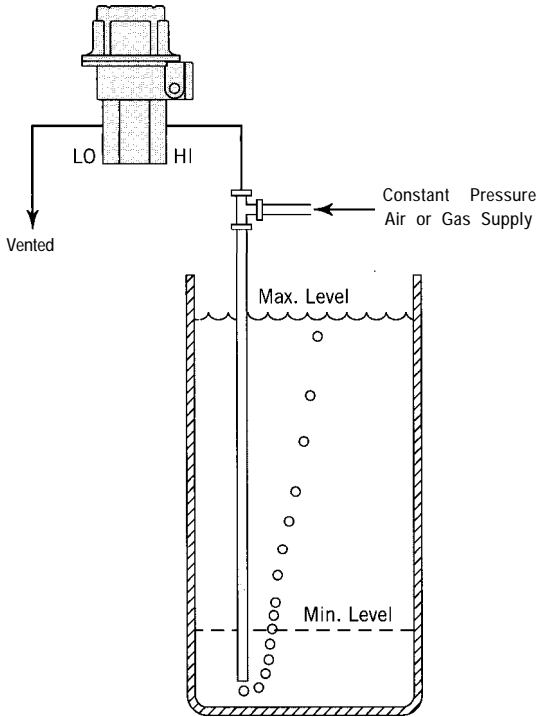
Figure C-3.3 Air bubbler level sensor. (Courtesy of ABB Kent-Taylor.)

as shown in Fig. C-3.2. The force required to keep the float in place, which is proportional to the liquid level, is then converted to a signal by the transmitter. Float sensors are less expensive than most other level sensors; however, a major disadvantage lies in their inability to change their zero and span. To change the zero requires relocation of the complete housing.

The *bubbler sensor* is another type of hydrostatic pressure sensor. As shown in Fig. C-3.3, it consists of an air or inert gas pipe immersed in the liquid. The air or inert gas flow through the pipe is regulated to produce a continuous stream of bubbles. The pressure required to produce this continuous stream is a measure of the hydrostatic head or liquid level.

There are some other ways to measure level in tanks, such as capacitance gauges, ultrasonic systems, and nuclear radiation systems. The last two sensors are also used to measure the level of solid material. The references cited at the beginning of this section are recommended for further reading.

C-4 TEMPERATURE SENSORS

Along with flow, temperature is the most frequently measured variable in the process industries. A simple reason is that very few physical phenomena are not affected by it. Temperature is also often used to infer other process variables. Two of the most common examples are in distillation columns and in chemical reactors. In distillation col-

Table C-4.1 Popular Sensors for Temperature
Measurement

I. Expansion thermometers
    A. Liquid-in-glass thermometers
    B. Solid-expansion thermometers (bimetallic strip)
    C. Filled-system thermometers (pressure thermometers)
        1. Gas-filled
        2. Liquid-filled
        3. Vapor-filled
II. Resistance-sensitive devices
    A. Resistance thermometers
    B. Thermistors
III. Thermocouples
IV. Noncontact methods
    A. Optical pyrometers
    B . Radiation pyrometers
    C. Infrared techniques

umns, temperature is commonly used to infer the purity of one of the exit streams. In chemical reactors, temperature is used as an indication of the extent of reaction or conversion.

Because of the many effects produced by temperature, numerous devices have been developed to measure it (see References 2, 3, and 6). With a few exceptions, the devices fall into four general categories, as shown in Table C-4.1. Quartz thermometers, py-rometric cones, and specialized paints are some of the sensors that do not fit the clas-sification scheme shown in Table C-4.1. Table C-4.2 (Zientara, 1972) shows some characteristics of typical sensors.

Liquid-in-glass thermometers indicate temperature change caused by the difference between the temperature coefficient of expansion for glass and the liquid employed. Mercury and alcohol are the most widely used liquids. Mercury-in-glass thermometers made from ordinary glass are useful between −35°F and 600°F. The lower limit is due to the freezing point of mercury and the upper limit to its boiling point. By filling the space above the mercury with an inert gas (usually nitrogen) to prevent boiling, the useful range may be extended to 950°F. Such thermometers usually bear the inscription "nitrogen filled." For temperatures below the freezing point of mercury (- 38°F), an-other liquid must be employed. Alcohol is the most widely used fluid for temperatures down to − 80°F, pentane for temperatures down to   200°F, and toluene for tempera-tures below   230°F.

The *bimetallic strip thermometer* works on the principle that metals expand with temperature and that the expansion coefficients are not the same for all metals. Figure C-4.1 shows a typical bimetallic strip thermometer. The temperature-sensitive element is a composite of two different metals fastened together into a strip. One metal has a high thermal expansion coefficient, and the other metal has a low thermal expansion coefficient. A common combination is invar (64% Fe, 36% Ni), which has a low co-efficient, and another nickel-iron alloy that has a high coefficient. Usually the expan-sion with temperature is low, and this is the reason for having the bimetallic strip wound

Table C-4.2 Characteristics of Typical Temperature Sensors

Sensor	Range, °F	Accuracy, °F	Advantages	Disadvantages
Glass-stem thermometer	Practical: − 200 to 600 Extreme: 321 to 1100	0.1-2.0	Low cost Simplicity Long life	Difficult to read Only local measurement, No automatic control or recording capability
Bimetallic thermometer	Practical: − 80 to 800 Extreme: − 100 to 1000	1.0-20	Less subject to breakage Dial reading Less costly than thermal or electrical	Less accurate than glass-stem thermometer Changes calibration with rough handling
Filled thermal elements	Practical: 300 to 1000 Extreme: − 450 to 1400	0.5−2% of full scale	Simplicity No auxiliary power needed Sufficient response times	Larger bulb size than electrical systems and greater minimum spans Bulb to readout distance is maximum of 50-200 ft. Factory repair only
Resistance thermometer	− 430 to 1800	0.1 (best)	System accuracy Low spans (10°F) available Fast response Small size	Self-heating may be a problem Long-term drift exceeds that of thermcouple Some forms expensive and difficult to mount

Thermocouple	− 440 to 5000	0.2 (best)	Small size, low cost Convenient mount Wide range	Not so simple as direct-reading thermometers Cold working on wires can affect calibration. 70°F nominal minimum span
Radiation py-rometer	0 to 7000	0.5 – 1.0% of full scale	No physical contact Wide range, fast response Measure small target or average over large area	More fragile than other electrical devices Nonlinear scale, relatively wide span required
Thermistor	− 150 to 600	0.1 (best)	Small size Fast response Good for narrow spans Low cost, stable No cold junction	Very nonlinear response Stability above 600°F is a problem. Not suitable for wide spans High resistance makes system prone to pick up noise from power lines
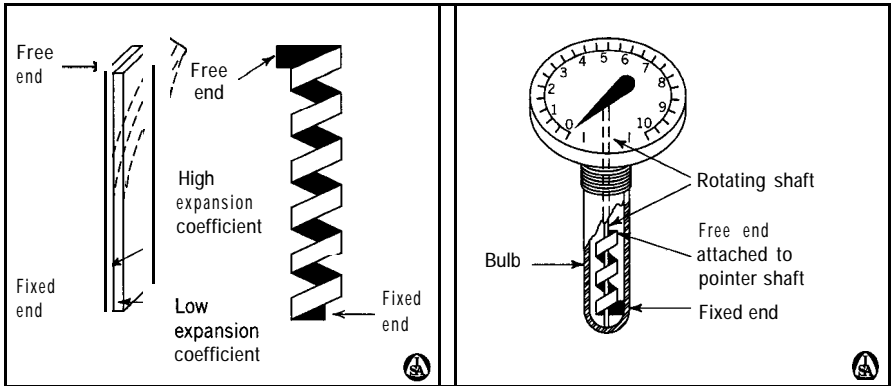
Figure C-4.1 Bimetallic strip thermometer. (Courtesy of the Instrument Society of America.)

in the form of a spiral. As the temperature increases, the spiral tends to bend toward the side of the metal with the low thermal coefficient.

Fig. C-4.2 shows the elements of a typical filled-system *thermometer.* Temperature variations cause the expansion or contraction of the fluid in the system, which is sensed by the Bourdon spring and transmitted to an indicator or transmitter. Because of their design simplicity, reliability, relatively low cost, and inherent safety, these elements are popular in the process industries. The Scientific Apparatus Manufacturer's Association **(SAMA)** has established four major classes, with subclassifications, of filled systems. Table C-4.3 lists these classifications. The most significant differences between the classifications are the fluid used and the compensation for temperature difference among the bulb, the capillary, and the Bourdon spring. For a more extensive description of these systems, see References 2 and 3.

Resistance thermometer devices **(RTDs)** are elements based on the principle that the
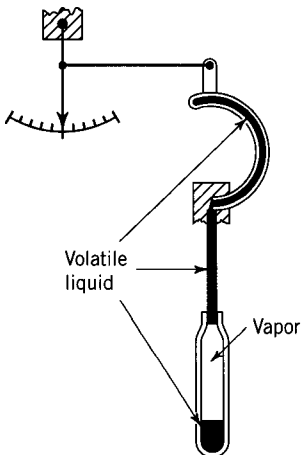
Figure C-4.2 Typical **filled-**system thermometer. (Courtesy of the **Instrument** Society of America.)

Table C-4.3 Scientific Apparatus Manufacturers' Association (SAMA)
Classification of Filled-System Thermometers

Class	Filling	Characteristics
I	Liquid other than mercury	Uncompensated
I A	Liquid other than mercury	Case and capillary compensated
IB	Liquid other than mercury	Case compensated
IIA	Vapor	For bulb above ambient applications
IIB	Vapor	For bulb below ambient applications
IIC	Vapor	For bulb either above or below ambient Large bulb used
IID	Vapor	For bulb either above or below ambient Nonvolatile liquid used for transmission
IIIA	Gas	Case and capillary compensated
IIIB	Gas	Case compensated
VA	Mercury	Case and capillary compensated
VB	Mercury	Case compensated

Note; There is no **SAMA** classification IV.

electrical resistance of pure metals increases with an increase in temperature. Because measurements of electrical resistance can be made with high precision, this also provides a very accurate way to make temperature measurements. The most commonly used metals are platinum, nickel, tungsten, and copper. Figure C-4.3 is a schematic of a typical RTD. A Wheatstone bridge is generally used for the resistance reading and, consequently, for the temperature reading.

Thermistor elements detect very small temperature changes. Thermistors are made of a sintered combination of ceramic material and some kind of semiconducting metallic oxide such as nickel, manganese, copper, titanium, or iron. Thermistors have a very
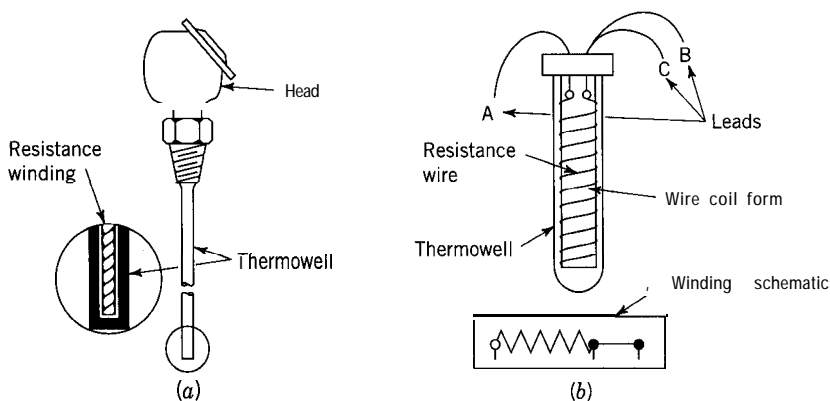
Figure C-4.3 Schematic of a resistance thermometer device. (a) Assembly. (b) Components. (Courtesy of the Instrument Society of America.)
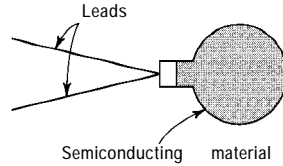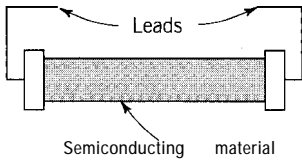
Figure C-4.4 Typical thermistor construction. (Courtesy of the Instrument Society of America.)
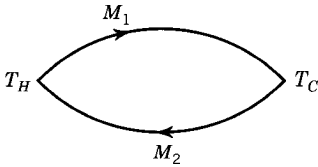
Figure C-4.5 Simple thermo-couple circuit.

Table C-4.4 Voltages Generated (millivolts) by Different Types of Thermocouples

°F	Millivolts
	0	2	4	6	8	10
100	- 2. 699	- 2. 736	- 2. 773	- 2. 810	- 2. 867	- 2. 883
10	- 0. 478	- 0. 435	- 0. 392	- 0. 349	- 0. 305	- 0. 262
100	1. 520	1. 566	1.611	1. 657	1. 703	1. 748
			Type K: Nickel–chromium vs. nickel–aluminum (chromel alumel)			
- 100	- 2. 581	- 2. 616	- 2. 650	- 2. 665	- 2. 719	- 2. 753
10	- 0. 467	- 0. 425	- 0. 383	- 0. 341	- 0. 299	- 0. 256
100	1. 518	1. 565	1.611	1. 450	1. 705	1. 752
			Type T: Copper vs. copper-nickel (copper-constantan)			
- 100	- 3. 492	- 3. 541	- 3. 590	- 3. 639	- 3. 688	- 3. 737
10	- 0. 611	- 0. 556	- 0. 501	- 0. 445	- 0. 390	- 0. 334
100	1. 942	2. 000	2. 058	2. 117	2. 175	2. 233
			Type J: Iron vs. copper-nickel (iron–constantan)			

high negative, or sometimes positive, temperature coefficient of resistivity. Figure C-4.4 shows some typical thermistors. Their advantages include small size and low cost. Their main disadvantages lie in their nonlinear relationship between temperature and resistance and in the fact that they usually require shielded power lines. Wheatstone bridges are generally used to measure the resistance and, therefore, temperature.

The last temperature element that we will discuss is the thermocouple, probably the best-known industrial temperature sensor. The thermocouple works on a principle discovered by T. J. Seebeck in 1821. The Seebeck effect, or Seebeck principle, states that an electric current flows in a circuit of two dissimilar metals if the two junctions are at different temperatures. Figure C-4.5 is a schematic of a simple circuit. M, and M_2 are the two metals, T_H is the temperature being measured, and T_C is the temperature of what is usually known as the cold, or reference, junction. The voltage produced by this thermoelectric effect depends on the temperature difference between the two junctions and on the metals used. Table C-4.4 shows some voltages generated by typical metals. The most common types of thermocouples are platinum-platinum/rhodium alloy, copper-constantan, iron-constantan, chromel-alumel, and chromel-constantan. Figure C-4.6 shows an assembly of an industrial thermocouple set-up. The protecting tube, also called a thermowell, is not necessary in all installations. This thermowell tends to slow down the response of the sensor system. For a more detailed discussion of thermocouples, see References 2 and 3.
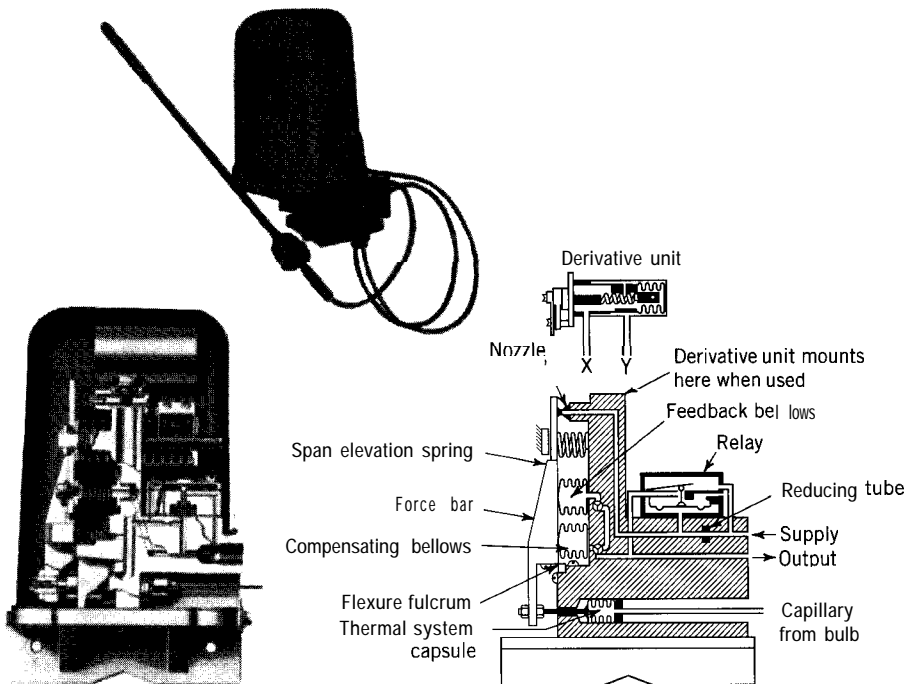
(a)

Figure C-4.6 Assembly of thermocouple systems. (a) (Courtesy of the Foxboro Co.)
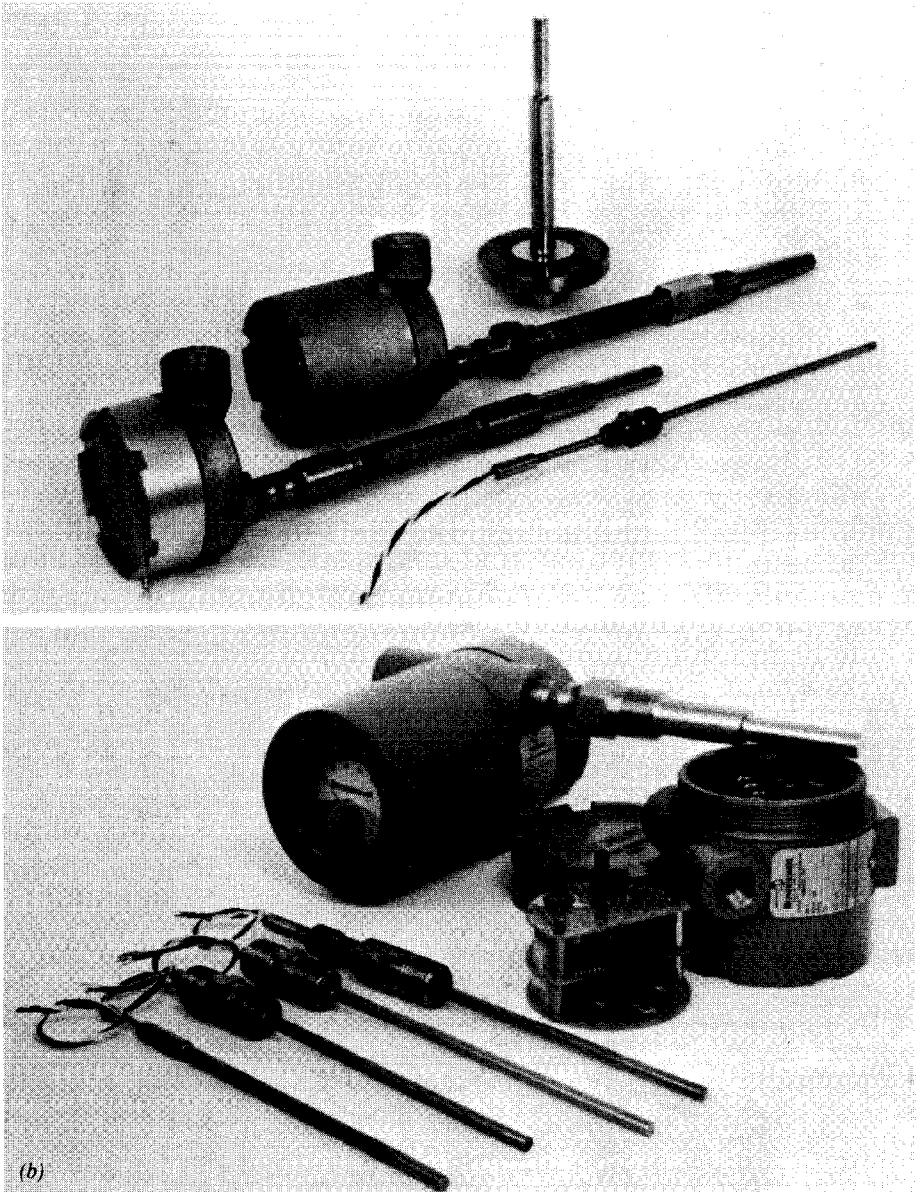
Figure C-4.6 (Continued) (b) (Courtesy of Rosemount, Inc.)

C-5 COMPOSITION SENSORS

Another important class of sensors consists of composition sensors (see References 2, 25, 26, 27, and 28). These sensors are used in environmental and product quality measurement and control. There are many different types of measurement sensors, such as density, viscosity, chromatography, pH, and ORP. Because of space limitations, we cannot present these sensors; however, we want to make the reader aware of their importance (see the references cited.)

C-6 TRANSMITTERS

This section presents an example of a pneumatic transmitter and an example of an electrical transmitter. The objective is to familiarize the reader with the working principles of these typical transmitters. The purpose of a transmitter is to convert the output from the sensor to a signal strong enough to be transmitted to a controller or any other receiving device. Most transmitters are either force-balance or motion-balance transmitters.

C-6.1 Pneumatic Transmitter

All pneumatic transmitters use a flapper-nozzle arrangement to produce an output signal proportional to the output from the sensor. A pneumatic differential pressure transmitter (see Reference 12), which is a force-balance transmitter, will be used to illustrate the working principles. This transmitter is shown in Fig. C-6.1.

The twin diaphragm capsule is the sensor. It senses the difference in pressure between the high- and low-pressure sides. Previously, we learned that this type of sensor is used to measure liquid level and flow. The diaphragm is connected to a force bar by a flexure. The force bar is connected to the body of the transmitter by a stainless steel diaphragm. This diaphragm serves as a seal to the measuring cavity and also as a positive fulcrum for the force bar. The top of the force bar is connected by a flexure strap to a range rod. This range rod has a range wheel that also serves as a fulcrum. A feedback bellows and a zero adjustment are located in the bottom part of the range rod. Above the range rod, a flapper-nozzle arrangement and a pneumatic relay are located. As shown in the figure, the flapper is connected to the combination of force bar and range rod.

As the diaphragm capsule senses a difference in pressure, this creates a tension or
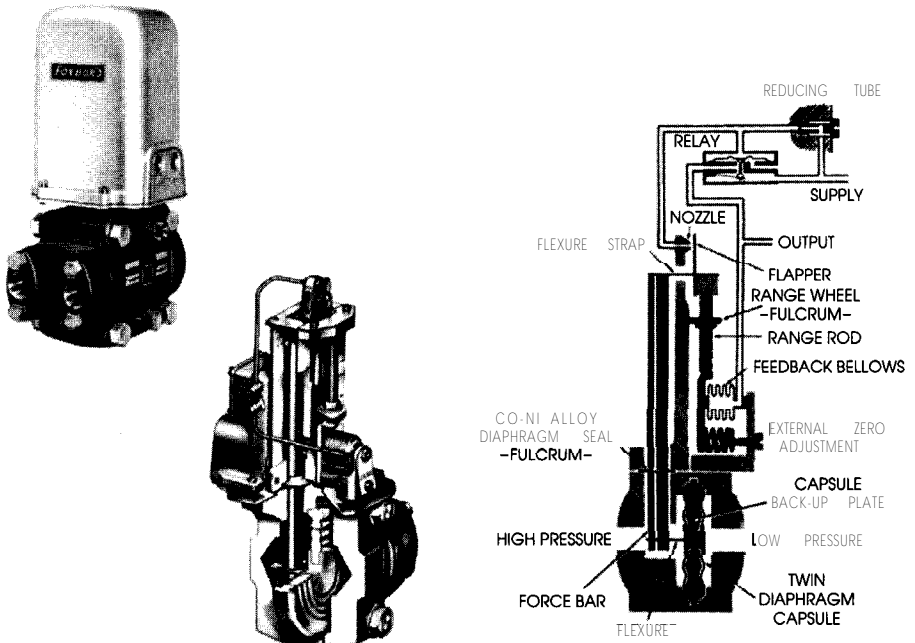
Figure C-6.1 Pneumatic differential pressure transmitter. (Courtesy of Foxboro Co.)

force on the lower end of the force bar. To be more specific, you may assume that the pressure on the high side increases, creating a pulling force on the force bar. This force results in a motion at the outer end of the bar, causing the flapper to move closer to the nozzle. In this case the output of the relay increases, and this increases the force that the feedback bellows exerts on the range rod. This force balances the force of the differential pressure across the diaphragm capsule. These balanced forces result in an output signal from the transmitter that is proportional to the difference in pressure.

The recommended supply pressure to most pneumatic instruments is between 20 and 25 psig. This ensures proper performance at the 15-psig output level. The calibration of these instruments requires adjustment of the zero and span (or range). In the instrument shown in Fig. C-6.1, this is done with the external zero adjustment screw and with the range wheel.

The preceding paragraphs have described the working principle of a typical pneumatic instrument. As we noted at the beginning, all pneumatic instruments use some
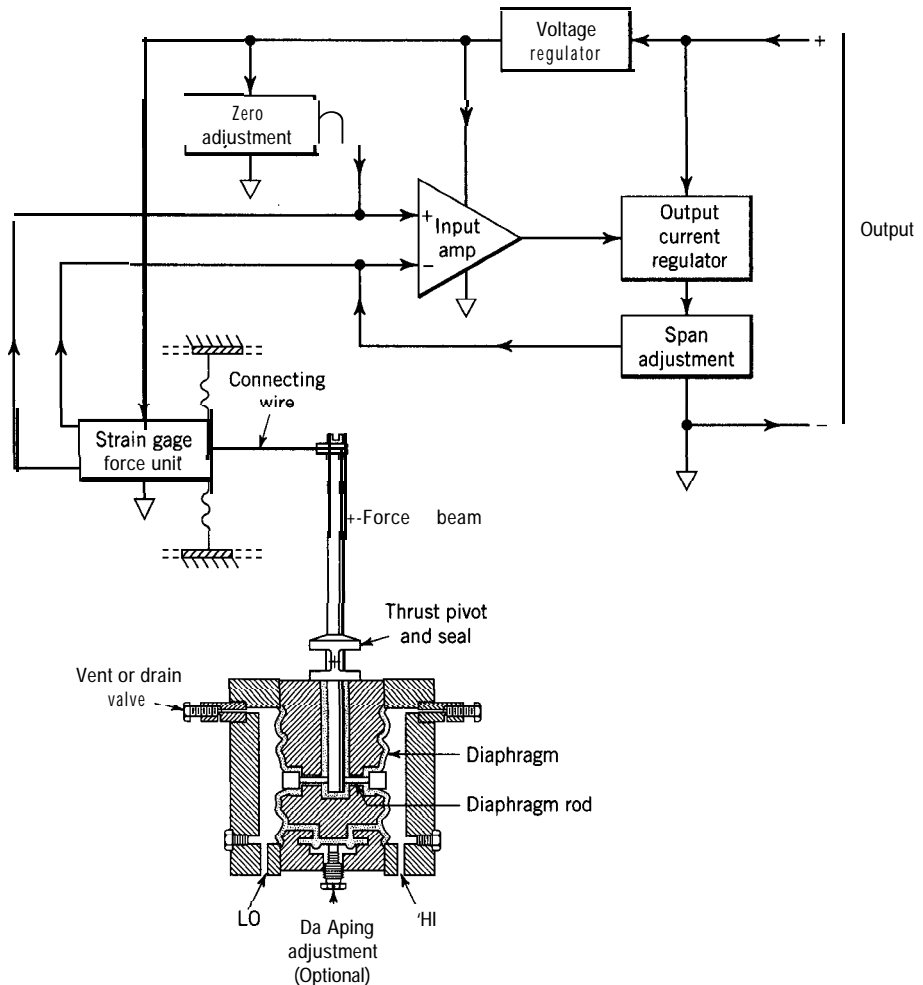
Figure   C-6.2   Electronic   differential   pressure   transmitter.   (Courtesy   of   ABB   Kent-Taylor.)

kind of flapper-nozzle arrangement to produce an output signal. This is a reliable and simple technique that has proved very successful for many years.

C-6.2 Electronic Transmitter

Figure C-6.2 shows a simplified diagram of an electronic differential pressure transmitter (see Reference 13). This motion-balance transmitter will be used to illustrate the working principles of typical electronic instrumentation.

An increase in differential pressure, acting on the measuring element diaphragms, develops a force that moves the lower end of the force beam to the left. This motion of the force beam is transferred to the strain gage force unit through the connecting wire. The strain gage force unit contains four strain gages connected in a bridge configuration. Movement of the force beam causes the strain gages to change resistance. This change in resistance produces a differential signal that is proportional to the input differential pressure. This differential signal is applied to the inputs of the input amplifier. One side of the signal is applied to the noninverting input through the zero network. This zero network provides the zero adjustment for the transmitter.

The signal from the input amplifier drives the output current regulator. The current regulator controls the transmitter output current through the span network and the output current sense circuit. The span network provides the span adjustment for the transmitter. The signal from the span network is fed back to the input circuit through a buffer amplifier and is used to control the gain of the input circuit. If the transmitter output current increases above 20 mA D.C., then the voltage across the output current sense resistor turns on the output current limiter, which limits the output.

C-7 TYPES OF CONTROL VALVES

There are many different types of control valves on the market (see References 14 through 20), so it is difficult to classify them. However, we will classify them into two broad categories: reciprocating stem and rotating stem control valves.

C-7.1 Reciprocating Stem

Figure C-7.1 shows a typical reciprocating stem control valve. This particular valve is called a *single-seated sliding stem globe valve. Globe valves* are a family of valves characterized by a closure member that travels in a line perpendicular to the valve seat. They are used primarily for throttling purposes and general flow control. Figure C-7.1 also shows in detail the different components of the valve. The valve is divided into two general areas: the actuator and the body. The actuator is the part of the valve that converts the energy input to the valve into mechanical motion to increase or decrease the flow restriction. Figure *C-7.2a* shows a *double-seated sliding stem globe valve.* Double-seated valves can handle high process pressure with a standard actuator. However, when tight shut-off is required, single-seated valves are generally used. Double-seated valves tend to have greater leakage when closed than single-seated valves.

Another type of body in common use is the *split-body valve* shown in Fig. C-7.2*b*. This type of body is frequently used in process lines where frequent changes of plug and seat are required because of corrosion.
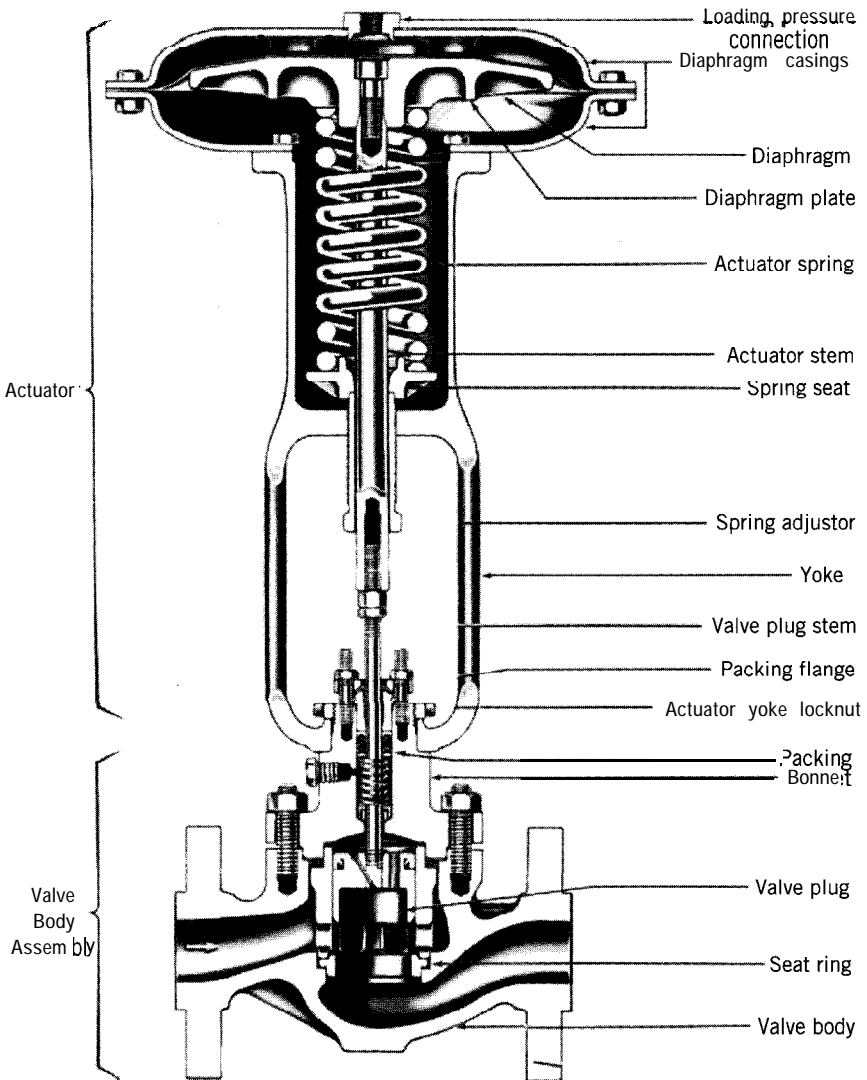
Figure C-7.1 Single-seated sliding stem globe valve. (Courtesy of Fisher Controls.)

Cage valves have hollow plugs with internal passages. The valve shown in Fig. 7-2.1 is a cage valve.

Three-way valves, shown in Fig. C-7.2c, are also reciprocating stem control valves. Three-way valves can be either diverging or converging, and consequently, they can either split one stream into two other streams or blend two streams into only one. They are sometimes used for control purposes.

There are some other types of reciprocating stem control valves. Most of them are used in specialized services. These include the Y-style *valve,* which is commonly used in molten metal or cryogenic service. *Pinch valves* or *diaphragm valves* consist of some
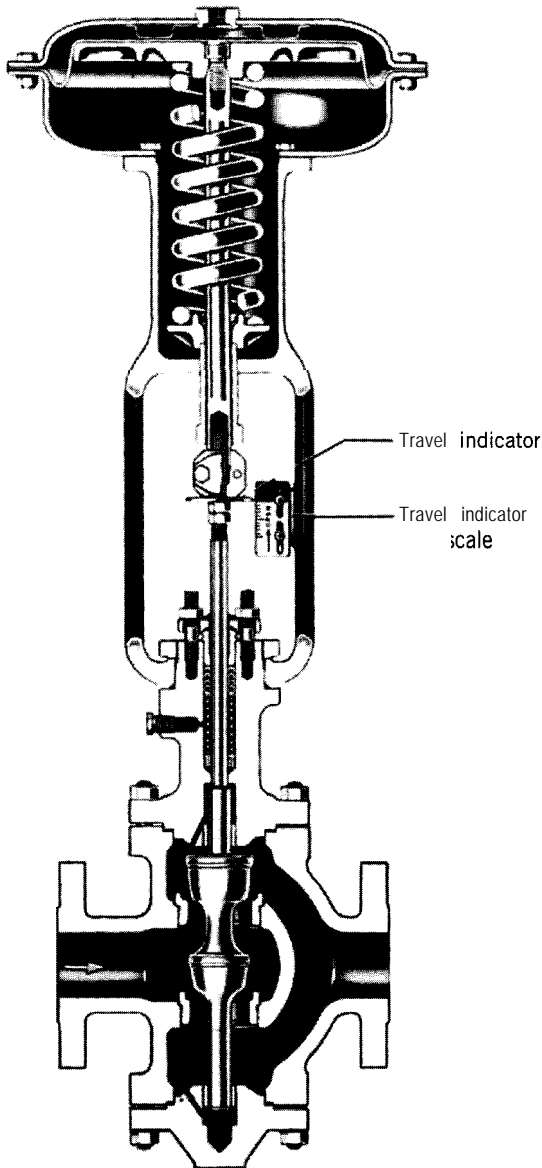
Figure C-7.2 (a) Double-seated sliding stem globe valve. (Courtesy of Fisher Controls.)

kind of flexure, such as a diaphragm, that can be moved together to open or close the area of flow. These valves are commonly used for highly corrosive fluids, slurries, and high-viscosity liquids, as well as in some food processing operations, such as the making of beer and wine. The *gate valve* is another type of reciprocating stem valve. It is used mainly as a block valve, for fully open or fully closed services. Gate valves are not used as automatic valves in throttling services.
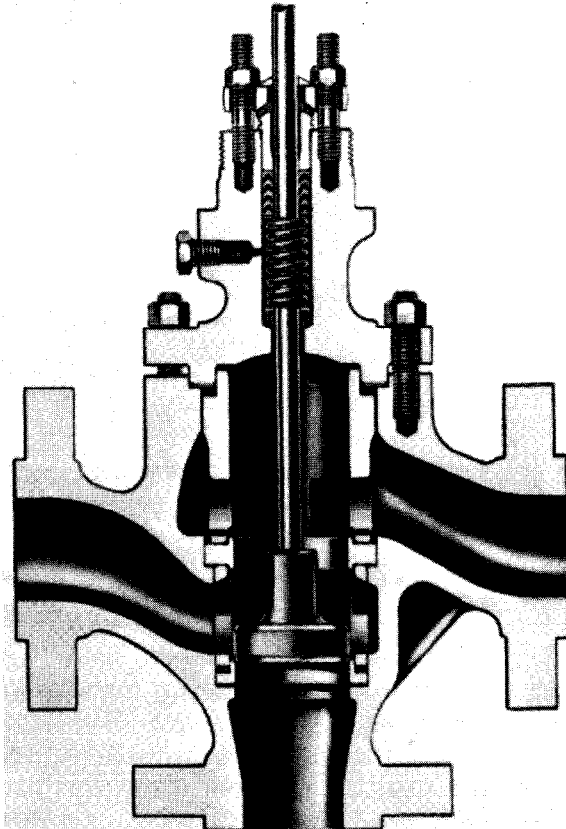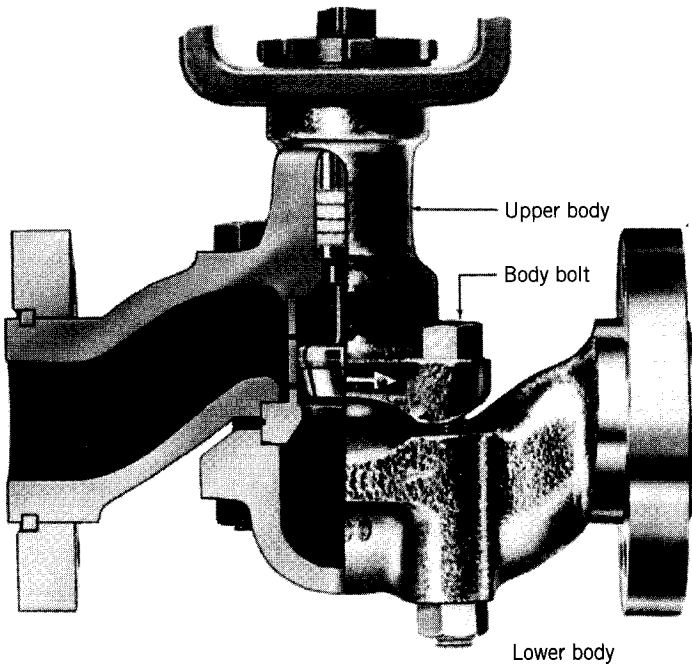
Figure C-7.2 *(Continued)* (b) Split-body valve. (Courtesy of Masoneilan Division, McGraw-Edison Co.) (c) Three-way valve. (Courtesy of Fisher Controls.)
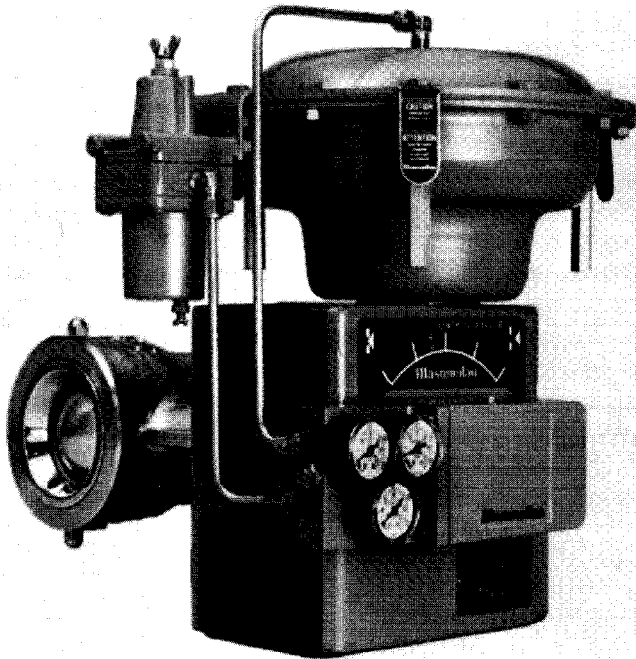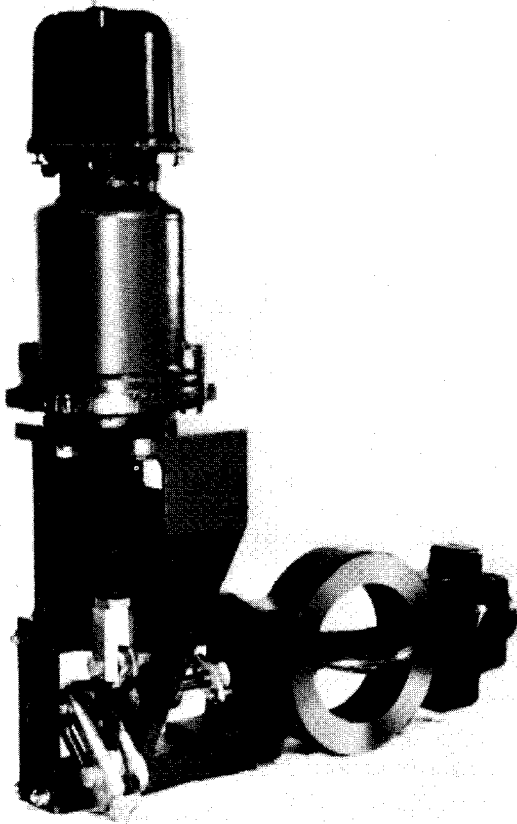
Figure C-7.3 (a) Butterfly valve. (Courtesy of Fisher Controls.)
(b) Ball valve with positioner. (Courtesy of Masoneilan Division,
McGraw-Edison Co.)

C-7.2 Rotating Stem

There are several very popular types of rotating stem valves. One of the most common is the *butterfly* valve, shown in Fig. C-7.3*a*. These valves consist of a disk rotating about a shaft. They require minimum space for installation and provide high-capacity flow at low cost.

Another common rotating stem *valve* is the *ball valve* shown in Fig. C-7.3*b*. Ball valves also provide high-capacity flow at low cost. They are commonly used to handle slurries or fibrous materials. They have low ieakage tendency and are small in size.

A very brief introduction to several types of control valves has been presented. However, these are by no means the only control valves, nor are they the only types of valves. There are a great number of valves available to meet requirements for specialized services as well as safety and other types of regulation.

C-8 CONTROL VALVE ACTUATORS

As previously defined, the actuator is the part of the valve that converts the energy input, either pneumatic or electrical, into mechanical motion to open or close the valve.

C-8.1 Pneumatically Operated Diaphragm Actuators

These are the most common actuators in the process industries. Figure C-8.1 shows a typical diaphragm actuator. These actuators consist of a flexible diaphragm placed between two casings. One of the chambers resulting from this arrangement must be made pressure-tight. The force generated within the actuator is opposed by a "range" spring. The controller air signal goes into the pressure-tight chamber, and an increase or decrease in air pressure produces a force that is used to overcome the force of the actuator's range spring and the forces within the valve body.

The action of the valve, FC or FO, is determined by the actuator. Figure C-8.1*a* shows a fail-closed or air-to-open valve. Figure C-8. lb shows a fail-open or air-to-close valve. Some valves can also have the action set at the body (reversed plug or cage) so that the stem always moves down. That is, in these cases the valve is either FC or FO when the stem moves down.

The size of the actuator depends on the process pressure against which it must move the stem and on the air pressure available. The most common air pressure range is 3 to 15 psig, but ranges of 6 to 30 psig and 3 to 27 psig are sometimes also used. These diaphragm actuators are simple in construction and also dependable and economical. Equations for sizing actuators are provided by manufacturers.

C-S.2 Piston Actuators

Piston actuators are normally used when maximum thrust output is required along with fast response. This usually occurs when working against high process pressure. These actuators operate using a high air pressure supply, up to 150 psig. The best designs are double-acting to give maximum thrust in both directions.
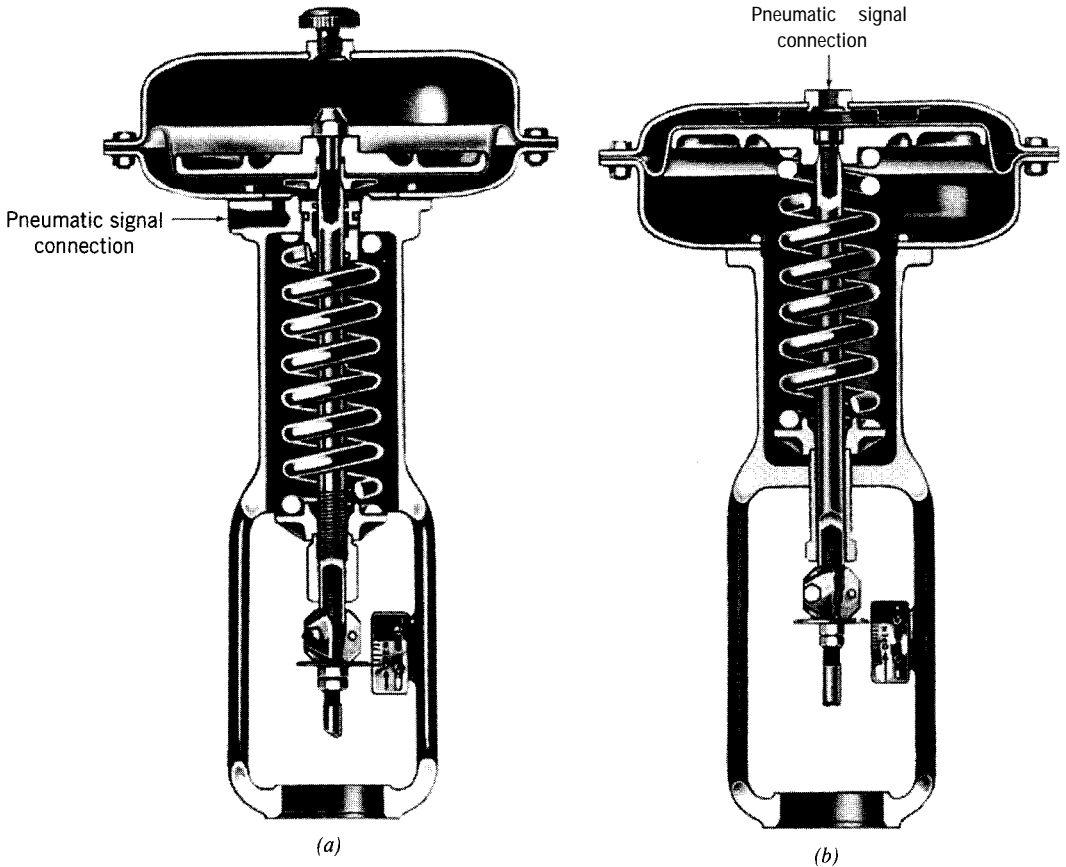
Figure C–8.1 (a) Diaphragm actuator: fail-closed, air-to-open. (b) Diaphragm actuator: fail-open, air-to-close. (Courtesy of Fisher Controls.)

C-8.3 Electrohydraulic and Electromechanical Actuators

Not so commonly used as the two previous types, electrohydraulic and electro-mechanical actuators are becoming more common with the use of electrical control signals. They require electric power to the motor and an electric signal from the controller.

In this family of actuators, the most common is probably the solenoid actuator. A solenoid valve can be used to actuate a double-acting piston actuator. By making or breaking an electric current signal, the solenoid switches the output of a connected hydraulic pump to either above or below the actuator piston. Accurate control of valve position can be obtained with this unit.

C-8.4 Manual-Handwheel Actuators

These actuators are used where automatic control is not required. They are available for reciprocating stem and rotary stem.

C-9 CONTROL VALVE ACCESSORIES

There are a number of devices, called accessories, that usually go along with control valves. This section presents a brief introduction to some of the most common of these accessories.

C-9.1 Positioners

A *positioner* is a device that acts very much like a proportional controller with very high gain. Its job is to compare the signal from the controller with the valve stem position. If the stem is not where the controller wants it to be positioned, then the positioner adds or exhausts air from the valve until the correct valve position is obtained. That is, when it is important to position the valve's stem accurately, a positioner is
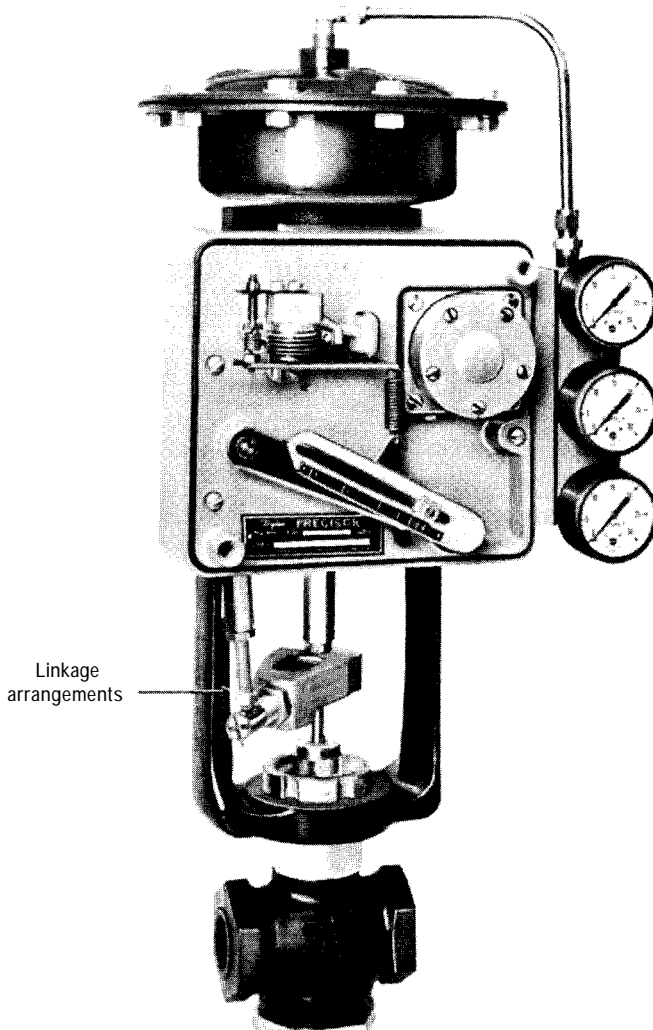
Figure C-9.1 Positioner installed in a valve. (Courtesy of ABB Kent-Taylor.)

normally used. Figure C-9.1 shows a valve with a positioner. The figure shows the bar-linkage arrangement by which the positioner senses the stem position. Another positioner is shown in Fig. C-7.3b.

The use of positioners tends to minimize the effects of

1. Lag in large-capacity actuators
2. Stem friction due to tight stuffing boxes
3. Friction due to viscous or gummy fluids
4. Changes in process line pressure
5. Hysteresis

Some control loops for which positioners are common are temperature, liquid level, concentration, and gas flow loops.

C-9.2 Boosters

Boosters, also called *air relays, are* used on valve actuators to speed up the response of the valve to a changing signal from a low-output-capacity pneumatic controller or transducer. It may also be noticed that for fast-responding control loops, such as liquid flow or liquid pressure, with which the use of positioners is discouraged, the use of boosters may be the proper choice (see Reference 14).

Boosters also have several other possible uses:

1. Amplify a pneumatic signal. Some typical amplification ratios are 1 : 2 and 1 : 3.
2. Reduce a pneumatic signal. Typical ratios are 5 : 1, 3 : 1, and 2 : 1.

C-9.3 Limit Switches

Limit switches are mounted on the side of the valves and are triggered by the position of the stem. These switches are generally used to drive alarms, solenoid valves, lights, or other such devices.

C-10 CONTROL VALVES-ADDITIONAL CONSIDERATIONS

This section presents a number of additional considerations to take into account when sizing and choosing a control valve. Thus this section complements Section 5-2.

Figures C-10.1a through C-10.1c show examples of manufacturer catalogs (Masoneilan and Fisher Controls). Once the C_v coefficient has been calculated using the equations presented in Chapter 5, these figures are used to the determine valve size.

C-lo.1   Viscosity   Corrections

Equation 5-2.1 does not take into consideration the effect of liquid viscosity in calculating the valve capacity, C,, coefficient. For liquids with the viscosity of water and light hydrocarbons, the viscous effects in valve capacity are negligible. However, for very viscous liquids, the viscous effects can lead to sizing errors.

Masoneilan (Reference 15) proposes calculating a turbulent C_v and a laminar C_v and then using the larger value as the required C,.
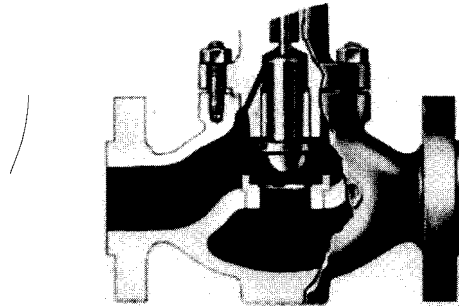
20000Series

ANSI Class **150-600** (Sch. 40)

Nominal Trim Size	¼	⅜	½	¾	1	1½	2	3	4	6	8	10
Orifice Dia. (in.)	.256	.375	.500	.750	.812	1.250	1.625	2.625	3.500	6.000	6.250	8.000

Valve Size (in.)	Reduced Trim			Full Capacity Trim								
¾	1.7	3.7	6.4	11								
1	1.7	3.7	6.4	11	12							
1½	1.7	3.8	6.6	12	13	25						
2	1.7	3.8	6.7	13	19*	26	46					
3				14		31	47	110				
4						32	49	113	195			
6							53	126	208	400		
8								133	224	415	640	
10									233	442	648	1000

Orifice Diameter .994   Refer to Bulletin 334E

(a)

Figure C-lo.1 (a) Example of Masoneilan's valve catalog. (Courtesy of Masoneilan Division, McGraw-Edison Co.)

Turbulent Flow

$$C_v = f\sqrt{\frac{G_f}{\Delta p}} \qquad \text{(C-10.1)}$$

Laminar Flow

$$C_v = 0\ 072\left(\frac{\mu f}{\Delta p}\right)^{2/3} \qquad \text{(C-10.2)}$$

where μ = viscosity, centipoise.

Fisher Controls (Reference 16) has developed a nomograph and procedure that provide a correction factor, F_v, that can be applied to the standard C_v coefficient to determine a corrected coefficient, C_{vr} .

Equal Percentage — Equal Percentage Characteristic

Coefficients	Body Size, In.	Port Diameter, In.	Total Travel, In.	10	20	30	40	50	60	70	80	90	100	K_m and C_f
C_v (Liquid)	1	1	3/4	.834	1.22	1.79	2.54	3.60	5.46	8.04	10.4	11.3	12.0	.92
	1-1/2	1-1/2	3/4	1.18	1.76	2.63	3.58	5.17	7.63	10.5	22.8	28.9	50.0	.85
	2	2	1-1/8	2.18	3.35	4.55	6.40	9.10	13.5	22.0	34.6	44.5	50.0	.94
	3	3	1-1/2	4.13	7.07	10.6	15.8	23.6	39.0	62.4	92.4	112	120	.83
	4	4	2	5.54	9.39	14.4	20.5	29.1	41.0	61.9	108	155	183	.86
	1*	3/4	3/4	.361	.597	.878	1.38	1.98	3.02	4.51	6.30	7.37	7.84	.97
	1-1/2*	3/4	3/4	.354	.559	.874	1.34	2.02	3.01	4.61	6.78	8.54	9.71	.96
		1	3/4	.803	1.13	1.61	2.50	3.58	4.95	8.20	12.3	15.7	16.9	.93
	2*	1	3/4	.772	1.17	1.65	2.47	3.53	5.38	8.20	11.2	13.7	14.9	.93
		1-1/2	3/4	1.11	1.64	2.42	3.46	4.86	7.22	12.0	17.4	24.0	29.1	.83
	3*	3/4	3/4	1.17	1.75	2.46	3.62	4.66	7.06	11.0	17.9	25.4	31.5	.79
		2	1-1/8	2.01	3.14	4.57	6.39	9.01	12.8	22.0	37.1	51.8	60.9	.84
	4*	2	1-1/8	2.18	3.27	4.61	6.28	8.78	12.8	23.1	40.2	53.8	64.1	.83
		3	1-1/2	3.47	6.12	9.66	14.8	22.6	37.1	62.2	92.5	117	136	.84
C_g (Gas)	1	1	3/4	27.0	37.5	53.0	77.6	112	162	242	344	434	482	40.1
	1-1/2	1-1/2	3/4	40.7	56.5	81.3	112	162	221	312	514	784	992	34.3
	2	2	1-1/8	72.9	111	149	207	293	392	639	1100	1580	1800	36.0
	3	3	1-1/2	127	219	317	471	706	1160	1970	3000	3870	4280	35.7
	4	4	2	176	293	445	619	838	1170	1810	3300	5960	6650	36.3
	1*	3/4	3/4	12.9	19.9	27.0	42.4	59.9	91.0	136	201	265	331	42.2
	1-1/2*	3/4	3/4	12.3	18.2	27.9	42.2	62.8	94.3	143	210	290	367	37.8
		1	3/4	20.4	32.0	48.4	73.4	108	162	244	358	463	617	36.5
	2*	1	3/4	27.1	38.5	53.0	77.4	111	162	236	348	466	582	39.1
		1-1/2	3/4	37.0	54.5	80.8	110	152	202	316	495	778	973	33.4
	3*	1-1/2	3/4	38.8	56.4	80.1	113	146	227	328	527	780	1040	33.0
		2	1-1/8	66.3	102	142	196	271	388	622	1060	1590	2040	33.5
	4*	2	1-1/8	65.1	107	147	193	281	386	617	1050	1530	2100	32.8
		3	1-1/2	114	200	305	450	704	1100	1850	2950	3960	4920	36.2
C_s (Steam)	1	1	3/4	1.35	1.88	2.65	3.88	5.60	8.10	12.1	17.2	21.7	24.1	40.1
	1-1/2	1-1/2	3/4	2.04	2.83	4.07	5.60	8.10	11.1	15.6	25.7	39.2	49.6	34.3
	2	2	1-1/8	3.65	5.55	7.45	10.4	14.7	19.6	32.0	55.0	79.0	90.0	36.0
	3	3	1-1/2	6.35	11.0	15.9	23.6	35.3	58.0	98.5	150	194	214	35.7
	4	4	2	8.80	14.7	22.3	31.0	41.9	58.5	90.5	165	298	333	36.3
	1*	3/4	3/4	.645	.995	1.35	2.12	3.00	4.55	6.80	10.1	13.3	16.6	42.2
	1-1/2*	3/4	3/4	.615	.910	1.40	2.11	3.14	4.72	7.15	10.5	14.5	18.4	37.8
		1	3/4	1.02	1.60	2.42	3.67	5.40	8.10	12.2	17.9	23.2	30.9	36.5
	2*	1	3/4	1.36	1.93	2.65	3.87	5.55	8.10	11.8	17.4	23.3	29.1	39.1
		1-1/2	3/4	1.85	2.73	4.04	5.50	7.60	10.1	15.8	24.8	38.9	48.7	33.4
	3*	1-1/2	3/4	1.94	2.82	4.01	5.65	7.30	11.4	16.4	26.4	39.0	52.0	33.0
		2	1-1/8	3.32	5.10	7.10	9.80	13.6	19.4	31.1	53.0	79.5	102	33.5
	4*	2	1-1/8	3.26	5.35	7.35	9.65	14.1	19.3	30.9	52.5	76.5	105	32.8
		3	1-1/2	5.70	10.0	15.3	22.5	35.2	55.0	92.5	148	198	246	36.2

*This column lists the K_m values for the C_v coefficients and the C_f values for the C_g and C_s coefficients at 100% travel.
* Restricted Trim

Linear — Linear Characteristic

Coefficients	Body Size, In.	Port Diameter, In.	Total Travel, In.	10	20	30	40	50	60	70	80	90	100	K_m and C_f
C_v (Liquid)	1	1	3/4	2.01	3.52	4.81	5.96	7.46	8.93	10.1	11.0	11.8	12.4	.90
	1-1/2	1-1/2	3/4	4.20	7.93	11.7	15.6	19.7	23.7	27.1	30.7	32.8	33.6	.87
	2	2	1-1/8	5.96	11.7	17.6	23.6	29.5	35.7	42.0	48.6	51.0	54.4	.96
	3	3	1-1/2	16.7	32.2	47.2	63.4	78.0	91.2	102	112	117	120	.85
	4	4	2	20.5	38.2	55.3	60.3	96.5	124	151	173	191	201	.84
	1-1/2*	1	3/4	1.90	3.32	4.80	5.93	7.58	9.00	10.6	12.6	14.7	16.2	.90
	2*	1	3/4	1.81	3.28	4.76	6.24	7.75	9.30	10.8	12.3	13.8	15.1	.90
		1-1/2	3/4	4.39	7.93	11.7	15.5	19.3	22.9	27.0	30.2	33.0	35.1	.81
	3*	1-1/2	3/4	4.24	7.97	11.7	15.4	19.3	23.2	27.2	31.3	35.1	38.5	.80
		2	1-1/8	5.63	11.4	17.7	24.0	30.8	37.6	47.5	57.7	65.1	68.7	.81
	4*	2	1-1/8	5.45	11.3	17.7	24.6	31.9	39.9	52.0	59.7	69.7	76.8	.81
		3	1-1/2	14.8	30.1	44.9	59.8	75.3	91.9	109	125	133	136	.87
C_g (Gas)	1	1	3/4	66.6	113	159	197	247	294	334	376	430	469	37.8
	1-1/2	1-1/2	3/4	131	251	371	507	633	762	888	1010	1120	1190	35.4
	2	2	1-1/8	173	365	553	753	960	1180	1440	1660	1840	1920	37.4
	3	3	1-1/2	512	999	1520	2040	2510	3000	3440	3850	4150	4380	36.4
	4	4	2	610	1210	1830	2490	3150	3870	4970	6290	7090	7480	37.2
	1-1/2*	1	3/4	51.6	99.7	147	195	242	289	335	395	492	573	35.4
	2*	1	3/4	55.9	99.8	146	195	242	291	343	388	473	570	37.7
		1-1/2	3/4	129	249	367	497	614	739	864	980	1110	1240	35.3
	3*	1-1/2	3/4	122	247	362	489	614	742	864	986	1130	1280	33.2
		2	1-1/8	172	327	520	763	951	1180	1450	1790	2180	2420	35.2
	4*	2	1-1/8	184	357	557	761	974	1220	1520	1840	2200	2590	33.7
		3	1-1/2	482	963	1470	1960	2440	2920	3460	4060	4710	5150	37.9
C_s (Steam)	1	1	3/4	3.33	5.65	7.95	9.85	12.4	14.7	17.2	18.8	21.5	23.5	37.8
	1-1/2	1-1/2	3/4	6.55	12.6	18.6	25.4	31.7	38.1	44.4	50.5	56.0	59.5	35.4
	2	2	1-1/8	8.65	18.3	27.7	37.7	48.0	59.0	72.0	83.0	92.0	96.0	37.4
	3	3	1-1/2	25.6	50.0	76.0	102	126	150	172	193	208	219	36.4
	4	4	2	30.5	60.5	91.5	125	158	194	249	315	355	374	37.2
	1-1/2*	1	3/4	2.58	4.99	7.35	9.75	12.1	14.5	16.8	19.8	24.6	28.7	35.4
	2*	1	3/4	2.80	4.99	7.30	9.75	12.1	14.6	17.2	19.4	23.7	28.5	37.7
		1-1/2	3/4	6.45	12.5	18.4	24.9	30.7	37.0	43.2	49.0	55.5	62.0	35.3
	3*	1-1/2	3/4	6.10	12.4	18.1	24.5	30.7	37.1	43.2	49.3	56.5	64.0	33.2
		2	1-1/8	8.60	16.4	26.0	38.2	47.6	59.0	72.5	89.5	109	121	35.2
	4*	2	1-1/8	9.20	17.9	27.9	38.1	48.7	61.0	76.0	92.0	110	130	33.7
		3	1-1/2	24.1	48.2	73.5	98.0	122	146	173	203	236	258	37.9

*This column lists the K_m values for the C_v coefficients and the C_f values for the C_g and C_s coefficients at 100% travel.
* Restricted Trim

Figure **C-lo.1** (Continued) (b) Example of Fisher's valve catalog. (Courtesy of Fisher Controls.) (c) Example of Fisher's valve catalog. (Courtesy of Fisher Controls.)

C-10.2 Flashing and Cavitation

The presence of either flashing or cavitation in a control valve can have significant effects on the operation of the valve and on the procedure for sizing it. It is important to understand the meaning and significance of these two phenomena. Figure C-lo.2 shows the pressure profile of a liquid flowing through a restriction (possibly a control valve).

To maintain steady-state mass flow, the velocity of the liquid must increase as the cross-sectional area for flow decreases. The liquid velocity reaches its maximum at a point just past the minimum cross-sectional area (the port area for a control valve). This point of maximum velocity is called the vena contracta. At this point, the liquid also experiences the lowest pressure. What happens is that the increase in velocity (kinetic energy) is accompanied by a decrease in "pressure energy." Energy is transformed from one form to another.

As the liquid passes the vena contracta, the flow area increases and the fluid velocity decreases and, in so doing, the liquid recovers part of its pressure. Valves such as butterfly valves, ball valves, and most rotary valves have a high-pressure recovery characteristic. Most reciprocating stem valves show a low-pressure recovery characteristic. The flow path through these reciprocating stem valves is more tortuous than through rotary type valves.

Looking again at Fig. C-10.2, let us suppose that the vapor pressure of the liquid at the flowing temperature is P_v. When the pressure of the liquid falls below its vapor pressure, some of the liquid starts changing phase from the liquid phase to the vapor phase. That is, the liquid flashes, and it can cause serious erosion damage to the valve plug and seat.

Aside from the physical damage to the valve, flashing tends to lower the flow capacity of the valve. As bubbles start forming, this tends to cause a "crowding condition" at the valve, which limits the flow. Furthermore, this crowding condition may get bad enough to "choke" the flow through the valve. That is, beyond this choked condition, increases in pressure drop across the valve will not result in an increased flow. It is important to recognize that the valve equation, Eq. 5-2.1, does not describe this condition. As the pressure drop increases, the equation predicts higher flow rates. This relationship is shown graphically in Fig. C-10.3, along with the choked-flow condition.
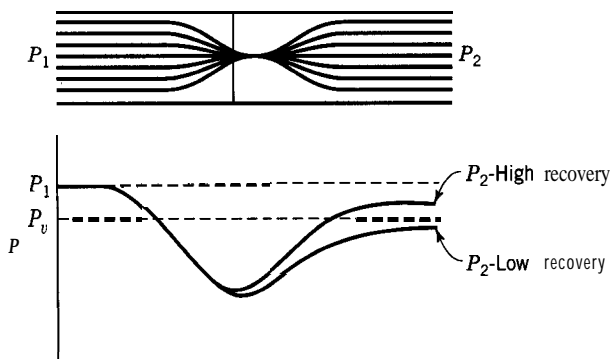
Figure C-lo.2 Pressure profile of a liquid across a restriction.
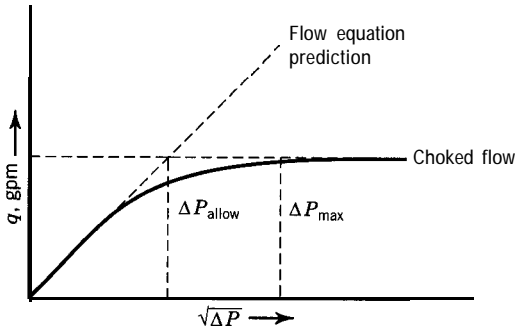
Figure   C-lo.3   Chocked-flow   condition.

Note from this figure that it is important for the engineer to know what maximum pressure drop, ΔP_{max}, is effective in producing flow. Instead of providing an equation for ΔP_{max}, manufacturers have chosen to provide an equation for ΔP_{allow} and to use this term to indicate when choke flow occurs. At higher pressure drops than ΔP_{allow}, choked flow results. ΔP_{allow} is a function not only of the fluid but also of the type of valve. Masoneilan (Reference 18) proposes the following equation:

$$\Delta P_{allow} = C_f^2 \Delta P_s \qquad\qquad \text{(C-10.3)}$$

and

$$\Delta P_s = P_1 - \left(0.96 - 0.28 \sqrt{\frac{P_v}{P_c}}\right) P_v \qquad\qquad \text{(C-10.4)}$$

or, if $P_v < 0.5P_1$, then

$$AP, = P_1 - P_v \qquad\qquad \textbf{(C-10.5)}$$

where

P_v = vapor pressure of liquid in psia
C_f = critical flow factor (see Reference 15)
P_c = critical pressure of liquid in psia

The critical flow factor, C_f, is shown in Fig. C-lo.4 for different types of valves. These values are the result of flow tests performed on the valves.
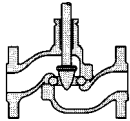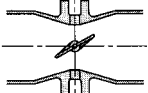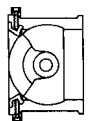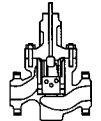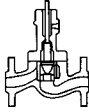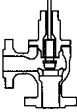
Fisher Controls (Reference 16) proposes the following equation for ΔP_{allow}:

$$\Delta P_{allow} = K_m(P_1 - r_c P_v) \qquad\qquad \text{(C-10.6)}$$

where

K_m = valve recovery coefficient (see Reference 16)
r_c = critical pressure ratio (see Reference 16)

The K_m coefficient depends on the type of valve and is also a result of flow tests. Figures C-lO.lb and C-10.1c show in the last column values of K_m for the particular type of valve. The r_c term is determined from Fig. C-10.5.

Valve Type	Trim Size	Flow To	C_f (F_L)	K_c*	C_{fr} (F_{LP}) D/d = 1.5 or greater	X_T
Split body globe valves	A	→ Close / Open ←	.80 .75	.51 .46	.77 .72	.54 .47
	B	→ Close / Open ←	.80 .90	.52 .65	.80 .89	.54 .68
37000 Series	A	Flow in Either Direction	.65	.32	.60	.35
Control ball valve	A	← Open	.60	.24	.55	.30
40000 Series balanced	1½"–4"	→ Close	.94	.71	.87	.74
	6"–16"	Close →	.92	.68	.89	.71
40000 Series unbalanced	A	← Open	.90	.65	.79	.68
	B	← Open	.90	.65	.86	.68
70000 Series	A	→ Close / Open ←	.81 .89	.53 .64	.78 .85	.55 .67
	B	→ Close / Open ←	.80† .90	.52† .65	.80 .90	.54 .68

(A) Full capacity trim, orifice dia. ≈ .8 valve size.

(B) Reduced capacity trim 50% of (A) and below.

† With Venturi Liner $C_f = 0.50$, $K_c = 0.19$.

Note $X_T = 0.84\ C_f^2$

Figure C-lo.4 Critical flow factor, C_f, at full opening. (Courtesy of Masoneilan Division, McGraw-Edison Co.)
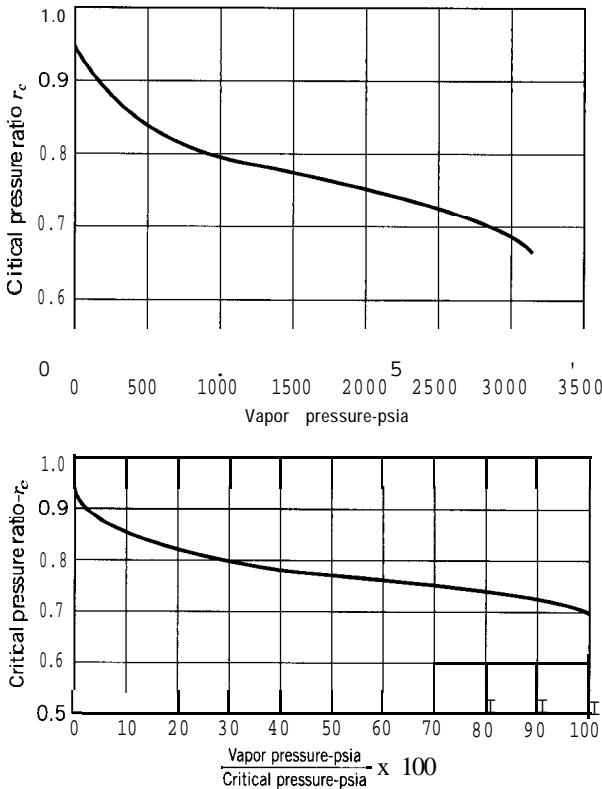
Figure C-lo.5 Fisher's critical pressure ratio. (Courtesy of Fisher Controls.)

If the pressure recovery experienced by the liquid is enough to raise the pressure above the vapor pressure of the liquid, then the vapor bubbles start collapsing, or imploding. This implosion is called cavitation. The energy released during cavitation produces noise, as though gravel were flowing through the valve (see Reference 14) and tears away the material of the valve. High-pressure recovery valves and rotary stem valves tend to experience cavitation more often than low-pressure recovery valves and reciprocating stem valves.

Tests have shown that for low-pressure recovery valves, such as rotary valves, choked flow and cavitation occur at nearly the same AP, so Eqs. C-lo.3 and C-lo.6 can also be used to calculate the pressure drop at which cavitation starts. For high-pressure recovery valves, cavitation can occur at pressure drops below ΔP_{allow}. For these types of valves, Masoneilan (see Reference 15) proposes the following equation:

$$\Delta p_{\text{cavitation}} = K_c(P_1 - P_v) \tag{C-10.7}$$

where K_c = coefficient of incipient cavitation, shown in Fig. C-10.4. Fisher Controls (see Reference 19) proposes the same equation, using the K_c term shown in Fig. C-10.6.
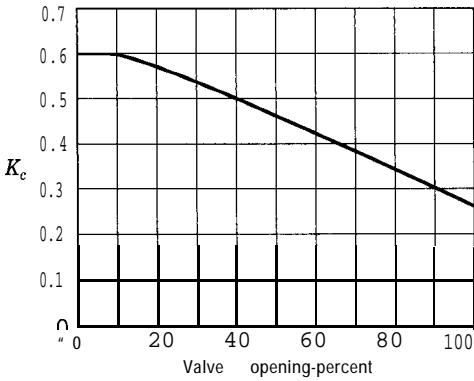
Figure C-lo.6 Fisher's coefficient of incipient cavitation. (Courtesy of Fisher Controls.)

Valve manufacturers produce special anticavitation trims that tend to increase the K_c term of the valve and, therefore, the pressure drop at which cavitation occurs.

C-11 SUMMARY

The purpose of this appendix was to introduce the reader to some of the instrumentation most commonly used for process control. The instrumentation shown included some of the hardware necessary for the measurement of process variables (primary elements) such as flow, pressure, temperature, and pressure. Two types of transmitters were also presented and their working principles discussed. Finally, some common types of valves (final control elements) used to take action were presented, along with their flow characteristics.

It is impossible to discuss in this book all of the details related to the different types of instruments; however, entire handbooks and an exhaustive collection of articles are available for this purpose. The reader is referred to the fine references listed at the end of this appendix. In addition to the many different types of instruments available today, new types of primary elements, transmitters, and final control elements are introduced on the market every month. In the primary elements area, new sensors that can measure difficult variables (such as concentration) more exactly, more repeatably, and faster are developed constantly.

In the transmitter area, the important phrase is *smart transmitters.* These are transmitters that, with the aid of microprocessors, present information to the controllers in a more readily understandable manner. The final control elements present another very active area of research. Not only are pneumatic control valves continually being upgraded, but electric actuators are also being developed and improved to allow interfacing with other electronic components such as controllers and computers. Other final control elements, such as drivers for variable-speed pumps and fans, are continuously being developed. The impetus behind this development is energy conservation. Lack of space prevents our examining the feasibility and justification of the use of these variable-speed pumps and fans for flow throttling. The reader is referred to References 21, 22, 23, and 24.

Certainly, the previous paragraph has shown that there is a lot of research being

conducted, principally by manufacturers in the instrumentation area, that should result in better measurement and control. This is one reason why process control is such a dynamic field.

REFERENCES

1. Baumann, H. D. 198 1. "Control Valve vs. Variable-Speed Pump." *Chemical Engineering,* June 24.
2. Casey, J. A., and D. Hammitt. 1978. "How to Select Liquid Flow Control Valves." *Chemical Engineering,* April 3.
3. Chalfin, S. 1974. "Specifying Control Valves." *Chemical Engineering,* October 14.
4. Considine, D. M., ed. 1961. *Handbook of Instrumentation and Controls. New* York: McGraw-Hill.
5. "Control Valve Handbook." Fisher Controls Co., Marshalltown, Iowa.
6. Creason, S. C. 1978. "Selection and Care of pH Electrodes." *Chemical Engineering,* October 23.
7. Dahlin, E., and A. Franci. 1985. "Practical Applications of a New Coriolis Mass Flowmeter." Paper presented at Control Expo '85, Control and Engineering Conference and Exposition, Rosemont, Ill., May 21-23.
8. "Fisher Catalog 10." Fisher Controls Co., Marshalltown, Iowa.
9. Fischer, K. A., and D. J. Leigh. 1983. "Using Pumps for Flow Control." *Instruments and Control Systems,* March.
10. Foster, R. 1975. "Guideline for Selecting Outline Process Analyzers." *Chemical Engineering,* March 17.
11. "Handbook Flowmeter Orifice Sizing," Handbook No. 10B9000, Fischer Porter Co., Warminster, Pa.
12. Hutchison, J. W., ed. "ISA Handbook of Control Valves." Research Triangle Park, N.C.: Instrument Society of America.
13. Jarc, D. A., and J. D. Robechek. 1981. "Control Pumps with Adjustable Speed Motors Save Energy." *Instruments and Control Systems,* May.
14. Kern, R. 1975. "Control Valves in Process Plants." *Chemical Engineering,* April 14.
15. Kern, R. 1975. "How to Size Flowmeters." *Chemical Engineering,* March 3.
16. Kern, R. 1975. "Measuring Flow in Pipes with Orifices and Nozzles." *Chemical Engineering,* February 3.
17. Liptak, B. *G.,* ed. 1995. *Instrument Engineers' Handbook.* Vol. I, 3rd ed. *Process Measurement.* New York: Chilton Book Co.
18. "Masoneilan Handbook for Control Valve Sizing." Masoneilan International, Inc., Norwood, Mass.
19. Ottmers, D. M., *et al.* 1979. "Instruments for Environment Monitoring." *Chemical Engineering,* October 15.
20. Perry, R., and D. W. Green. 1984. *Perry's Chemical Engineering Handbook.* 6th ed. New York: McGraw-Hill.
21. Pritchett, D. H. 1981. "Energy-Efficient Pump Drive." *Chemical Engineering Progress,* October.
22. "Process Control Instrumentation." Publication 105A-15M-4/71. Foxboro, Mass.: The Foxboro Co.
23. Ryan, J. B. 1975. "Pressure Control." *Chemical Engineering,* February 3.
24. Smith, C. L. 1978. "Liquid Measurement Technology." *Chemical Engineering,* April 3.
25. Spink, L. K. 1976. "Principles and Practice of Flow Meter Engineering." 9th ed. Foxboro, Mass.: The Foxboro Co.

26. Taylor Instrument Co., Rochester NY, "Differential Pressure Transmitter Manual, **1B-12B215.**"
27. Utterback, V. C. 1976. "Online Process Analyzers." *Chemical Engineering*, June 21.
28. Wallace, L. M. 1976. "Sighting In on Level Instruments." *Chemical Engineering*, February 16.
29. Young, A. M. 1985. "Coriolis-Based Mass Flow Measurement." *Sensors*, December.
30. Zientara, D. E. 1972. "Measuring Process Variables." *Chemical Engineering*, September 11.

Index

ISBN 0-471-57588-7

90000>

9 780471 575887