Optimal Control

PRENTICE HALL INFORMATION
AND SYSTEM SCIENCES SERIES

Thomas Kailath, Editor

Optimal Control

LINEAR QUADRATIC METHODS

BRIAN D. O. ANDERSON

JOHN B. MOORE

Department of Systems Engineering
Australian National University, Canberra

Contents

3

The Standard Regulator Problem—II   *35*

4

Tracking Systems   *68*

Part II. Properties and Application of the Optimal Regulator

5

Properties of Regulator Systems with a Classical Control Interpretation   *101*

6

Asymptotic Properties and Quadratic Weight Selection 139

7

State Estimator Design 164

8

System Design Using State Estimators 207

9

Frequency Shaping 262

—— **10** ————————————————————————————

Controller Reduction   *289*

—— **11** ————————————————————————————

Digital Controllers   *323*

Appendices   336

Preface

Despite the development of a now vast body of knowledge known as modern control theory, and despite some spectacular applications of this theory to practical situations, it is quite clear that some of the theory has yet to find application, and many practical control problems have yet to find a theory that will successfully deal with them. No one book, of course, can remedy the situation. The aim of this book is to construct bridges that are still required for the student and practicing control engineer between the familiar classical control results and those of modern control theory. It attempts to do so by consistently adopting the viewpoints that:

1. many modern control results have interpretation in terms of classical control notions;
2. many modern control results do have practical engineering significance, as distinct from applied mathematical significance;
3. to achieve practical designs, classical design insights and modern control tools are synergistic.

As a consequence, linear systems are very heavily emphasized, and the discussion of some results deemed fundamental in the general theory of optimal control has been kept to the barest minimum, thereby allowing emphasis on those particular optimal control results having application to linear systems. It may therefore seem strange to present a book on optimal control which does not expound in detail the Pontryagin Minimum Principle, but it is nonetheless consistent with the general aims of the book.

In selecting the material for the book, the aim has not been to locate optimal control theory of linear systems within the broader framework of optimal control theory per se. Rather, the aim is to present results of linear optimal control theory interesting from an engineering point of view, consistent with the ability of students to follow the material.

For the most part, continuous-time systems are treated, since engineering systems operate in continuous time, and a good deal more of the discussion is on time-invariant than is on time-varying systems. Infinite-time optimization problems for time-varying systems involve concepts such as uniform complete controllability, which the authors consider to be in the nature of advanced rather than core material, and accordingly discussion of such material is kept to a minimum. For completeness, some mention is also made of discrete-time systems, including implementation of continuous-time controller designs in discrete time, but it seemed to us that an extended discussion of discrete-time systems would involve undue repetition.

The text is aimed at the first or later year graduate student. The background assumed of any reader is, first, an elementary control course, covering such notions as transfer functions, Nyquist plots, root locus, etc.; second, an elementary introduction to the state-space description of linear systems and the dual notions of complete controllability and complete observability; and third, an elementary introduction to linear algebra. However, exposure to a prior or concurrent course in optimal control is not assumed.

The book contains three major parts. Part One introduces and outlines the basic theory of the linear regulator/tracker for time-invariant and time-varying systems, emphasizing the former. The actual derivation of the optimal control law is via the Hamilton-Jacobi equation which is introduced using the Principle of Optimality. The infinite-time problem is considered. Part Two outlines the engineering properties of the regulator. Degree of stability, phase and gain margin, tolerance of time delay, effect of nonlinearities, asymptotic properties and various sensitivity problems are all considered. Part Three considers state estimation and robust controller design using state estimate feedback. Loop transmission recovery, frequency shaping, and techniques of controller reduction and implementation are considered.

The problems at the end of each section are, in the main, extensions and completions of theoretical material for which hints are often provided. The solutions are available in a Solutions Manual. There are also open-ended problems set in some chapters which require computer studies. No solutions are provided for these problems.

We would like to emphasize that the manuscript was compiled as a truly joint effort. We wish to acknowledge discussions with Boeing flight control engineers Dagfinn Gangsaas, Jim Blight, and Uy-Loi Ly. These people motivated us in revising our original manuscript of 1971, published as *Linear Optimal Control.*

For those readers who are familiar with our earlier work, *Linear Optimal Control,* we record briefly the main changes:

1. We have omitted material on relay control systems, dual-mode controllers, *and so-called specific optimal regulator problems.*

2. We have added material on second variation theory, frequency shaping, loop recovery, and controller reduction.

3. We have rewritten many sections, but especially the material on robustness and tracking.

4. We have included new appendices on the Pontryagin Principle, Lyapunov Stability, and Riccati equations, deleting from the main text a chapter on Riccati equations which appeared in *Linear Optimal Control.*

5. The title is changed to focus on linear-quadratic methods, as opposed to the so-called H^∞ and L_1 methods now being developed.

We are appreciative of the typists Kay Hearder, Dong Feng Li, Marilyn Holloway, and Dawn Jarrett for their efforts, particularly in reading our handwriting.

BRIAN D.O. ANDERSON
JOHN B. MOORE

Optimal Control

Part 1. Basic Theory of the Optimal Regulator

1

Introduction

1.1 LINEAR OPTIMAL CONTROL

The methods and techniques of what is now known as "classical control" will be familiar to most readers. In the main, the systems or plants that can be considered by using classical control ideas are linear and time invariant, and have a single input and a single output. The primary aim of the designer using classical control design methods is to stabilize a plant, whereas secondary aims may involve obtaining a certain transient response, bandwidth, disturbance rejection, steady state error, and robustness to plant variations or uncertainties. The designer's methods are a combination of analytical ones (e.g., Laplace transform, Routh test), graphical ones (e.g., Nyquist plots, Nichols charts), and a good deal of empirically based knowledge (e.g., a certain class of compensator works satisfactorily for a certain class of plant). For higher-order systems, multiple-input systems, or systems that do not possess the properties usually assumed in the classical control approach, the designer's ingenuity is generally the limiting factor in achieving a satisfactory design.

Two of the main aims of modern, as opposed to classical, control are to de-empiricize control system design and to present solutions to a much wider class of control problems than classical control can tackle. One of the major ways modern control sets out to achieve these aims is by providing an array of analytical design procedures that facilitate the design task.

In the early stages of a design, the designer must use his familiarity with the engineering situation, and understanding of the underlying physics, to formulate a

sensible mathematical problem. Then the analytical design procedures, often implemented these days with commercial software packages, yield a solution—which usually serves as a first cut in a trial and error iterative process.

Optimal control is one particular branch of modern control that sets out to provide analytical designs of a specially appealing type. The system that is the end result of an optimal design is not supposed merely to be stable, have a certain bandwidth, or satisfy any one of the desirable constraints associated with classical control, but it is supposed to be the *best* possible system of a particular type—hence, the word *optimal.* If it is both optimal and possesses a number of the properties that classical control suggests are desirable, so much the better.

Linear optimal control is a special sort of optimal control. The plant that is controlled is assumed linear, and the controller, the device that generates the optimal control, is constrained to be linear. Linear controllers are achieved by working with quadratic performance indices. These are quadratic in the control and regulation/tracking error variables. Such methods that achieve linear optimal control are termed Linear-Quadratic (LQ) methods. Of course, one may well ask: why linear optimal control, as opposed simply to optimal control? A number of justifications may be advanced; for example, many engineering plants are linear prior to addition of a controller to them; a linear controller is simple to implement physically, and will frequently suffice.

Other advantages of optimal control, when it is specifically linear, follow.

1. Many optimal control problems do not have computable solutions, or they have solutions that may be obtained only with a great deal of computing effort. By contrast, nearly all linear optimal control problems have readily computable solutions.

2. Linear optimal control results may be applied to nonlinear systems operating on a small signal basis. More precisely, suppose an optimal control has been developed for some nonlinear system with the assumption that this system will start in a certain initial state. Suppose, however, that the system starts in a slightly different initial state, for which there exists some other optimal control. Then a first approximation to the difference between the two optimal controls may normally be derived, if desired, by solving a linear optimal control problem (with all its attendant computational advantages). This holds independently of the criterion for optimality for the nonlinear system. (We list two references [1] and [2] that outline this important result.[†])

3. The computational procedures required for linear optimal design may often be carried over to nonlinear optimal problems. For example, the nonlinear optimal design procedures based on the theory of the second variation [1–3]

[†]References are located at the end of each chapter.

and quasilinearization [3, 4] consist of computational algorithms replacing the nonlinear problem by a sequence of linear problems.

4. Linear optimal control designs where the plant states are measurable turn out to possess a number of properties, other than simply optimality of a quadratic index, which classical control suggests are attractive. Examples of such properties are good gain margin and phase margin, and good tolerance of non-linearities. Such robustness properties can frequently be achieved even when state estimation is required. The robustness properties suggest that controller designs for nonlinear systems may sometimes be achieved by designing with the assumption that the system is linear (even though this may not be a good approximation), and by relying on the fact that an optimally designed linear system can tolerate nonlinearities—actually quite large ones—without impairment of all its desirable properties. Hence, linear optimal design methods are in some ways applicable to nonlinear systems.

5. Linear optimal control provides a framework for the unified treatment of the control problems studied via classical methods. At the same time, it vastly extends the class of systems for which control designs may be achieved.

Linear optimal control design for time-invariant systems is largely a matter of control law synthesis; see the flow chart of Figure 1.1-1 for the approach emphasized in this text. Recall that the designer's first task is to use his or her engineering understanding to formulate a mathematical problem. This is embodied in the top two blocks. If we disregard the iteration in this and later steps (the iterations are illustrated in the flowchart), there are three essential steps covered by the modern analytical procedures of this text. These are

   full-state feedback design (where it is assumed that all states are measured and available for feedback)

   state estimator design (where the concern is to estimate values of the states when they cannot all be measured directly, but certain measurements are available)

   controller reduction (where the concern is to approximate a complicated state estimate feedback controller obtained from the above two steps by a simpler one—complication usually being measured by the state dimension)

The final major stage of design, involving the implementation of the controller, may involve the derivation of a discrete-time approximation to the controller.

   In the second step (state estimator design), a variation is to estimate only the state feedback control signal, rather than the full state vector.

   Linear quadratic methods that from the start build in controller constraints such as controller order are dealt with only briefly in this text. For full details see, for example, [5, 6].
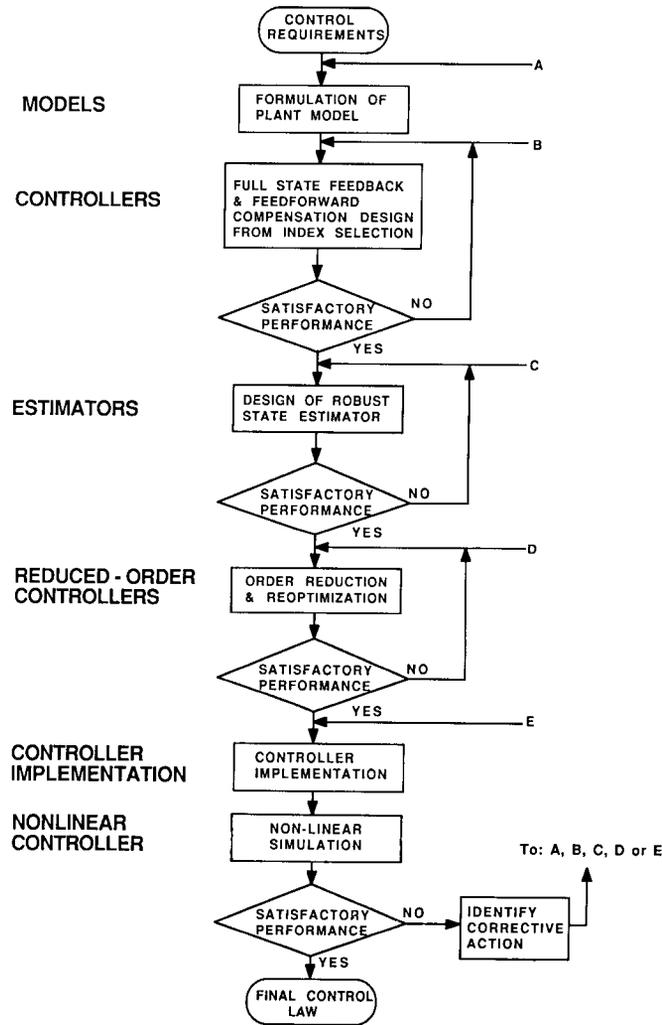
Figure 1.1-1   Control Law Synthesis Process

1.2 ABOUT THIS BOOK IN PARTICULAR

This is not a book on optimal control, but a book on optimal control via linear quadratic methods. Accordingly, it reflects very little of the techniques or results of general optimal control. Rather, we study a basic problem of linear optimal control, the "regulator problem," and attempt to relate mathematically all other problems discussed to this one problem. If the reader masters the mathematics of the regulator problem, he should find most of the remainder of the mathematics easy going.

We aim to analyze the engineering properties of the solution to the problems presented. We thus note the various connections to classical control results and ideas, which, in view of their empirical origins, are often best for providing a framework for a modern control design and assessing a practical design.

1.3 PART AND CHAPTER OUTLINE

In this section, we briefly discuss the breakdown of the book into parts and chapters. There are three parts, listed below with brief comments.

Part I—Basic theory of the optimal regulator.   These chapters serve to introduce the linear regulator problem and to set up the basic mathematical results associated with it. Chapter 1 is introductory. Chapter 2 sets up the problem by translating into mathematical terms the physical requirements on a regulator. It introduces the Principle of Optimality and the Hamilton–Jacobi equation for solving optimal control problems, and then obtains a solution for problems where performance over a finite (as opposed to infinite) time interval is of interest. The infinite-time interval problem is considered in Chapter 3, which includes stability properties of the optimal regulators, and shows how to achieve a regulator design with a prescribed degree of stability. Also considered is the formulation of an optimal linear regulator problem by linearization of a nonlinear system and computation of the second variation of an optimized index. Chapter 4 considers tracking problems by building on the regulator theory. In tracking, one generally wishes the plant output to follow a specific prescribed time function or a signal from a class, for example, a step function of unknown magnitude.

Part II—Properties of the optimal regulator.   In Chapter 5, frequency domain formulas are derived to deduce sensitivity and robustness properties. In particular, the return difference relation is studied along with its interpretation as a spectral factorization. Robustness measures in terms of sensitivity and complementary sensitivity functions are introduced, and for the multivariable case, the role of singular values is explored. Gain and phase margins and tolerance of sector nonlinearities are optimal regulator properties studied. The inverse problem of optimal control is briefly mentioned. In Chapter 6, the relationship between quadratic index weight selection and closed-loop properties is studied, with emphasis on the asymptotic properties as the control cost weight approaches infinity or zero.

Part III—State estimation and linear-quadratic-gaussian design.   When the states of a plant are not available, then the certainty equivalence principle suggests that state estimates be used instead of states in a state feedback design. Chapter 7 deals with state estimation, including the case when measurements are noisy, with, in the ideal case, additive gaussian noise.

Design methods and properties are developed, to some extent exploiting the theory of Parts I and II. Chapter 8 deals with control law synthesis using full state feedback designs and state estimation. It is pointed out that when the plant is precisely known, then in the linear-quadratic-gaussian (LQG) case, certainty equivalence is the optimal approach. This is the separation theorem. Otherwise, there can be poor robustness properties, unless loop recovery and frequency shaping techniques are adopted, as studied in Chapters 8 and 9, respectively. State estimate feedback designs, particularly when frequency shaped, may result in controllers of unacceptably high order. Controller reduction methods are studied in Chapter 10. These attempt to maintain controller performance and robustness properties while reducing controller complexity. Finally, in Chapter 11, some practical aspects concerning implementation of controllers via digital computers are studied.

   Appendices.   Results in matrix theory, linear system theory, the Minimum Principle, stability theory and Riccati equations relevant to the material in the book are summarized in the appendices.   .

REFERENCES

[1] J. V. Breakwell, J. L. Speyer, and A. E. Bryson, "Optimization and Control of Nonlinear Systems Using the Second Variation," *SIAM J. Control,* Vol. 1, No. 2 (1963), pp. 193–223.

[2] H. J. Kelley, "Guidance Theory and Extremal Fields," *IRE Trans. Auto. Control,* Vol. AC-7, No. 4 (October 1962), pp. 75–82.

[3] A. P. Sage, *Optimum Systems Control.* Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1968.

[4] R. E. Bellman and R. E. Kalaba, *Quasilinearization and Nonlinear Boundary Value Problems.* New York: Elsevier, 1965.

[5] W. S. Levine and M. Athans, "On the Determination of the Optimal Constant Output Feedback Gains for Linear Multivariable Systems," *IEEE Trans. Auto. Control,* Vol. AC-15, No. 1 (February 1970), pp. 44–48.

[6] D. C. Hyland and D. S. Bernstein, "The Optimal Projection Equations for Fixed-Order Dynamic Compensation," *IEEE Trans. on Auto. Control,* Vol. AC-29, No. 11 (November 1984), pp. 1034–1037.

2

The Standard Regulator Problem—1

2.1 A REVIEW OF THE REGULATOR PROBLEM

We shall be concerned almost exclusively with linear finite-dimensional systems, which frequently will also be time invariant. The systems may be represented by equations of the type

$$\dot{x}(t) = F(t)x(t) + G(t)u(t) \qquad (2.1\text{-}1)$$

$$y(t) = H'(t)x(t) \qquad (2.1\text{-}2)$$

Here, $F(t)$, $G(t)$, and $H(t)$ are matrix functions of time, in general with continuous entries. If their dimensions are respectively $n \times n$, $n \times m$, $n \times p$, the n vector $x(t)$ denotes the *system state* at time t, the m vector $u(t)$ the *system input* or *system control* at time t, and the p vector $y(t)$ the *system output* at time t. The superscript prime denotes matrix transposition.

In classical control work, usually systems with only one input and output are considered. With these restrictions in (2.1-1) and (2.1-2), the vectors $u(t)$ and $y(t)$ become scalars, and the matrices $G(t)$ and $H(t)$ become vectors, and accordingly will often be denoted by lowercase letters to distinguish their specifically vector character. The systems considered are normally also *time–invariant*. In terms of (2.1-1) and (2.1-2), this means that the input $u(t)$ and output $y(t)$ for an initially zero state are related by a time-invariant impulse response. Furthermore, the most common state-space descriptions of time-invariant systems are those where $F(t)$, $g(t)$, and $h(t)$ are constant with time. Note, though, that nonconstant $F(t)$, $g(t)$, and

$h(t)$ may still define a time-invariant impulse response—e.g., $F(t) = 0$, $g(t) = e^t$, $h(t) = e^{-t}$ defines a time-invariant impulse response via the map

$$y(t) = \int_{t_0}^{t} \exp\left[-(t - \tau)\right] u(\tau)\, d\tau$$

The classical description of a system is normally in terms of its transfer function matrix, which we denote by $W(s)$, s being the Laplace transform variable. The well-known connection between $W(s)$ and the matrices of (2.1-1) and (2.1-2), if these are constant, is

$$W(s) = H'(sI - F)^{-1}G \qquad (2.1\text{-}3)$$

A common class of control problems involves a *plant*, for which a control is desired to achieve one of the following aims:

1. *Qualitative statement of the regulator problem.* Suppose that initially the plant output, or any of its derivatives, is nonzero. Provide a plant input to bring the output and its derivatives to zero. In other words, the problem is to apply a control to take the plant from a nonzero state to the zero state. This problem may typically occur where the plant is subjected to unwanted disturbances that perturb its output (e.g., a radar antenna control system with the antenna subject to wind gusts).

2. *Qualitative statement of the tracking (or servomechanism) problem.* Suppose that the plant output, or a derivative, is required to track some prescribed function. Provide a plant input that will cause this tracking (e.g., when a radar antenna is to track an aircraft, such a control is required).

In a subsequent chapter, we shall discuss the tracking problem. For the moment, we restrict our attention to the more fundamental regulator problem; thus no external input is applied.

When considering the regulator problem using classical control theory, we frequently seek a solution that uses *feedback* of the output and its derivatives to generate a control. A *controller* with a transfer function description is interposed between the plant output and plant input. The plant output is the controller input, and the controller output is the plant input. The feedback arrangement is shown in Fig. 2.1-1. Both the plant and controller have a single input and output, and are time-invariant. Each possesses a transfer function.

In the optimal control approach of this text, it is assumed in the first instance that the plant states are available for measurement. If this is not the case, it is generally possible to construct a physical device called a *state estimator* driven by both the plant input and output. This produces at its output estimates of the plant states, and these may be used in lieu of the states. This will be discussed in a later chapter. In addition to assuming availability of the states, it is usual in the first instance to seek controllers that are nondynamic, or memoryless. In other words, the controller output or plant input $u(t)$ is assumed to be an instantaneous function
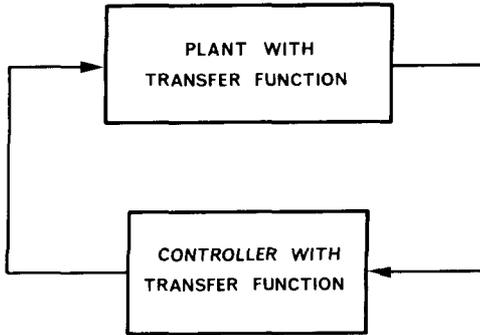
Figure 2.1-1   Classical feedback arrangement.

of the plant state $x(t)$. The nature of this function may be permitted to vary with time, in which case we could write down a *control law*

$$u(t) = k(x(t), t) \tag{2.1-4}$$

to indicate the dependence of $u(t)$ on both $x(t)$ and t.

Of interest from the viewpoint of ease of implementation is the case of the *linear control law*, given by

$$u(t) = K'(t)x(t) \tag{2.1-5}$$

for some matrix K of appropriate dimension. [When $K(t)$ is a constant matrix, (2.1-5) becomes a constant or time-invariant control law, and, as will become clear, a number of connections with the classical approach can be made.] Figure 2.1-2 shows the arrangement resulting from combining (2.1-5) with a state estimator for generating $x(t)$, or an estimate $\hat{x}(t)$ of $x(t)$. The plant is assumed to be linear, but may have multiple inputs and outputs and may be time varying. The state estimator constructs the plant state vector or an estimate of it from the input and output vectors, and is actually a linear, finite-dimensional system itself. Linear combinations of the states are fed back to the system input in accordance with (2.1-5).

When attempting to construct a controller for the regulator problem, we might imagine that the way to proceed would be to search for a control scheme that would take an arbitrary nonzero initial state to the zero state, preferably as fast as possible. Could this be done? If F and G are constant, and if the pair $[F, G]$ is
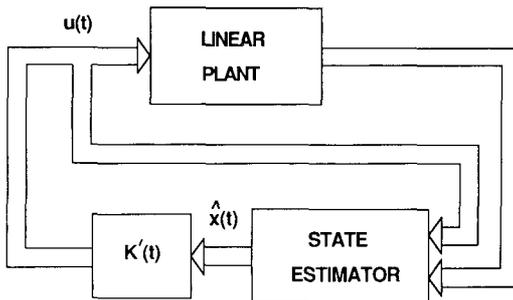
Figure 2.1-2   State estimate feedback arrangement.

completely controllable, the answer is certainly yes [1]. Recall (see Appendix B), that the definition of complete controllability requires that there be a control taking any nonzero state $x(t_0)$ at time t_0 to the zero state at some time T. In fact, if F and G are constant, T can be taken as close to t_0 as desired, and likewise for some classes of time-varying $F(t)$, $G(t)$. What, therefore, would be wrong with such a scheme? Two things. First, the closer T is to t_0, the greater is the amount of control energy (and the greater is the magnitude of the control) required to effect the state transfer. In any engineering system, an upper bound is set on the magnitude of the various variables in the system by practical considerations. Therefore, one could not take T arbitrarily close to t_0 without exceeding these bounds. Second, as reference to [1] shows, the actual control cannot be implemented as a linear feedback law for finite T, unless one is prepared to tolerate infinite entries in $K(T)$, that is, the controller gain at time T. In effect, a linear feedback control is ruled out.

Any other control scheme for which one or both of these objections is valid is equally unacceptable. In an effort to meet the first objection, one could conceive that it is necessary to keep some measure of control magnitude bounded or even small during the course of a control action; such measures might be

$$\int_{t_0}^{T} u'(t)\,u(t)\,dt, \qquad \int_{t_0}^{T} [u'(t)\,u(t)]^{1/2}\,dt, \qquad \max_{t\,\in\,[t_0,\,T]} \|u(t)\|,$$

or

$$\int_{t_0}^{T} u'(t)\,R(t)\,u(t)\,dt$$

where $R(t)$ is a positive definite matrix for all t, which, without loss of generality, can always be taken as symmetric. We shall discuss subsequently how to meet the second objection. Meanwhile, we shall make further adjustments to our original aim of regulation.

First, in recognition of the fact that "near enough is often good enough" for engineering purposes, we shall relax the aim that the system should actually achieve the zero state, and require merely that the state as measured by some norm should become small. If there were some fixed time T by which this was required, we might ask that $x'(T)Ax(T)$, with A some positive definite matrix, be made small.

Second, it is clearly helpful from the control point of view to have $\|x(t)\|$ small for any t in the interval over which control is being exercised, and we can express this fact by asking, for example, that $\int_{t_0}^{T} x'(t)Q(t)x(t)\,dt$ be small, where $Q(t)$ is symmetric positive definite. In some situations, as we shall see, it proves sufficient to have $Q(t)$ nonnegative definite.

In defining our optimal regulation task, we seek one that has engineering sense in that the appropriate quantities are penalized, is tractable of solution, and yields an optimal controller that is suitable for implementation, preferably linear. In static optimization, with quadratic indices and linear constraints, the linearity property is observed. For example, consider the minimization through choice of u_0 of the index $x_1'Qx_1 + u_0'Ru_0$ when $x_1 = Fx_0 + Gu_0$, with x_0 prescribed. A straightforward

calculation shows that the minimum is achieved with $u_0 = -(G'QG + R)^{-1}G'QFx_0$, which means that u_0 depends linearly on x_0.

All this suggests that we could get a linear control law if for the dynamic constraint (2.1-1) we seek to minimize the *quadratic performance index*

$$V(x(t_0), u(\cdot), t_0) = \int_{t_0}^{T} (u'Ru + x'Qx)\, dt + x'(T)Ax(T) \qquad (2.1\text{-}6)$$

As the notation implies, the value taken by V depends on the initial state $x(t_0)$ and time t_0, and the control over the interval $[t_0, T]$.

Let us consider the following as a formal statement of the regulator problem.

> ***Regulator problem.***   Consider the system (2.1-1), where the entries of $F(t)$ and $G(t)$ are assumed to be continuous. Let the matrices $Q(t)$ and $R(t)$ have continuous entries, be symmetric, and be nonnegative and positive definite, respectively. Let A be a nonnegative definite symmetric matrix. Define the *performance index* $V(x(t_0), u(\cdot), t_0)$ as in (2.1-6) and the *minimization problem* as the task of finding an *optimal control* $u^*(t)$, $t \in [t_0, T]$, minimizing V, and the associated *optimum performance index* $V^*(x(t_0), t_0)$—that is, the value of V obtained by using the optimal control.

With T finite, the problem is termed a finite time, or sometimes finite horizon, problem. With T infinite, an infinite time or infinite horizon problem is obtained. We shall postpone consideration of infinite time problems to the next chapter.

Notice that earlier it was suggested that A should be positive definite, whereas the statement of the regulator problem merely suggests that it should be nonnegative definite. As we shall see subsequently, the size of the final state $x(T)$ can frequently be made small merely by the relaxed requirement. Indeed, the choice $A = 0$ will often lead to a satisfactory result.

As already foreshadowed, the minimization of (2.1-6) turns out to be achievable with a linear feedback law. All of the other nonquadratic measures suggested do not, in general, lead to linear feedback laws.

Before studying the minimization problem further, we note the following references. Books [2] and [3] are but two of a number of excellent treatments of the regulator problem, and, of course, optimal control in general. Several older papers dealing with the regulator problem could be read with benefit [4-7].

Main points of the section.   The quadratic performance index of the regulator problem penalizes nonzero states and controls. Its minimization will be shown to lead to a linear state feedback law, which may require a state estimator for its implementation.

Problem 2.1-1.   Consider the system

$$\dot{x} = F(t)x + G(t)u$$

with F, G possessing continuous entries. Show that there does not exist a control law

$$u = K'(t)x(t)$$

with the entries of $K(t)$ continuous, such that with arbitrary $x(t_0)$ and some finite T, $x(T) = 0$. [*Hint:* Use the fact that if $\dot{x} = \bar{F}(t)x$ where $\bar{F}$ has continuous entries, then a transition matrix exists.]

Problem 2.1-2. Electrical networks composed of a finite number of interconnected resistors, capacitors, inductors, and transformers can normally be described by state-space equations of the form

$$\dot{x} = Fx + Gu$$

$$y = H'x + Ju$$

The entries of the state vector will often correspond to capacitor voltages and inductor currents, the entries of the input vector to the currents at the various ports of the network, and the entries of the output vector to the voltages at the various ports of the network (sign convention is such that current × voltage = inflowing power). Assuming the initial $x(t_0)$ is nonzero, give a physical interpretation to the problem of minimizing

$$\int_{t_0}^{T} (u'Ju + x'Hu)\,dt$$

2.2 THE HAMILTON–JACOBI EQUATION

In this section, we temporarily move away from the specific regulator problem posed in the last section to consider a wider class of optimization problems requiring the minimization of a performance index. We shall, in fact, derive a partial differential equation, *the Hamilton–Jacobi equation*, satisfied by the optimal performance index under certain differentiability and continuity assumptions. Moreover, it can be shown that if a solution to the Hamilton–Jacobi equation has certain differentiability properties, then this solution is the desired performance index. But since such a solution need not exist, and not every optimal performance index satisfies the Hamilton–Jacobi equation, the equation represents only a sufficient, rather than a necessary, condition on the optimal performance index.

In this section, we shall also show how the optimal performance index, if it satisfies the Hamilton–Jacobi equation, determines an optimal control. This will allow us in Sec. 2.3 to combine the statements of the regulator problem and the Hamilton–Jacobi theory, to deduce the optimal performance index and associated optimal control for the regulator problem.

Other approaches will lead to the derivation of the optimal control and optimal performance index associated with the regulator problem, notably the use of the Minimum Principle of Pontryagin, combined with the Euler–Lagrange equations, [2], [3], and [8]. The Minimum Principle and Euler–Lagrange equations are

lengthy to derive, although their application to the regulator problem is straight-forward, as shown in Appendix C. The simplest route to take without quoting results from elsewhere appears to be the development of the Hamilton–Jacobi equation with subsequent application to the regulator problem. Actually, the Hamilton–Jacobi equation has so far rarely proved useful except for linear regulator problems, to which it seems particularly well suited.

The treatment we follow in deducing the Hamilton–Jacobi equation is a blend of treatments to be found in [3] and [7]. We start by posing the following optimal control problem. For the system

$$\dot{x} = f(x, u, t) \qquad x(t_0) \text{ given} \qquad\qquad (2.2\text{-}1)$$

find the optimal control $u^*(t)$, $t \in [t_0, T]$, which minimizes

$$V(x(t_0), u(\cdot), t_0) = \int_{t_0}^{T} l(x(\tau), u(\tau), \tau) \, d\tau + m(x(T)) \qquad (2.2\text{-}2)$$

Without explicitly defining for the moment the degree of smoothness—that is, the number of times quantities should be differentiable—we shall restrict f, l, and m to being smooth functions of their arguments. Otherwise, $f(x, u, t)$ can be essentially arbitrary, whereas $l(x(\tau), u(\tau), \tau)$, and $m(x(T))$ will often be nonnegative, to reflect some physical quantity the minimization of which is desired. As the notation implies, the performance index depends on the initial state $x(t_0)$ and time t_0, and the control $u(t)$ for all $t \in [t_0, T]$. The optimal control $u^*(\cdot)$ may be required a priori to lie in some special set, such as the set of piecewise continuous functions, square-integrable functions bounded by unity, and so forth.

Let us adopt the notation $u_{[a, b]}$ to denote a function $u(\cdot)$ restricted to the interval $[a, b]$. Let us also make the definition

$$V^*(x(t), t) = \min_{u[t, T]} V(x(t), u(\cdot), t) \qquad\qquad (2.2\text{-}3)$$

That is, if the system starts in state $x(t)$ at time t, the minimum value of the performance index (2.2-2) is $V^*(x(t), t)$. Notice that $V^*(x(t), t)$ is independent of $u(\cdot)$, precisely because knowledge of the initial state and time abstractly determines the particular control, by the requirement that the control minimize $V(x(t), u(\cdot), t)$. Rather than just searching for the control minimizing (2.2-2) and for the value of $V^*(x(t_0), t_0)$ for various $x(t_0)$, we shall study the evaluation of (2.2-3) for all t and $x(t)$, and the determination of the associated optimum control. Of course, assuming we have a functional expression for V^* in terms of $x(t)$ and t, together with the optimal control, we solve the optimization problem defined by (2.2-1) and (2.2-2) by setting $t = t_0$.

Now, for arbitrary t in the range $[t_0, T]$ and t_1 in the range $[t, T]$, recognize that $u_{[t, T]}$ is the concatenation of $u_{[t, t_1]}$ and $u_{[t_1, T]}$, so that minimizing over $u_{[t, T]}$ is equivalent to minimizing over $u_{[t, t_1]}$ and $u_{[t_1, T]}$. Then

$$V^*(x(t), t) = \min_{u[t, T]} \left[\int_t^T l(x(\tau), u(\tau), \tau) \, d\tau + m(x(T)) \right]$$

$$= \min_{u[t,\,t_1]} \left\{ \min_{u[t_1,\,T]} \left[\int_t^{t_1} l(x(\tau),\,u(\tau),\,\tau)\,d\tau \right.\right.$$

$$\left.\left. + \int_{t_1}^{T} l(x(\tau),\,u(\tau),\,\tau)\,d\tau + m(x(T)) \right] \right\}$$

where the inner minimum and the minimizing $u_{[t_1,T]}$ necessarily depend on $u_{[t,t_1]}$. Now further examine this inner minimum. The first summand is independent of $u_{[t_1,T]}$, while the second summand is itself an optimal performance index. Thus

$$V^*(x(t),\,t) = \min_{u[t,\,t_1]} \left\{ \int_t^{t_1} l(x(\tau),\,u(\tau),\,\tau)\,d\tau \right.$$

$$\left. + \min_{u[t_1,\,T]} \left[\int_{t_1}^{T} l(x(\tau),\,u(\tau),\,\tau)\,d\tau + m(x(T)) \right] \right\}$$

or

$$V^*(x(t),\,t) = \min_{u[t,\,t_1]} \left[\int_t^{t_1} l(x(\tau),\,u(\tau),\,\tau)\,d\tau + V^*(x(t_1),\,t_1) \right] \qquad (2.2\text{-}4)$$

Equation (2.2-4) is an expression of the Principle of Optimality [9, 10], which is in some ways self-evident, but nevertheless needs to be carefully considered. Consider various trajectories resulting from different controls, all commencing with state $x(t)$ at time t. Three are shown in the Fig. 2.2-1. Suppose that over $[t_1,\,T]$ the control is optimal. Thus the cost incurred in optimally traversing from $x_i(t_1)$ to $x_i(T)$ is $V^*(x_i(t_1),\,t_1)$ for each i. At this stage, the trajectory over $[t,\,t_1]$ is arbitrary. What (2.2-4) says is that the optimal cost for trajectories commencing at t and finishing at T is incurred by minimizing the sum of the cost in transiting to $x_i(t_1)$, and the optimal cost from there onwards, that is,

$$\int_t^{t_1} l(x(\tau),\,u(\tau),\,\tau)\,d\tau \quad \text{and} \quad V^*(x(t_1),\,t_1)$$
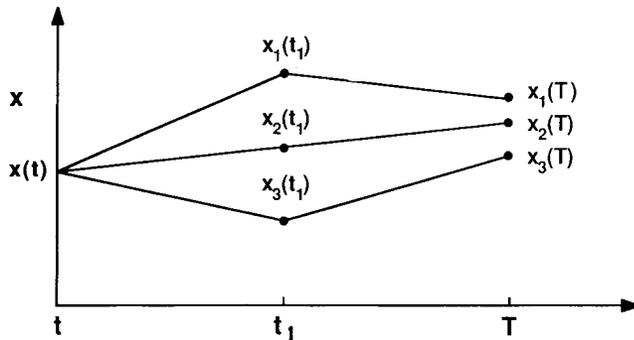
Figure 2.2-1   Illustration of Principle of Optimality; trajectories over $[t_1,\,T]$ are all optimal.

This statement of the Principle of Optimality focuses on costs. Below, we shall give a restatement focusing on controls.

In (2.2-4), let us now set $t_1 = t + \Delta t$, where Δt is small. Applying Taylor's theorem to expand the right-hand side (noting that the smoothness assumptions permit us to do this), we obtain

$$
\begin{aligned}
V^*(x(t), t) = \min_{u[t, t+\Delta t]} \Big\{ & \Delta t \, l(x(t + \alpha \, \Delta t), u(t + \alpha \, \Delta t), t + \alpha \, \Delta t) \\
& + V^*(x(t), t) + \left[\frac{\partial V^*}{\partial x}(x(t), t) \right]' \frac{dx(t)}{dt} \Delta t \\
& + \frac{\partial V^*}{\partial t}(x(t), t) \, \Delta t + 0 \, (\Delta t)^2 \Big\}
\end{aligned}
$$

where α is some constant lying between 0 and 1. Immediately,

$$
\begin{aligned}
\frac{\partial V^*}{\partial t}(x(t), t) = -\min_{u[t, t+\Delta t]} \Big\{ & l(x(t + \alpha \, \Delta t), u(t + \alpha \, \Delta t), t + \alpha \, \Delta t) \\
& + \left[\frac{\partial V^*}{\partial x}(x(t), t) \right]' f(x(t), u(t), t) + 0 \, (\Delta t) \Big\}
\end{aligned}
$$

Now, let Δt approach zero, to conclude that

$$
\frac{\partial V^*}{\partial t}(x(t), t) = -\min_{u(t)} \left\{ l(x(t), u(t), t) + \left[\frac{\partial V^*}{\partial x}(x(t), t) \right]' f(x(t), u(t), t) \right\}
$$

In this equation, f and l are known functions of their arguments, whereas V^* is unknown. In order to emphasize this point, we shall rewrite the equation as

$$
\frac{\partial V^*}{\partial t} = -\min_{u(t)} \left[l(x(t), u(t), t) + \frac{\partial V^{*\,'}}{\partial x} f(x(t), u(t), t) \right] \tag{2.2-5}
$$

This is one statement of the Hamilton–Jacobi equation. In this format, it is not precisely a partial differential equation but a mixture of a functional and a partial differential equation.

The value of $u(t)$ minimizing the right-hand side of (2.2-5) will depend on the values taken by $x(t)$, $\partial V^*/\partial x$, and t. We shall denote it by $\bar{u}(x(t), \partial V^*/\partial x, t)$.

Note also that to minimize $V(x(t), u(\cdot), t)$, the value of the minimizing control at time t is precisely $\bar{u}(x(t), \partial V^*/\partial x, t)$. To achieve our objective of expressing the optimal control as an explicit known function of $x(t)$ and t, we shall have to determine $\partial V^*/\partial x$ as an explicit known function of $x(t)$ and t.

With this definition of $\bar{u}(\cdot, \cdot, \cdot)$, (2.2-5) becomes

$$
\begin{aligned}
\frac{\partial V^*}{\partial t} = & -l[x(t), \bar{u}(x(t), \frac{\partial V^*}{\partial x}, t), t] \\
& -\frac{\partial V^{*\,'}}{\partial x} f[(x(t), \bar{u}(x(t), \frac{\partial V^*}{\partial x}, t), t]
\end{aligned} \tag{2.2-6}
$$

Despite the bewildering array of symbols, (2.2-6) is but a first-order partial differential equation with one dependent variable, V^*, and two independent variables, $x(t)$ and t, because l, f, and $\bar{u}$ are known functions of their arguments.

A boundary condition for (2.2-6) is very simply derived. Reference to the performance index (2.2-2) shows that $V(x(T), u(\cdot), T) = m(x(T))$ for all $u(\cdot)$, and, accordingly, the minimum value of this performance index with respect to $u(\cdot)$ is also $m(x(T))$. That is,

$$V^*(x(T), T) = m(x(T)) \tag{2.2-7}$$

The pair (2.2-6) and (2.2-7) may also be referred to as the Hamilton–Jacobi equation, and constitute a true partial differential equation.

If the minimization implied by (2.2-5) is impossible—that is, if $\bar{u}(\cdot, \cdot, \cdot)$ does not exist—then the whole procedure is invalidated and the Hamilton–Jacobi approach cannot be used in tackling the optimization problem.

We now consider how to determine the optimal control for the problem defined by (2.2-1) and (2.2-2). We assume that (2.2-6) and (2.2-7) have been solved so that $V^*(x(t), t)$ is a known function of $x(t)$ and t, and define

$$\hat{u}(x(t), t) = \bar{u}[x(t), \frac{\partial V^*}{\partial x}(x(t), t), t] \tag{2.2-8}$$

That is, $\hat{u}$ is the same as $\bar{u}$, except that the second variable on which $\bar{u}$ depends itself becomes a specified function of the first and third variables.

This new function $\hat{u}(\cdot, \cdot)$ has two important properties. The first and more easily seen is that $\hat{u}(x(t), t)$ is the value at time t of the optimal control minimizing

$$V(x(t), u(\cdot), t) = \int_t^T l(x(\tau), u(\tau), \tau)\, d\tau + m(x(T)) \tag{2.2-9}$$

That is, to achieve the optimal performance index $V^*(x(t), t)$, the optimal control to implement at time t is $\hat{u}(x(t), t)$.

The second property is that the optimal control $u^*(\cdot)$ for the *original* minimization problem defined by (2.2-1) and (2.2-2)—with t_0 as the initial time and t as an intermediate value of time—is related to $\hat{u}(\cdot, \cdot)$ simply by

$$u^*(t) = \hat{u}(x(t), t) \tag{2.2-10}$$

when $x(t)$ is the state at time t arising from application of $u^*(\cdot)$ over $[t_0, t)$. To some, this result will be intuitively clear, being in fact one restatement of the Principle of Optimality, to the effect that a control policy, optimal over an interval $[t_0, T]$, is Optimal over all subintervals $[t, T]$. To demonstrate it rigorously, we examine a variant of the arguments leading to (2.2-4). By definition,

$$V^*(x(t_0), t_0) = \min_{u[t_0, T]} \left[\int_{t_0}^T l(x(\tau), u(\tau), \tau)\, d\tau + m(x(T)) \right]$$

and the minimum is achieved by $u^*(\cdot)$. With $u^*(\cdot)$ regarded as the sequential use of $u^*_{[t_0, t)}$ and $u^*_{[t, T]}$, and *with the assumption that $u^*_{[t_0, t)}$ is applied until time t*, evidently

$$V^*(x(t_0), t_0) = \min_{u[t, T]} \left[\int_{t_0}^{T} l(x(\tau), u(\tau), \tau)\, d\tau + m(x(T)) \right] \tag{2.2-11}$$

$$= \int_{t_0}^{t} l(x(\tau), u(\tau), \tau)\, d\tau$$

$$+ \min_{u[t, T]} \left[\int_{t}^{T} l(x(\tau), u(\tau), \tau)\, d\tau + m(x(T)) \right] \tag{2.2-12}$$

The minimization in (2.2-11), and therefore (2.2-12), is achieved by $u^*_{[t, T]}$. In other words, $u^*_{[t, T]}$ is the optimal control for the system (2.2-1) with performance index

$$\int_{t}^{T} l(x(\tau), u(\tau), \tau)\, d\tau + m(x(T)) \tag{2.2-13}$$

with initial state $x(t)$, where $x(t)$ is derived by starting (2.2-1) at time t_0 in state $x(t_0)$, and applying $u^*_{[t_0, t)}$. But $\hat{u}(x(t), t)$ is the value of the optimal control at time t for the performance index (2.2-13), and so

$$\hat{u}(x(t), t) = u^*_{[t, T]}(t) = u^*(t) \tag{2.2-14}$$

Several points should now be noted. First, because of the way it is calculated, $\hat{u}(x(t), t)$ is independent of t_0. The implication is, then, that the optimal control at an arbitrary time σ for the minimization of

$$V(x(\sigma), u(\cdot), \sigma) = \int_{\sigma}^{T} l(x(\tau), u(\tau), \tau)\, d\tau + m(x(T)) \tag{2.2-15}$$

is also $\hat{u}(x(\sigma), \sigma)$. Put another way, the control $\hat{u}(x(\cdot), \cdot)$ is the optimal control for the whole class of problems (2.2-15), with variable $x(\sigma)$ and σ.

The second point to note is that the optimal control at time t is given in terms of the state $x(t)$ at time t, although, because its functional dependence on the state may not be constant, it is in general a time-variable function of the state. It will be theoretically implementable with a feedback law, as in Fig. 2.2-2. (Other schemes, such as the Minimum Principle and Euler–Lagrange equations, for computing the optimal control do not necessarily have this useful property; in these schemes, the optimal control may often be found merely as a certain function of time.)
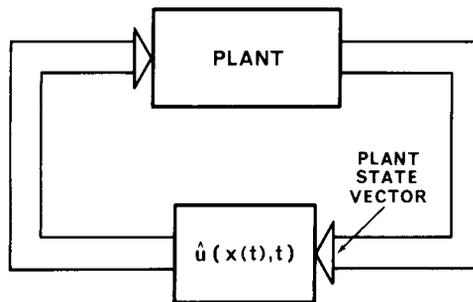
Figure 2.2-2 Feedback implementation of the optimal control.

The third point is that the remarks leading to the Hamilton–Jacobi equation are reversible, in the sense that if a suitably smooth solution of the equation is known, this solution has to be the optimal performance index $V^*(x(t), t)$.

Finally, rigorous arguments, as in, for example, [4] and [5], pin down the various smoothness assumptions precisely and lead to the following conclusion, which we shall adopt as a statement of the Hamilton–Jacobi results.

Hamilton–Jacobi equation.   Consider the system

$$\dot{x} = f(x, u, t) \tag{2.2-1}$$

and the performance index

$$V(x(t), u(\cdot), t) = \int_t^T l(x(\tau), u(\tau), \tau)\, d\tau + m(x(T)) \tag{2.2-9}$$

Suppose that f, l, and m are continuously differentiable in all arguments, that there exists a unique minimum[†] of $l(x, u, t) + \lambda' f(x, u, t)$ with respect to $u(t)$ of the form $\bar{u}(x(t), \lambda, t)$, and that $\bar{u}$ is continuously differentiable in all its arguments. Furthermore, suppose that $V^*(x(t), t)$ is a solution of the Hamilton–Jacobi equation (2.2-5) or (2.2-6) with boundary condition (2.2-7). Then $V^*(\cdot, \cdot)$ is the optimal performance index for (2.2-9), and the control given by (2.2-8) is the optimal control at time t for the class of problems with performance index (2.2-15).

Conversely, suppose that f, l, and m are continuously differentiable in all arguments, that there exists an optimal control, and that the corresponding minimum value of (2.2-9), $V^*(x(t), t)$, is twice continuously differentiable in its arguments. Suppose also that

$$l(x, u, t) + (\partial V^*/\partial x)' f(x, u, t)$$

has a unique minimum with respect to $u(t)$ at $\hat{u}(x(t), t)$, and that $\hat{u}(\cdot, \cdot)$ is differentiable in x and continuous in t. Then $V^*(\cdot, \cdot)$ satisfies the Hamilton–Jacobi equation (2.2-5) or (2.2-6) with boundary condition (2.2-7).

We conclude this section with a simple example illustrating the derivation of the Hamilton–Jacobi equation in a particular instance. Suppose we are given the system

$$\dot{x} = u$$

with performance index

$$V(x(0), u(\cdot), 0) = \int_0^T (u^2 + x^2 + \tfrac{1}{2}x^4)\, dt$$

[†] Though the point is inessential to our development, we should note that $u(t)$ may be constrained a priori to lie in some set $U(t)$ strictly contained in the Euclidean space of dimension equal to the dimension of u. All minimizations are then subject to the constraint $u(t) \in U(t)$.

Using Eq. (2.2-5), we have

$$\frac{\partial V^*}{\partial t} = -\min_{u(t)} \left\{ u^2 + x^2 + \tfrac{1}{2}x^4 + \frac{\partial V^*}{\partial x} u \right\}$$

The minimizing $u(t)$ is clearly

$$\bar{u} = -\frac{1}{2}\frac{\partial V^*}{\partial x}$$

and we have

$$\frac{\partial V^*}{\partial t} = \frac{1}{4}\left(\frac{\partial V^*}{\partial x}\right)^2 - x^2 - \frac{1}{2}x^4$$

as the Hamilton–Jacobi equation for this problem, with boundary condition $V^*(x(T),\ T) = 0$. The question of how this equation might be solved is quite unresolved by the theory presented so far. In actual fact, it is rarely possible to solve a Hamilton–Jacobi equation, although for the preceding example, a solution happens to be available [11]. It is extraordinarily complex, and its repetition here would serve no purpose.

Main points of the section.   Under smoothness and other conditions, the optimal performance index, as a function of initial time and state, satisfies a partial differential equation. The optimal control can be expressed using the optimal performance index. The Principle of Optimality is a major intuitive aid in understanding optimal control.

Problem 2.2-1.   Consider a system of the form

$$\dot{x} = f(x) + gu$$

with performance index

$$V(x(t), u(\cdot), t) = \int_t^T (u^2 + h(x))\ dt$$

Show that the Hamilton–Jacobi equation is linear in $\partial V^*/\partial t$ and quadratic in $\partial V^*/\partial x$.

Problem 2.2-2.   Let u be an r vector, and let p and x be n vectors. Let A, B, C be constant matrices of appropriate dimensions such that the following function of u, x, and p can be formed:

$$Q(u, x, p) = u'Au + 2x'Bu + 2u'Cp$$

Show that Q has a unique minimum in u for all x and p if, and only if, $\frac{1}{2}(A + A')$ is positive definite.

2.3 SOLUTION OF THE FINITE-TIME REGULATOR PROBLEM

In this section, we return to the solution of the regulator problem, which we restate for convenience.

Regulator problem. Consider the system

$$\dot{x} = F(t)x(t) + G(t)u(t) \qquad x(t_0) \text{ given} \qquad (2.3\text{-}1)$$

with the entries of $F(t)$, $G(t)$ assumed continuous. Let the matrices $Q(t)$ and $R(t)$ have continuous entries, be symmetric, and be nonnegative and positive definite, respectively. Let A be a nonnegative definite matrix. Define the *performance index*.

$$V(x(t_0), u(\cdot), t_0) = \int_{t_0}^{T} (u'Ru + x'Qx)\, dt + x'(T)Ax(T) \qquad (2.3\text{-}2)$$

and the *minimization problem* as the task of finding an *optimal control* $u^*(t)$, $t \in [t_0, T]$, minimizing V and the associated *optimum performance index* $V^*(x(t_0), t_0)$.

For the moment, assume that T is finite.

To solve the problem, we shall make use of the results on the Hamilton–Jacobi equation summarized at the end of the last section. An outline of the problem solution follows.

1. We shall show by simple arguments independent of the Hamilton–Jacobi theory that the optimal performance index $V^*(x(t), t)$, if it exists, must be of the form $x'(t)P(t)x(t)$, where $P(t)$ is a symmetric matrix.
2. With the assumption that $V^*(x(t), t)$ exists, result 1 will be used together with the Hamilton–Jacobi theory to show that $P(t)$ satisfies a nonlinear differential equation—in fact, a matrix Riccati equation.
3. We shall establish existence of $V^*(x(t), t)$.
4. We shall find the optimal control.

To carry out this program, it is necessary to make the following temporary assumption.

Temporary Assumption 2.3-1

Assume that $F(t)$, $G(t)$, $R(t)$, and $Q(t)$ have entries that are continuously differentiable.

This assumption is removed in Prob. 2.3-2.

We note that some treatments of the regulator problem assume a priori the form $x'(t)P(t)x(t)$ for the optimal performance index. It is therefore interesting to observe a simple derivation of this form, most of which appears in [12].

The quadratic form of $V^*(x(t), t)$.   The necessary and sufficient conditions for a function $V^*(x(t), t)$ to be a quadratic form are that $V^*(x(t), t)$ is continuous in $x(t)$, and

$$V^*(\lambda x, t) = \lambda^2 V^*(x, t) \qquad \text{for all real } \lambda \tag{2.3-3}$$

$$V^*(x_1, t) + V^*(x_2, t) = \tfrac{1}{2}[V^*(x_1 + x_2, t) + V^*(x_1 - x_2, t)] \tag{2.3-4}$$

(The student is asked to verify this claim in Problem 2.3-1.) To show that (2.3-3) and (2.3-4) hold, we adopt the temporary notation u_x^* to denote the optimal control over $[t, T]$ when the initial state is $x(t)$ at time t. Then the linearity of (2.3-1) and the quadratic nature of (2.3-2) imply the following equalities, whereas the inequalities follow directly from the fact that an optimal index is the minimum index. We have that

$$V^*(\lambda x, t) \leq V(\lambda x, \lambda u_x^*(\cdot), t) = \lambda^2 V^*(x, t)$$

$$\lambda^2 V^*(x, t) \leq \lambda^2 V(x, \lambda^{-1} u_{\lambda x}^*(\cdot), t) = V^*(\lambda x, t)$$

for all real constants λ. These imply (2.3-3) directly. Similar reasoning gives the inequality

$$\begin{aligned}
V^*(x_1, t) + V^*(x_2, t) &= \tfrac{1}{4}[V^*(2x_1, t) + V^*(2x_2, t)] \\
&\leq \tfrac{1}{4}[V(2x_1, u_{x_1+x_2}^* + u_{x_1-x_2}^*, t) \\
&\quad + V(2x_2, u_{x_1+x_2}^* - u_{x_1-x_2}^*, t)] \\
&= \tfrac{1}{2}[V(x_1 + x_2, u_{x_1+x_2}^*, t) \\
&\quad + V(x_1 - x_2, u_{x_1-x_2}^*, t)] \\
&= \tfrac{1}{2}[V^*(x_1 + x_2, t) + V^*(x_1 - x_2, t)]
\end{aligned} \tag{2.3-5}$$

By making use of the controls $u_{x_1}^*$ and $u_{x_2}^*$, we establish the following inequality in a like manner:

$$\tfrac{1}{2}[V^*(x_1 + x_2, t) + V^*(x_1 - x_2, t)] \leq V^*(x_1, t) + V^*(x_2, t) \tag{2.3-6}$$

Then (2.3-5) and (2.3-6) imply (2.3-4).

It is trivial to show that $V^*(x(t), t)$ is continuous in $x(t)$. We conclude that $V^*(x(t), t)$ has the form

$$V^*(x(t), t) = x'(t)P(t)x(t) \tag{2.3-7}$$

for some matrix $P(t)$, without loss of generality symmetric. {If $P(t)$ is not symmetric, it may be replaced by the symmetric matrix $\tfrac{1}{2}[P(t) + P'(t)]$ without altering (2.3-7)}.

Derivation of the matrix Riccati equation.   Now we shall show, using the Hamilton–Jacobi equation, that the symmetric matrix $P(t)$ satisfies a matrix Riccati equation.

The first form of the Hamilton–Jacobi equation is now repeated:

$$\frac{\partial V^*(x(t), t)}{\partial t} = -\min_{u(t)} \left\{ l(x(t), u(t), t) + \left[\frac{\partial V^*}{\partial x}(x(t), t) \right]' f(x(t), u(t), t) \right\} \quad (2.3\text{-}8)$$

In our case, $l(x(t), u(t), t)$ is $u'(t)R(t)u(t) + x'(t)Q(t)x(t)$; $[(\partial V^*/\partial x)(x(t), t)]'$ from (2.3-7) is $2x'(t)P(t)$, whereas $f(x(t), u(t), t)$ is $F(t)x(t) + G(t)u(t)$. The left side of (2.3-8) is simply $x'(t)\dot{P}(t)x(t)$. Hence, Eq. (2.3-8) becomes, in the special case of the regulator problem,

$$x'\dot{P}x = -\min_{u(t)} [u'Ru + x'Qx + 2x'PFx + 2x'PGu] \quad (2.3\text{-}9)$$

To find the minimum of the expression on the right-hand side of (2.3-9), we note the following identity, obtained by completing the square:

$$u'Ru + x'Qx + 2x'PFx + 2x'PGu = (u + R^{-1}G'Px)'R(u + R^{-1}G'Px)$$
$$+ x'(Q - PGR^{-1}G'P + PF + F'P)x$$

Because the matrix $R(t)$ is positive definite, it follows that (2.3-9) is minimized by setting

$$\bar{u}(t) = -R^{-1}(t)G'(t)P(t)x(t) \quad (2.3\text{-}10)$$

in which case one obtains

$$x'(t)\dot{P}(t)x(t) = -x'(t)[Q(t) - P(t)G(t)R^{-1}(t)G'(t)P(t)$$
$$+ P(t)F(t) + F'(t)P(t)]x(t)$$

Now this equation holds for all $x(t)$; therefore,

$$-\dot{P}(t) = P(t)F(t) + F'(t)P(t) - P(t)G(t)R^{-1}(t)G'(t)P(t) + Q(t) \quad (2.3\text{-}11)$$

where we use the fact that both sides are symmetric.

Equation (2.3-11) is the matrix Riccati equation we are seeking. It has a boundary condition following immediately from the Hamilton–Jacobi boundary condition. We recall that $V^*(x(T), T) = m(x(T))$, which implies in the regulator problem that $x'(T)P(T)x(T) = x'(T)Ax(T)$. Since both $P(T)$ and A are symmetric, and $x(T)$ is arbitrary,

$$P(T) = A \quad (2.3\text{-}12)$$

Before we proceed further, it is proper to examine the validity of the preceding manipulations in the light of the statement of the Hamilton–Jacobi equation at the end of the last section. Observe the following:

1. The minimization required in (2.3-9) is, in fact, possible, yielding the continuously differentiable minimum of (2.3-10). [In the notation of the last section, the expression on the right of (2.3-10) is the function $\bar{u}(\cdot, \cdot, \cdot)$ of $x(t)$, $\partial V^*/\partial x$, and t.]

2. The loss function $x'Qx + u'Ru$ and function $Fx + Gu$ appearing in the basic

system equation have the necessary differentiability properties, this being guaranteed by temporary assumption 2.3-1.

3. If $P(t)$, the solution of (2.3-11), exists, both $\dot{P}(t)$ and $\ddot{P}(t)$ exist and are continuous, the former because of the relation (2.3-11), the latter because differentiation of both sides of (2.3-11) leads to $\ddot{P}(t)$ being equal to a matrix with continuous entries (again, temporary assumption 2.3-1 is required). Consequently, $x'(t)P(t)x(t)$ is twice continuously differentiable.

Noting that Eqs. (2.3-11) and (2.3-12) imply the Hamilton–Jacobi equation (2.3-8), with appropriate initial conditions, we can then use the statement of the Hamilton–Jacobi equation of the last section to conclude the following.

1. If the optimal performance index $V^*(x(t), t)$ exists, it is of the form $x'(t)P(t)x(t)$, and $P(t)$ satisfies (2.3-11) and (2.3-12).
2. If there exists a symmetric matrix $P(t)$ satisfying (2.3-11) and (2.3-12), then the optimal performance index $V^*(x(t), t)$ exists, satisfies the Hamilton–Jacobi equation, and is given by $x'(t)P(t)x(t)$.

In theory, $P(t)$, and in particular $P(t_0)$, can be computed[†] from (2.3-11) and (2.3-12). Thus, aside from the existence question, the problem of finding the optimal performance index is solved.

Existence of the optimal performance index $V^*(x(t), t)$.

Here we shall argue that $V^*(x(t), t)$ must exist for all $t \le T$. Suppose it does not. Then, by the preceding arguments, Eqs. (2.3-11) and (2.3-12) do not have a solution $P(t)$ for all $t \le T$.

The standard theory of differential equations yields the existence of a solution of (2.3-11) and (2.3-12) in a neighborhood of T. For points sufficiently far distant from T, a solution may not exist, in which case (2.3-11) exhibits the phenomenon of a finite escape time. That is, moving back earlier in time from T, there is a first time T' such that $P(t)$ exists for all t in $(T', T]$, but as t approaches T', some entry or entries of $P(t)$ become unbounded. Then $P(t)$ fails to exist for $t \le T'$. Moreover, the only way that the solution of (2.3-11) and (2.3-12) can fail to exist away from T is if there is a finite escape time.

Since our assumption that $V^*(x(t), t)$ does not exist for all $t \le T$ implies that there is a finite escape time, we shall assume existence of a finite escape time $T' < T$ and show that this leads to a contradiction. We have $V^*(x(t), t)$ exists for all t in $(T', T]$, and, in particular, $V^*(x(T' + \epsilon), T' + \epsilon)$ exists for all positive ϵ less than $(T - T')$. Now

$$0 \le V^*(x(T' + \epsilon), T' + \epsilon) = x'(T' + \epsilon)P(T' + \epsilon)x(T' + \epsilon)$$

[†] By the word *computed*, we mean obtainable via numerical computation. There is no implication that an analytic formula yields $P(t)$.

the inequality holding because of the nonnegativity of both the integrand for all u and the final value term in (2.3-2). Hence, $P(T' + \epsilon)$ is a nonnegative definite matrix. As ϵ approaches zero, some entry becomes unbounded; without loss of generality, we can conclude that at least one diagonal entry becomes unbounded. If this were not the case, a certain 2×2 principal minor of $P(T' + \epsilon)$ must become negative as ϵ approaches zero, which contradicts the nonnegative definite property of $P(T' + \epsilon)$. Therefore, we suppose that a diagonal element—say, the ith—is unbounded as ϵ approaches zero; let e_i be a vector with zeros for all entries except the ith, where the entry is 1. Then

$$V^*(e_i, T' + \epsilon) = p_{ii}(T' + \epsilon)$$

which approaches infinity as ϵ approaches zero. (Here, p_{ij} denotes the entry in the ith row and jth column of P.)

But the optimal performance index is never greater than the index resulting from using an arbitrary control. In particular, suppose the zero control is applied to the system (2.3-1), and let $\Phi(t, \tau)$ denote the transition matrix.

Starting in state e_i at time $T' + \epsilon$, the state at time τ is $\Phi(\tau, T' + \epsilon)e_i$, and the associated performance index is

$$V(e_i, 0, T' + \epsilon) = \int_{T' + \epsilon}^{T} e_i'\Phi'(\tau, T' + \epsilon) Q(\tau)\Phi(\tau, T' + \epsilon) e_i \, d\tau$$

$$+ e_i'\Phi'(T, T' + \epsilon) A \Phi(T, T' + \epsilon) e_i$$

which must not be smaller than $p_{ii}(T' + \epsilon)$. But as ϵ approaches zero, $V(e_i, 0, T' + \epsilon)$ plainly remains bounded, whereas $p_{ii}(T' + \epsilon)$ approaches infinity. Hence, we have a contradiction that rules out the existence of a finite escape time for (2.3-11).

Thus, (2.3-11) and (2.3-12) define $P(t)$ for all $t \leq T$, and therefore the index $V^*(x(t), t) = x'(t)P(t)x(t)$ exists for all $t \leq T$.

The optimal control.   In the course of deriving the Riccati equation, we found the optimal control at time t for the regulator problem with initial time t when constructing the minimizing $u(t)$ of (2.3-9) in Eq. (2.3-10). But, as pointed out in the last section, this gives the optimal control $u^*(\cdot)$ for an arbitrary initial time [see (2.2-10)]; thus,

$$u^*(t) = -R^{-1}(t)G'(t)P(t)x(t) \tag{2.3-13}$$

Note that in (2.3-10), $P(t)$ is unknown. The product $2P(t)x(t)$ represents $\partial V^*/\partial x$, and $\bar{u}(t)$ is to be regarded as being defined by independent variables $x(t)$ (actually absent from the functional form for $\bar{u}$), $\partial V^*/\partial x$, and t. Subsequently, we are able to express $\partial V^*/\partial x$ explicitly in terms of t and x, since $P(t)$ becomes explicitly derivable from (2.3-11). This leads to the feedback law of (2.3-13). Notice, too, that Eq. (2.3-13) is a linear feedback law, as promised.

Problem 2.3-2 allows removal of temporary assumption 2.3-1, and, accordingly, we may summarize the results as follows.

Solution of the regulator problem.   The optimal performance index for the regulator problem with initial time t and initial state $x(t)$ is $x'(t)P(t)x(t)$, where $P(t)$ is given by the solution of the Riccati equation (2.3-11) with initial condition (2.3-12). The matrix $P(t)$ exists for all $t \le T$. The optimal control for the regulator problem with arbitrary initial time is given by the linear feedback law (2.3-13) for any time t in the interval over which optimization is being carried out.

To illustrate the previous concepts, we consider the following examples. The system equation we consider first is the scalar system

$$\dot{x} = fx + gu, \qquad g \ne 0$$

and the performance index is

$$\int_{t_0}^{T} (ru^2 + qx^2)\, dt, \qquad q > 0, \qquad r > 0$$

To find the control, we solve the scalar Riccati equation

$$-\dot{P} = 2fP - r^{-1}g^2P^2 + q, \qquad P(T) = 0$$

Since $r^{-1}g^2 \ne 0$, then $2f\bar{P} - r^{-1}g^2\bar{P}^2 + q = 0$ has two solutions $\bar{P}_1 \le 0$, $\bar{P}_2 \ge 0$, both real when $q \ge 0$. Since $qr^{-1}g^2 \ne 0$, then $\bar{P}_1$, $\bar{P}_2 \ne 0$, so that $\bar{P}_1 < 0$, $\bar{P}_2 > 0$ and $(\bar{P}_1 - \bar{P}_2) < 0$. The Riccati equation can now be organised as $l^2(P - \bar{P}_1)(P - \bar{P}_2) = \dot{P}$, where $l^2 = r^{-1}g^2$, or

$$\int_t^T l^2\, dt = \int_{P(t)}^0 \frac{dP}{(P - \bar{P}_1)(P - \bar{P}_2)}$$

$$= \frac{1}{\bar{P}_1 - \bar{P}_2}\left[\int_{P(t)}^0 \frac{dP}{P - \bar{P}_1} - \int_{P(t)}^0 \frac{dP}{P - \bar{P}_2}\right]$$

Carrying out the integrations gives

$$l^2(T - t) = \frac{1}{\bar{P}_1 - \bar{P}_2}\, ln\left[\frac{\bar{P}_1}{\bar{P}_2}\frac{P - \bar{P}_2}{P - \bar{P}_1}\right]$$

whence

$$P(t) = \frac{\bar{P}_1\bar{P}_2\{\exp[(\bar{P}_1 - \bar{P}_2)l^2(T - t)] - 1\}}{-\bar{P}_1 + \bar{P}_2 \exp[(\bar{P}_1 - \bar{P}_2)l^2(T - t)]}$$

Observe that $P(t)$ is well defined for all $t \le T$, and for large $(T - t)$, $P(t) \approx \bar{P}_2 > 0$, where, we recall, $\bar{P}_2 = [(f + \sqrt{f^2 + r^{-1}g^2q})/(r^{-1}g^2)]$. The optimal control is

$$u^*(t) = -r^{-1}gP(t)x(t)$$

As a second example, consider the system equation

$$\dot{x} = \tfrac{1}{2}x + u$$

and the performance index is

$$\int_{t_0}^{T} (2e^{-t}u^2 + \tfrac{1}{2}e^{-t}x^2)\, dt$$

We require, of course, the optimal control and associated optimal performance index. The Riccati equation associated with this problem is

$$-\dot{P} = P - \tfrac{1}{2}e^{t}P^2 + \tfrac{1}{2}e^{-t} \qquad P(T) = 0$$

The solution of this equation may be verified to be

$$P(t) = (1 - e^{t}e^{-T})(e^{t} + e^{2t}e^{-T})^{-1}$$

The optimal control is thus

$$u(t) = -\tfrac{1}{2}(1 - e^{t}e^{-T})(1 + e^{t}e^{-T})^{-1}x(t)$$

and the optimal performance index is

$$V^*(x(t_0), t_0) = x'(t_0)(1 - e^{t_0}e^{-T})(e^{t_0} + e^{2t_0}e^{-T})^{-1}x(t_0)$$

The above examples give little insight into how one might solve the Riccati equation. Some discussion of this point occurs in Appendix E, and we note here several key points. First, there is in general no analytic formula for solving the equation. Second, an $n \times n$ Riccati equation can be connected with a $2n$-dimensional linear differential equation (see Problem 2.3-6), so that a formula expresses the solution of the Riccati equation in terms of the transition matrix and boundary condition of the linear equation. Third, when F, G, Q, and R are all constant, the exponential formula for the transition matrix of a linear equation gives a type of analytic solution for the Riccati equation. Fourth, if F, G, Q, and R are all constant and of dimension 1, an analytic solution is available, as per the first example above.

Main points of the section.   The optimal performance index is a quadratic form in the initial state, computable by solving a matrix Riccati equation. The optimal feedback law is linear, and involves the Riccati equation solution.

Problem 2.3-1

(a) Suppose that $W(x) = x'Px$ for some symmetric P. Obviously, W is continuous in x. Show that $W(\lambda x) = \lambda^2 W(x)$ for all x and scalar λ, and $W(x_1) + W(x_2) = \tfrac{1}{2}[W(x_1 + x_2) + W(x_1 - x_2)]$ for all x_1, x_2.

(b) (This is a technically difficult problem.) Suppose that W is continuous in x, that $W(\lambda x) = \lambda^2 W(x)$ for all x and scalar λ, and $W(x_1) + W(x_2) = \tfrac{1}{2}[W(x_1 + x_2) + W(x_1 - x_2)]$. Prove that W is necessarily a quadratic in x of the form $x'Px$ for some symmetric P. Do this as follows:
   (i)   Define $p_{ii} = W(e_i)$, $p_{ji} = p_{ij} = \tfrac{1}{4}[W(e_i + e_j) - W(e_i - e_j)]$
   (ii) Show $W(e_1 \pm e_2) = p_{11} + p_{22} \pm 2p_{12}$.

(iii) By induction, show that $W(me_1 \pm e_2) = m^2 p_{11} + p_{22} \pm 2mp_{12}$ for all positive integers m.

(iv) By induction, show that $W(me_1 \pm ne_2) = m^2 p_{11} + n^2 p_{22} \pm 2mnp_{12}$ for all positive integers m, n.

(v) Using continuity, show that $W(\alpha_1 e_1 + \alpha_2 e_2) = \alpha_1^2 p_{11} + \alpha_2^2 p_{22} + 2\alpha_1 \alpha_2 p_{12}$ for all real α_1, α_2.

(vi) By induction on r, extend the calculation of $W\left(\sum_{i=1}^{r} \alpha_i e_i\right)$ from $r = 2$ to $r = $ dimension x. Use the fact that

$$\sum_{i=1}^{r+1} \alpha_i e_i = \left(\sum_{i=1}^{r-1} \alpha_i e_i + \frac{\alpha_{r+1}}{2} e_{r+1}\right) + \left(\alpha_r e_r + \frac{\alpha_{r+1}}{2} e_{r+1}\right)$$

Problem 2.3-2.   Consider the regulator problem posed at the start of the section, without Temporary Assumption 2.3-1, and let $P(t)$ be defined by (2.3-11) and (2.3-12). [Continuity of the entries of $F(t)$, and so on is sufficient to guarantee existence of a unique solution in a neighborhood of T.] Define the control law (*not* known to be optimal)

$$u^{**}(t) = -R^{-1}(t)G'(t)P(t)x(t)$$

and show that

$$V(x(t_0), u(\cdot), t_0) = x'(t_0)P(t_0)x(t_0) + \int_{t_0}^{T} (u - u^{**})'R(u - u^{**})\, d\tau$$

Conclude that if $P(t)$ exists for all $t \in [t_0, T]$, the optimal control is, in fact, u^{**}.

Problem 2.3-3.   Find the optimal control for the system (with scalar u and x)

$$\dot{x} = u \qquad x(t_0) \text{ given}$$

and with performance index

$$V(x(t_0), u(\cdot), t_0) = \int_{t_0}^{T} (u^2 + x^2)\, dt + x^2(T)$$

Problem 2.3-4.   Find the optimal control for the system (with scalar u and x)

$$\dot{x} = x + u \qquad x(t_0) \text{ given}$$

and with performance index

$$V(x(t_0), u(\cdot), t_0) = \int_{t_0}^{T} (u^2 + 2x^2)\, dt$$

Sketch (or compute) the solution for $T = 10$, $t_0 = 0$.

Problem 2.3-5.   Another version of the regulator problem imposes a further constraint—namely, that $x(T) = 0$ at the final time T. A class of performance indices implies intrinsic satisfaction of this constraint:

$$V(x(t_0), u(\cdot), t_0) = \int_{t_0}^{T} (u'Ru + x'Qx) \, dt$$

$$+ \lim_{n \to \infty} nx'(T)x(T)$$

Show that if P satisfies the Riccati equation

$$-\dot{P} = PF + F'P - PGR^{-1}G'P + Q, \qquad P(T) = nI$$

and if $W_n = P^{-1}$ exists, then W_n satisfies a Riccati equation, associated with the optimization problem

$$\dot{x} = -F'x + Du$$

$$V = \int_{t_0}^{T} (u'u + x'GR^{-1}G'x) \, dt + n^{-1}x'(T)x(T)$$

Here, D is any matrix such that $DD' = Q$. Show that $W(T) = 0$ gives a solution $W_\infty(t)$ to the Riccati equation for W, with $W_n(t) \to W_\infty(t)$ as $n \to \infty$. It turns out that $x'(t)W_\infty^{-1}(t)x(t)$ defines the optimal performance index for the constrained regulator problem [2].

Problem 2.3-6. Consider the Riccati equation $-\dot{P} = PF + F'P - PGR^{-1}G'P + Q$ with $P(T, T) = A$ and the linear differential equation

$$\begin{bmatrix} \dot{X} \\ \dot{Y} \end{bmatrix} = \begin{bmatrix} F(t) & -G(t)R^{-1}(t)G'(t) \\ -Q(t) & -F'(t) \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix} \qquad \begin{bmatrix} X(T) \\ Y(T) \end{bmatrix} = \begin{bmatrix} I \\ A \end{bmatrix}$$

Show that if $X(t)$ is nonsingular on $[t_0, T]$, then $Y(t)X^{-1}(t)$ is a solution of the Riccati equation satisfying the boundary condition. [*Hint:* Recall that $d/dt \, X^{-1}(t) = -X^{-1}(t)\dot{X}(t)X^{-1}(t)$]. Show further that if $\Phi(t, s)$ is the $2n \times 2n$ transition matrix of the linear differential equation and it is partitioned into $n \times n$ submatrices, then

$$P(t, T) = [\Phi_{21}(t, T) + \Phi_{22}(t, T)A][\Phi_{11}(t, T) + \Phi_{12}(t, T)A]^{-1}$$

2.4 DISCRETE TIME SYSTEMS

This section is devoted to a brief exposition of linear regulator design for discrete-time systems. Our starting point is the state-space equation

$$x(t + 1) = F(t)x(t) + G(t)u(t) \qquad x(t_0) \text{ given} \tag{2.4-1}$$

and the performance index

$$V(x(t_0), u(\cdot), t_0) = \sum_{t=t_0+1}^{T} [x'(t)Q(t)x(t) + u'(t-1)R(t)u(t-1)]$$

$$= \sum_{t=t_0}^{T-1} [x'(t)Q(t)x(t) + u'(t)R(t+1)u(t)] \tag{2.4-2}$$

$$+ x'(T)Q(T)x(T) - x'(t_0)Q(t_0)x(t_0)$$

In these equations, $x(t)$ is the state at time t and $u(t)$ the control at time t. Generally, but not always, t is assumed to take on integer values. The plant (2.4-1) is initially—that is, at time t_0—in state $x(t_0)$, and the aim is to return the plant state to the origin, or a state close to the origin. To do this, we set up a performance index (2.4-2), in which $Q(t)$ and $R(t)$ are nonnegative definite symmetric matrices. [*Note:* We do *not* assume $R(t)$ to be positive definite.] The performance index has the property that "large" values of the state will tend to make the performance index large. Hence, by choosing the control sequence $u(t_0)$, $u(t_0 + 1), \ldots$, which minimizes the performance index, we can expect to achieve the desired regulator effect.

It might be thought curious that $R(t)$ is not positive definite, since in the corresponding continuous time performance index, the corresponding matrix is positive definite. In the latter case, the presence of the matrix rules out the possibility of using an infinitely large control to take the state to zero in an infinitely short time. In the discrete time case, it is not possible to take the state to zero in an infinitely short time, and the possibility of an infinitely large control occurring does not even arise. Hence, there is no need to prevent such a possibility by using a positive definite $R(t)$.

We shall solve the optimization problem for the case of finite T, showing that the optimal control is a state feedback law, and that the optimal performance index is quadratic in the initial state $x(t_0)$—results that are, of course, analogous with the corresponding continuous-time results, although existence questions are much simplified.

The route to a derivation of the optimal control is via the Principle of Optimality. Thus, if until time t optimal controls $u(t_0)$, $u(t_0 + 1), \ldots, u(t - 1)$ have been applied, leading to a state $x(t)$, then the remaining terms in the optimal control sequence, $u(t)$, $u(t + 1), \ldots, u(T - 1)$ must also be optimal in the sense of minimizing $V(x(t), u(\cdot), t)$.

Now let $V^*(x(t), t)$ denote the optimal performance index associated with an initial state $x(t)$ at time t. Then, by the Principle of Optimality,

$$
\begin{aligned}
V^*(x(t), t) = \min_{u(t)} \{ & [F(t)x(t) + G(t)u(t)]'Q(t+1)[F(t)x(t) + G(t)u(t)] \\
& + u'(t)R(t+1)u(t) \\
& + V^*(F(t)x(t) + G(t)u(t), t+1) \} \\
= \min_{u(t)} \{ & u'(t)[G'(t)Q(t+1)G(t) + R(t+1)]u(t) \\
& + 2x'(t)F'(t)Q(t+1)G(t)u(t) + x'(t)F'(t)Q(t+1)F(t)x(t) \\
& + V^*(F(t)x(t) + G(t)u(t), t+1) \}
\end{aligned}
$$

(2.4-3)

Bearing in mind the corresponding continuous time results, *it would be reasonable to guess that $V^*(x(t), t)$ would be of the form $x'(t)P(t)x(t)$*. Since it proves convenient to make use of this result almost immediately, we build into the following argument an inductive proof of the result.

We require first the following assumption.

Assumption 2.4-1

For all t, $G'(t - 1)Q(t)G(t - 1) + R(t)$ is positive definite.

This assumption may, in fact, be relaxed. However, when it does not hold, the optimal control law becomes nonunique (although still linear), and we wish to avoid this complication. Notice that the assumption will very frequently hold, for example, if $R(t)$ is positive definite for all t, or if $Q(t)$ is positive definite for all t and the columns of $G(t)$ are linearly independent, and so forth.

With Assumption 2.4-1, it is easy to evaluate the "starting point" for the induction hypothesis—viz., $V^*(x(T - 1), T - 1)$. We have

$$V(x(T - 1), u(\cdot), T - 1) = x'(T)Q(T)x(T) + u'(T - 1)R(T)u(T - 1)$$

and, in view of the system equation (2.4-1), this becomes

$$V(x(T - 1), u(\cdot), T - 1) = x'(T - 1)F'(T - 1)Q(T)F(T - 1)x(T - 1)$$
$$+ 2x'(T - 1)F'(T - 1)Q(T)G(T - 1)u(T - 1)$$
$$+ u'(T - 1)[G'(T - 1)Q(T)G(T - 1) + R(T)]u(T - 1)$$

Evidently, the control $u(T - 1)$ that minimizes this performance index is a linear function of $x(T - 1)$—that is,

$$u^*(T - 1) = K'(T - 1)x(T - 1) \tag{2.4-4}$$

for a certain matrix $K(T - 1)$. Moreover, the resulting optimal index $V^*(x(T - 1), T - 1)$ becomes quadratic in $x(T - 1)$—that is,

$$V^*(x(T - 1), T - 1) = x'(T - 1)P(T - 1)x(T - 1) \tag{2.4-5}$$

a certain nonnegative definite symmetric $P(T - 1)$. The actual expressions for $K(T - 1)$ and $P(T - 1)$ are

$$K'(T - 1) = -[G'(T - 1)Q(T)G(T - 1)$$
$$+ R(T)]^{-1}G'(T - 1)Q(T)F(T - 1) \tag{2.4-6}$$

$$P(T - 1) = F'(T - 1)\{Q(T) - Q(T)G(T - 1)[G'(T - 1)Q(T)G(T - 1)$$
$$+ R(T)]^{-1}G'(T - 1)Q(T)\}F(T - 1) \tag{2.4-7}$$

Observe that $V^*(x(T - 1), T - 1)$, $P(T - 1)$, $K'(T - 1)$ exists under Assumption 2.4-1.

We now turn to the calculation of the matrices $K(t)$, determining the optimal control law, and $P(t)$, determining the optimal performance index, for arbitrary values of t. As part of the inductive hypothesis, we assume that $V^*(x(t + 1), t + 1) = x'(t + 1)P(t + 1)x(t + 1)$ for a certain matrix $P(t + 1)$. By proving that $V^*(x(t), t)$ is of the form $x'(t)P(t)x(t)$ for a certain $P(t)$, we will have established

the quadratic nature of the performance index. [Of course, the expression for $V^*(x(T-1), T-1)$ derived in Eq. (2.4-5) serves as the first step in the induction.]

Applying the inductive hypothesis to (2.4-3), we have

$$V^*(x(t), t) = \min_{u(t)} \{u'(t)[G'(t)Q(t+1)G(t) + R(t+1)]u(t)$$
$$+ 2x'(t)F'(t)Q(t+1)G(t)u(t) + x'(t)F'(t)Q(t+1)F(t)x(t)$$
$$+ x'(t)F'(t)P(t+1)F(t)x(t) + 2x'(t)F'(t)P(t+1)G(t)u(t)$$
$$+ u'(t)G'(t)P(t+1)G(t)u(t)\}$$

Again, the minimizing $u(t)$, which is the optimal control at time t, is a linear function of $x(t)$,

$$u^*(t) = K'(t)x(t) \tag{2.4-8}$$

and the optimal performance index $V^*(x(t),\ t)$, resulting from use of $u^*(t)$, is quadratic in $x(t)$—that is,

$$V^*(x(t), t) = x'(t)P(t)x(t) \tag{2.4-9}$$

The expression for $K'(t)$ is

$$K'(t) = -[G'(t)Q(t+1)G(t) + R(t+1)$$
$$+ G'(t)P(t+1)G(t)]^{-1}[G'(t)Q(t+1)F(t) + G'(t)P(t+1)F(t)]$$
$$= -[G'(t)S(t+1)G(t) + R(t+1)]^{-1}G'(t)S(t+1)F(t) \tag{2.4-10}$$

where

$$S(t+1) = Q(t+1) + P(t+1) \tag{2.4-11}$$

The expression for $P(t)$ is

$$P(t) = F'(t)\{S(t+1) - S(t+1)G(t)[G'(t)S(t+1)G(t)$$
$$+ R(t+1)]^{-1}G'(t)S(t+1)\}F(t) \tag{2.4-12}$$

Observe that $V^*(x(t), t)$, $P(t)$, $S(t)$, $K(t)$ exist for all t under Assumption 2.4-1. Of course, (2.4-11) and (2.4-12) together allow a backwards recursive determination of $P(t)$, $S(t)$. More specifically, setting $P(T) = 0$, (2.4-11) gives $S(T) = Q(T)$, (2.4-12) gives $P(T-1)$, compare with (2.4-7), then (2.4-11) gives $S(T-1)$ and so on. Equation (2.4-10) expresses the optimal feedback law in terms of known quantities and either the sequence $P(t)$ or, more conveniently perhaps, the sequence $S(t)$.

If in (2.4-12), $S(t+1)$ is replaced by $Q(t+1) + P(t+1)$, a recursion in one matrix only (which one can term a discrete-time Riccati equation) is obtained for $P(t)$. If in (2.4-11) with $t+1$ replaced by t, the expression for $P(t)$ is replaced by (2.4-12), we obtain a Riccati equation for $S(t)$:

$$S(t) = F'(t)\{S(t+1) - S(t+1)G(t)[G'(t)S(t+1)G(t)$$
$$+ R(t+1)]^{-1}G'(t)S(t+1)\}F(t) + Q(t) \qquad S(T) = Q(T) \tag{2.4-13}$$

This is perhaps the most convenient vehicle for recursion. The boundary condition can be replaced by $S(T + 1) = 0$ if desired.

Two primary references to the quadratic minimization problem for discrete-time systems are [13] and [14]. For another textbook discussion, see [2]. For convenience, we shall summarize the principal results.

Finite time regulator.   Consider the system

$$x(t + 1) = F(t)x(t) + G(t)u(t) \qquad x(t_0) \text{ given} \qquad (2.4-1)$$

Let $Q(t)$ and $R(t)$ be nonnegative definite matrices for all t, with $G'(t - 1)Q(t)G(t - 1) + R(t)$ nonsingular for all t. Define the performance index

$$V(x(t_0), u(\cdot), t_0) = \sum_{t=t_0+1}^{T} [x'(t)Q(t)x(t) + u'(t - 1)R(t)u(t - 1)] \qquad (2.4-2)$$

Then the minimum value of the performance index is

$$V^*(x(t_0), t_0) = x'(t_0)P(t_0)x(t_0)$$

where $P(t)$ is defined via

$$P(t) = S(t) - Q(t)$$

and $S(t)$ satisfies the recursion

$$S(t) = F'(t)\{S(t + 1) - S(t + 1)G(t)[G'(t)S(t + 1)G(t)$$
$$+ R(t + 1)]^{-1}G'(t)S(t + 1)\}F(t) + Q(t) \qquad S(T) = Q(T) \qquad (2.4-13)$$

The associated optimal control law is given by

$$u^*(t) = -[G'(t)S(t + 1)G(t) + R(t + 1)]^{-1}G'(t)S(t + 1)F(t)x(t) \qquad (2.4-14)$$

Main points of the section.   The optimal performance index for the linear-quadratic problem is a quadratic form in the state, with weighting matrix computable via a recursion commencing at the terminal time and evolving backwards in time. The optimal feedback law is linear in the state.

Problem 2.4-1.   Find the optimal control law and optimal performance index for the system

$$x(t + 1) = x(t) + u(t) \qquad x(0) \text{ given}$$

[where $x(\cdot)$ and $u(\cdot)$ are scalar quantities] with the performance index

$$\sum_{t-1}^{3} [2x^2(t) + u^2(t - 1)]$$

Problem 2.4-2.   When Assumption 2.4-1 fails—that is, $G'(t - 1)Q(t)$ $G(t - 1) + R(t)$ is singular for some t—in the case of single-input systems this

quantity is zero. Now the inverse of this quantity occurs in the formulas for $K'(T-1)$ and $P(T-1)$ in (2.4-6) and (2.4-7). Show, by examining the derivation of the optimal control $u^*(T-1)$, that if $G'(T-1)Q(T)G(T-1) + R(T)$ is zero, $u^*(T-1) = 0$ will be a value, but not the unique value, of the optimal control, and that the inverse may be replaced by its pseudo-inverse, namely zero. Show that for $t < T-1$, if $G'(t-1)Q(t)G(t-1) + G'(t-1)P(t)G(t-1) + R(t)$ is zero, the inverse of this quantity in the expressions for $K(t)$ and $P(t)$ may be replaced by its pseudo-inverse, namely zero.

Problem 2.4-3.   Given an n-dimensional completely controllable time-invariant system, show that the optimal control minimizing the performance index

$$V(x(t_0), u(\cdot), t_0) = x'(t_0 + n)Qx(t_0 + n)$$

where Q is positive definite, will result in a deadbeat ($x(t_0 + n) = 0$) response. [*Hint:* First prove that under controllability there exists a K such that $(F + GK)^n = 0$.]

Problem 2.4-4   (Open-ended problem requiring computer). Study the transient solution of a first-order Riccati equation and associated gains for various (say three) scalar time invariant plants and various (say two) performance indices with control Q, R in discrete (or continuous time). In each case plot transient response for the closed-loop system starting from a nonzero initial state. Also examine one case for T increasing. *Suggestion:* Perhaps choose a stable plant, an integrator and an unstable plant for this exercise. (This exercise is continued in Problems 3.3-3, 5.4-7.)

REFERENCES

[1] R. E. Kalman, Y. Ho, and K. S. Narendra, "Controllability of Linear Dynamical Systems," *Contributions to Differential Equations*, Vol. 1. New York: John Wiley & Sons, Inc., 1963.

[2] F. L. Lewis, *Optimal Control*. New York: John Wiley and Sons, 1986.

[3] A. P. Sage, *Optimal Systems Control*. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1968.

[4] R. E. Kalman, "Contributions to the Theory of Optimal Control," *Bol. Soc. Matem. Mex.* (1960), pp. 102–119.

[5] R. E. Kalman, "The Theory of Optimal Control and the Calculus of Variations," *Mathematical Optimization Techniques*, R. Bellman, ed. University of California Press, 1963, Chap. 16.

[6] R. E. Kalman, "When is a Linear Control System Optimal?" *Trans. ASME Ser. D: J. Basic Eng.*, Vol. 86 (March 1964), pp. 1–10.

[7] A. M. Letov, "Analytical Controller Design, I," *Automation and Remote Control*, Vol. 21, No. 4 (April 1960), pp. 303–306.

[8] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mischenko, *The*

Mathematical Theory of Optimal Processes, K. N. Trirogoff (transl.), L. W. Neustadt ed. New York: Interscience Publishers, 1962.

[9] S. E. Dreyfus, *Dynamic Programming and the Calculus of Variations*. New York: Academic Press Inc., 1965.

[10] R. E. Bellman, and S. E. Dreyfus, *Applied Dyanamic Programming*. Princeton, N.J.: Princeton University Press, 1962.

[11] R. Bellman, and R. Bucy, "Asymptotic Control Theory," *SIAM J. Control*, Vol. 2, No. 1 (1964), pp. 11–18.

[12] P. A. Faurre, "Sur les points conjugués en commande optimale," *Compt. Rend. Acad. Sci. Paris, Ser. A*, Vol. 266 (June 24, 1968), pp. 1294–1296.

[13] R. E. Kalman, and T. S. Englar, "A User's Manual for ASP C," *RIAS Rep. NASA*, Baltimore, Md., 1965.

[14] R. E. Kalman, and R. W. Koepcke "Optimal Synthesis of Linear Sampling Control Systems Using Generalized Performance Indices," *Trans. ASME*, Vol. 80 (November 1958), pp. 1820–1826.

3

The Standard Regulator Problem—II

3.1 THE INFINITE-TIME REGULATOR PROBLEM

In this section, we lift the restriction imposed in the last chapter that the final time (the right-hand endpoint of the optimization interval), T, be finite. We thus have the following problem.

Infinite-time regulator problem. Consider the system

$$\dot{x} = F(t)x(t) + G(t)u(t) \qquad x(t_0) \text{ given} \qquad (3.1\text{-}1)$$

with the entries of $F(t)$, $G(t)$ assumed continuous. Let the matrices $Q(t)$ and $R(t)$ have continuous entries, be symmetric, and be nonnegative and positive definite, respectively. Define the performance index

$$V(x(t_0), u(\cdot), t_0) = \int_{t_0}^{\infty} (u'(t)R(t)u(t) + x'(t)Q(t)x(t))\, dt \qquad (3.1\text{-}2)$$

and the minimization problem as the task of finding an optimal control $u^*(t)$, $t \geq t_0$, minimizing V and the associated optimum performance index $V^*(x(t_0), t_0)$.

It is not always possible to solve this problem as it is stated. To give some idea of the difficulty, consider the system

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u \qquad \begin{bmatrix} x_1(t_0) \\ x_2(t_0) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

with performance index defined by $R = [1]$ and

$$Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

It is readily established that

$$V(x(t_0), u(\cdot), t_0) = \int_{t_0}^{\infty} (u^2 + e^{2t - 2t_0}) \, dt$$

In a sense, V is minimized by taking $u = 0$; but the resulting value of V is still infinite.

   With the finite-time problem, the optimal V is always finite; this may not be so in the infinite-time case. For the example given, it is clear that V becomes infinite for the following three reasons:

1. The state $x_1(t_0)$ is uncontrollable.
2. The uncontrollable part of the system trajectory is unstable $(x_1(t) = e^t)$.
3. The unstable part of the system trajectory is reflected in the system performance index (e^{2t} is integrated).

The difficulty would not have arisen if one of these reasons did not apply.

   It is intuitively clear that any regulator problem where one had a situation corresponding to that described by 1, 2, and 3 cannot have a finite optimal performance index. Therefore, to ensure that our problems are solvable, we shall make the following assumption.

Assumption 3.1-1

   System (3.1-1) is completely controllable for every time t. (See Appendix B for definition.)

   Actually, this condition can be relaxed to requiring complete stabilizability for all t (and even this condition can sometimes be further relaxed). However, for the sake of simplicity, relaxation to complete stabilizability is considered in this text only for the time-invariant case—see next section.

   We now state the solution to the infinite-time regulator problem, under Assumption 3.1-1.

   Solution to the infinite-time regulator problem.   Let $P(t, T)$ be the solution of the equation

$$-\dot{P} = PF + F'P - PGR^{-1}G'P + Q \qquad (3.1\text{-}3)$$

with boundary condition $P(T, T) = 0$. Then $\lim_{T \to \infty} P(t, T) = \bar{P}(t)$ exists for all t and is a solution of (3.1-3). Moreover, $x'(t)\bar{P}(t)x(t)$ is the optimal perform-

ance index $V^*(x(t), t)$, when the initial time is t and the initial state is $x(t)$. The optimal control at time t when the initial time is arbitrary is uniquely† given by

$$u^*(t) = -R^{-1}(t)G'(t)\bar{P}(t)x(t) \tag{3.1-4}$$

(assuming t lies in the optimization interval).

Evidently, we need to establish four separate results: (1) the existence of $\bar{P}(t)$; (2) the fact that it is a solution of (3.1-3); (3) the formula for the optimal performance index; and (4) the formula for the optimal control. The proof of (3) and (4) will be combined.

Existence of $\bar{P}(t)$.   Since (3.1-1) is completely controllable at time t, there exists for every $x(t)$ a control $\bar{u}(\cdot)$ and a time t_2 such that $\bar{u}(\cdot)$ transfers $x(t)$ to the zero state at time t_2. Although $\bar{u}(\cdot)$ is initially defined only on $[t, t_2]$, we extend the definition to $[t, \infty)$ by taking $\bar{u}(\cdot)$ to be zero after t_2. This ensures that the system will remain in the zero state after time t_2. The notation $V(x(t), u(\cdot), t, T)$ will be used to denote the performance index resulting from initial state $x(t)$ at time t, a control $u(\cdot)$, and a final time T, which is finite rather than infinite as in (3.1-2). Then $P(t, T)$ exists for all T and $t \le T$. Since the performance index is the integral of a nonnegative quantity, $P(t, T)$ is nonnegative symmetric. Moreover

$$
\begin{aligned}
x'(t)P(t, T)x(t) &= V^*(x(t), t, T) \\
&\le V(x(t), \bar{u}_{[t, T]}, t, T) \\
&\le V(x(t), \bar{u}_{(t, \infty)}, t, \infty) \\
&= V(x(t), \bar{u}_{[t, t_2]}, t, t_2) \\
&< \infty
\end{aligned}
$$

Since $x(t)$ is arbitrary, it follows that the entries of $P(t, T)$ are bounded independently of T. [Note that as T approaches infinity, if any entry of $P(t, T)$ became unbounded, there would have to be a diagonal entry of $P(t, T)$ which became unbounded, tending to infinity, or else $P(t, T)$ could not be nonnegative symmetric; consequently, for a suitable $x(t)$, $x'(t)P(t, T)x(t)$ would be unbounded.]

Reference to (3.1-2), with T replacing $+\infty$ on the integral, and use of the nonnegative definite and positive definite character of Q and R, respectively, shows that

$$x'(t)P(t, T_0)x(t) \le x'(t)P(t, T_1)x(t)$$

for any $T_1 \ge T_0$. The bound on $P(t, T)$ together with this monotonicity relation guarantees existence of the limit $\bar{P}(t)$. In more detail, the existence of the limit $\bar{p}_{ii}(t)$

---

† At least, $u^*(t)$ is uniquely defined up to a set of measure zero, unless we insist on some property such as continuity. Henceforth, we shall omit reference to this qualification. (Those unfamiliar with measure theory may neglect this point.)

will follow by taking $x(t) = e_i$, for each i, and applying the well-known result of analysis that a bounded monotonically increasing function possesses a limit. (Recall that e_i is defined as a vector with zeros for all entries except the ith, where the entry is 1.) The existence of $\bar{p}_{ij}(t)$ will follow by observing that

$$2p_{ij}(t, T) = (e_i + e_j)'P(t, T)(e_i + e_j) - p_{ii}(t, T) - p_{jj}(t, T)$$

and that each term on the right possesses a limit as T approaches infinity.

$\bar{P}(t)$ satisfies the Riccati equation (3.1-3). Denote the solution of (3.1-3) satisfying $P(T) = A$ by $P(t, T; A)$. [Then $P(t, T)$, defined earlier, is $P(t, T; 0)$.] Now a moment's reflection shows that

$$P(t, T; 0) = P(t, T_1; P(T_1, T; 0))$$

for $t \leq T_1 \leq T$, and thus

$$\bar{P}(t) = \lim_{T \to \infty} P(t, T; 0) = \lim_{T \to \infty} P(t, T_1; P(T_1, T; 0))$$

For fixed time T_1, the solution $P(t, T_1; A)$ of (3.1-3) depends continuously on A; therefore,

$$\bar{P}(t) = P(t, T_1; \lim_{T \to \infty} P(T_1, T; 0))$$

$$= P(t, T_1; \bar{P}(T_1)),$$

which proves that $\bar{P}(t)$ is a solution of (3.1-3) defined for all t.

Optimal performance index and control formulas. We show first that if the control defined by (3.1-4) is applied (where there is no assumption that this control is optimal), then

$$V(x(t), u^*(\cdot), t, \infty) = \lim_{T \to \infty} V(x(t), u^*(\cdot), t, T) = x'(t)\bar{P}(t)x(t) \qquad (3.1\text{-}5)$$

Direct substitution of (3.1-4) into the performance index (3.1-2), with the initial time replaced by t and the final time by T, leads to

$$V(x(t), u^*(\cdot), t, T) = x'(t)\bar{P}(t)x(t) - x'(T)\bar{P}(T)x(T)$$

$$\leq x'(t)\bar{P}(t)x(t)$$

and, therefore,

$$\lim_{T \to \infty} V(x(t), u^*(\cdot), t, T) \leq x'(t)\bar{P}(t)x(t)$$

Also,

$$V(x(t), u^*(\cdot), t, T) \geq V^*(x(t), t, T) = x'(t)P(t, T)x(t)$$

and, therefore,

$$\lim_{T \to \infty} V(x(t), u^*(\cdot), t, T) \geq x'(t)\bar{P}(t)x(t)$$

The two inequalities for $\lim_{T\to\infty} V(x(t), u^*(\cdot), t, T)$ then imply (3.1-5). Since $u^*(\cdot)$ has not yet been shown to be optimal, it follows that

$$V^*(x(t), t, \infty) \le V(x(t), u^*(\cdot), t, \infty) \tag{3.1-6}$$

We now show that the inequality sign is impossible. Suppose that strict inequality holds. Then there is a control u_1, different from u^*, such that

$$\lim_{T\to\infty} V(x(t), u_1, t, T) = V^*(x(t), t, \infty)$$

Since, also, from the first and third members of (3.1-5),

$$\lim_{T\to\infty} V^*(x(t), t, T) = V(x(t), u^*(\cdot), t, \infty)$$

it follows that strict inequality in (3.1-6) implies

$$\lim_{T\to\infty} V(x(t), u_1, t, T) < \lim_{T\to\infty} V^*(x(t), t, T)$$

This, in turn, requires for suitably large T

$$V(x(t), u_1, t, T) < V^*(x(t), t, T)$$

This is plainly impossible by the definition of the optimal performance index as the minimum over all possible indices. Consequently, we have established that $x'(t)\bar{P}(t)x(t)$ is the optimal performance index for the infinite-time problem, and that $-R^{-1}(t)G'(t)\bar{P}(t)x(t)$ is the unique optimal control because it achieves this performance index, thereby completing the formal solution to the infinite-time regulator problem.

Time-invariant case.   It is of interest, for practical reasons, to determine whether time-invariant plants (3.1-1) will give rise to time-invariant linear control laws of the form

$$u(t) = K'x(t) \tag{3.1-7}$$

For finite-time optimization problems of the type considered in the last chapter, no choice of T, $R(\cdot)$, and $Q(\cdot)$ will yield a time-invariant control law when F and G are constant, unless the matrix A takes on certain special values. (Problem 3.1-2 asks for this fact to be established.) For the infinite-time problem, the case is a little different. Let us state an infinite-time problem that will be shown to yield a constant control law. Later in the chapter, we shall consider a variation on this problem statement which also yields a constant (and linear) control law, even though the weighting matrices become time varying.

Time-invariant regulator problem.   Consider the system

$$\dot{x} = Fx + Gu \qquad x(t_0) \text{ given} \tag{3.1-8}$$

where F and G are constant. Let the constant matrices Q and R be non-negative and positive definite, respectively. Define the performance index

$$V(x(t_0), u(\cdot), t_0) = \int_{t_0}^{\infty} (u'Ru + x'Qx)\, dt \qquad (3.1\text{-}9)$$

and the minimization problem as the task of finding an optimal control $u^*(\cdot)$ minimizing V, together with the associated optimum performance index.

To ensure solvability of the problem, it is necessary, as before, to make some additional restrictions. We shall require the following relaxation of Assumption 3.1-1 specialized to the time-invariant case.

Assumption 3.1-2

   The system (3.1-8) is completely stabilizable.

Solution to the time-invariant regulator problem.   Let $P(t, T)$ be the solution of the equation

$$-\dot{P} = PF + F'P - PGR^{-1}G'P + Q \qquad (3.1\text{-}3)$$

with initial condition $P(T, T) = 0$. Then $\lim_{T \to \infty} P(t, T) = \bar{P}$ exists and is constant; also, $\bar{P} = \lim_{t \to -\infty} P(t, T)$. Furthermore, $\bar{P}$ satisfies (3.1-3); that is,

$$\bar{P}F + F'\bar{P} - \bar{P}GR^{-1}G'\bar{P} + Q = 0 \qquad (3.1\text{-}10)$$

and $x'(t)\bar{P}x(t)$ is the optimal performance index when the initial time is t and the initial state is $x(t)$. The optimal control at time t when the initial time is arbitrary is uniquely given by the constant control law

$$u^*(t) = -R^{-1}G'\bar{P}x(t) \qquad (3.1\text{-}11)$$

First, we show that when F, G are constant, the solution to the infinite-time regulator problems (3.1-3) and (3.1-4) is still obtainable on relaxing Assumption 3.1-1 to Assumption 3.1-2.

The proof of the first part of the claim rests on the existence, under Assumption 3.1-2, of a feedback control law $\bar{u}(t) = \bar{K}'x(t)$ which achieves exponential stability of the system $\dot{x}(t) = (F + G\bar{K}')x(t)$. It is clear then that, following the earlier proof,

$$x'(t)P(t, T)x(t) \leq V(x(t), \bar{u}_{[t,\infty)}, t, \infty) < \infty$$

which leads to existence of $\bar{P}(t)$ as in the earlier proof. (Similar arguments can be applied for the time-varying case in terms of $\bar{K}(t)$ under a stabilizability assumption, but we do not formally state results for this case.)

Now using the time-varying results, we can straightforwardly establish the various other claims just made.

First, $\lim_{T \to \infty} P(t, T)$ certainly exists. Now the plant is time invariant, and the function under the integral sign of the performance index is not specifically time

dependent. This means that the choice of the initial time is arbitrary—that is, all initial times must give the same performance index, which is to say that

$$\bar{P}(t) = \lim_{T \to \infty} P(t, T)$$

is independent of t. We denote it by $\bar{P}$. Likewise, because the initial time is arbitrary,

$$\bar{P} = \lim_{T \to \infty} P(t, T) = \lim_{T \to \infty} P(0, T - t) = \lim_{t \to -\infty} P(0, T - t) = \lim_{t \to -\infty} P(t, T)$$

To illustrate the preceding concepts, we can consider the simple scalar example discussed in the previous chapter. The prescribed system is

$$\dot{x} = fx + gu \qquad g \neq 0$$

and the performance index is

$$V(x(t_0), u(\cdot), t_0) = \int_{t_0}^{\infty} (ru^2 + qx^2) \, dt \qquad q > 0 \qquad r > 0$$

To find the optimal control, we solve the scalar equations

$$-\dot{P} = 2fP - r^{-1}g^2P^2 + q \qquad P(T, T) = 0$$

The steady state solution is the positive solution of $2f\bar{P} - r^{-1}g^2\bar{P}^2 + q = 0$, which is identical to the limiting solution of $P(t, T)$ as $T \to \infty$ studied in the previous chapter, namely

$$\bar{P}_2 = \frac{f + (f^2 + g^2q/r)^{1/2}}{r^{-1}g^2} > 0$$

The associated control law is

$$u^* = K'x \qquad K' = -r^{-1}g\bar{P}_2 = -\frac{f + (f^2 + g^2q/r)^{1/2}}{g}$$

and closed-loop system is

$$\dot{x} = -(f^2 + g^2q/r)^{1/2}x$$

Notice that it is only the ratio of q to r that influences the feedback gain K. Also, as q/r increases, the gain increases in magnitude and approximates $-(\text{sgn } g)(q/r)^{1/2}$. Then the closed-loop equations approximate $\dot{x} = -|g|(q/r)^{1/2}x$ with a high magnitude negative real closed-loop pole. That is, as (q/r) increases, the system bandwidth increases.

As (q/r) decreases, the feedback gain approximates $K = -[f + |f|]/g$ when the open loop is stable ($f < 0$), so that $K \approx 0$ and the closed loop approximates the open loop. When the open loop is unstable ($f > 0$), then $K \approx -2f/g$ and the closed-loop system approximates $\dot{x} = -fx$. That is, the closed-loop pole reflects the open-loop pole into the left half-plane, and the system bandwidth remains the same.

This example foreshadows more general results studied in later chapters, namely that for multivariable situations, as $Q \to \infty$ (or $R \to 0$) then the feedback gain

K becomes large and the bandwidth of the resulting closed-loop system become high. In practical designs, the required bandwidth is often known a priori, so that the designer searches for appropriate Q, R to achieve these. A second more general result is that for a stable open-loop plant, as $Q \to 0$ (or $R \to \infty$), then $K \to 0$ and the closed-loop poles move to the open-loop poles, whereas for an unstable plant, $K \nrightarrow 0$, but the closed-loop poles move to the stable reflections through the imaginary axis of the unstable open-loop poles.

In the last chapter, we found an expression for the solution of the Riccati differential equation $P(t, T)$. We see that one can write

$$P(t, T) = [a + b \exp(-2\bar{f}(t - T))] [c + d \exp(-2\bar{f}(t - T))]^{-1}$$

where a, b, c, d are certain constants and $\bar{f}$ is a certain quantity which turns out to be identical with the closed-loop mode $f + gk$ (see Problem 3.1-4). It follows that as $t \to -\infty$, $P(t, T) \to \bar{P}_2 = ac^{-1}$ with a time constant given by half the time constant of the closed-loop system. This is in fact a general phenomenon, that the Riccati differential equation converges to its steady state solution with time constants one-half those associated with the closed-loop system. Proving this result would take us too far afield at this point.

Let us consider now the very simple plant (open-loop integrator)

$$\dot{x} = u$$

so that $f = 0$, $g = 1$ above. Then $\bar{P}_2 = (rq)^{1/2}$, $K = -(q/r)^{1/2}$ so that the closed-loop system is

$$\dot{x} = -\tau^{-1} x, \qquad \tau = (r/q)^{1/2}$$

Then

$$\int_0^\infty ru^2 \, dt = \int_0^\infty qx^2 \, dt = \tfrac{1}{2}V^* = \tfrac{1}{2}(rq)^{1/2} x_0^2$$

That is, the control cost is identical to the state cost, each being half the total optimal cost $V^* = (rq)^{1/2} x_0^2$.

A further example is provided by a voltage regulator problem, discussed in [1]. The open-loop plant consists of a cascade of single-order blocks, of transfer functions

$$\frac{3}{0.1s + 1}, \frac{3}{0.04s + 1}, \frac{6}{0.07s + 1}, \frac{3.2}{2s + 1}, \frac{2.5}{5s + 1}$$

In other words, the Laplace transform $Y(s)$ of the deviation from the correct output is related to the Laplace transform $U(s)$ of the deviation from the reference input by

$$Y(s) = \frac{3}{0.1s + 1} \cdot \frac{3}{0.04s + 1} \cdot \frac{6}{0.07s + 1} \cdot \frac{3.2}{2s + 1} \cdot \frac{2.5}{5s + 1} U(s)$$

State-space equations relating $u(\cdot)$ to $y(\cdot)$ are readily found to be

$$\dot{x} = \begin{bmatrix} -0.2 & 0.5 & 0 & 0 & 0 \\ 0 & -0.5 & 1.6 & 0 & 0 \\ 0 & 0 & -\frac{100}{7} & \frac{600}{7} & 0 \\ 0 & 0 & 0 & -25 & 75 \\ 0 & 0 & 0 & 0 & -10 \end{bmatrix} x + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 30 \end{bmatrix} u$$

[Here, $x_1, x_2, \ldots, x_5$ have physical significance; x_5 is the output of a block of transfer function $3(0.1s + 1)^{-1}$ with input u, x_4 the output of a block of transfer function $3(0.4s + 1)^{-1}$ with input x_5, and so on. Also, $x_1 = y$ is the tracking error voltage to be regulated to zero.]

As a first attempt at optimization, the performance index

$$\int_{t_0}^{\infty} (x_1^2 + u^2)\, dt$$

is considered, that is, it penalizes directly the tracking error cost. (Of course, $x_1^2 = x'Qx$, where Q is a matrix the only nonzero entry of which is unity in the $1-1$ position.)

Appendix E describes methods for solving the Riccati equation, and there is also some discussion in the next section. In this example, the important quantity calculated is not so much the matrix $\bar{P}$ but the optimal gain vector. The optimal control law is, in fact, found to be

$$u = -[0.9243 \quad 0.1711 \quad 0.0161 \quad 0.0492 \quad 0.2644]x$$

Stabilizability has been crucial in our arguments which show that $\bar{P}$ exists, and it is reasonable to ask whether stabilizability is necessary for existence. Indeed, it is not, at least when one considers contrived examples. For example, suppose $\dot{x} = x + u$, $R = 1$, $Q = 0$. Then the optimal control is $u^* \equiv 0$. Once, however, the performance index "observes" all states, stabilizability is essential. Problem 3.1-3 asks for this to be established. The point is reexamined at the end of the next section.

Main points of the section.   Complete controllability or stabilizability of the state space equations ensure existence of optimal linear quadratic indices and associated Riccati equation solutions in the infinite-time case.

Problem 3.1-1.   Consider the time-invariant problem, and suppose F and G have the form

$$F = \begin{bmatrix} F_{11} & F_{12} \\ 0 & F_{22} \end{bmatrix} \qquad G = \begin{bmatrix} G_1 \\ 0 \end{bmatrix}$$

so that the pair $[F, G]$ is not completely controllable and $[F_{11}, G_1]$ is completely controllable; see Appendix B. Suppose that the solution $P(t)$ of the Riccati equation associated with the finite-time regulator problem is partitioned as

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P'_{12} & P_{22} \end{bmatrix}$$

Show that P_{11} satisfies a Riccati equation, that P_{12} satisfies a linear differential equation—assuming P_{11} is known—and that P_{22} satisfies a linear differential equation—assuming P_{12} is known.

Problem 3.1-2. Consider the finite-time regulator problem posed in the previous chapter, where the linear system is time-invariant. Show that if $A = 0$, there is no way, even permitting time-varying $R(t)$ and $Q(t)$, of achieving a constant nonzero control law. Show that if R and Q are constant and if $[F, G]$ is stabilizable, then there is at least one particular A that will achieve a constant control law, and that this A is independent of the final time T. [*Hint:* For the second part of the problem, make use of Eq. (3.1-10).]

Problem 3.1-3. Consider $\dot{x} = Fx + Gu$, assumed to have an uncontrollable unstable mode. Suppose that $R > 0$ and $Q > 0$. Show that $P(0, T)$ is unbounded with T and conclude that the infinite time problem has no finite solution. [*Hint:* Change the coordinate basis to that used in Problem 3.1-1, choose an initial state $[x'_1(0) \quad x'_2(0)]'$ such that $\exp(F_{22}t)x_2(0)$ diverges, and argue that $x'(0)P(0, T)x(0)$ diverges as $T \to \infty$.]

Problem 3.1-4. For the system $\dot{x} = fx + gu$ with performance index $V(x(t_0), t_0, u(\cdot)) = \int_{t_0}^{T} (ru^2 + qx^2)\, dt$, we found in the last section that $P(t, T) = [a + b \exp(-2\bar{f}(t - T))][c + d \exp(-2\bar{f}(t - T))]^{-1}$ for certain a, b, c, d and $\bar{f}$. Show that $\bar{f} = f + gk$ with k the gain for the steady state problem.

Problem 3.1-5. Consider the system $\dot{x} = Gu$ where G is $n \times n$ and non-singular. Show that the infinite-time problem is solvable, and the optimal control causes

$$\int_{t_0}^{\infty} x'Qx\, dt = \int_{t_0}^{\infty} u'Ru\, dt$$

Problem 3.1-6. An angular position control system is described by the equation

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & -10 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t)$$

(Here, x_1 denotes angular position, x_2 angular velocity, and the control is a torque). Choose $R = 1$, and $Q = \text{diag}[q, 0]$, penalizing angular position but not velocity. Show that the optimal control is

$$u^* = -[\sqrt{q} \quad \sqrt{100 + 2\sqrt{q}} - 10]x$$

and the closed-loop poles are the zeros of $s^2 + \sqrt{100 + 2\sqrt{q}}\, s + \sqrt{q}$. [*Hint:* Write down the steady state Riccati equation and by examining each term, solve for the

entries of $\bar{P}$, applying the nonnegative definiteness of $\bar{P}$ to eliminate certain solutions].

3.2 STABILITY OF THE TIME-INVARIANT REGULATOR

In this section, we shall be concerned with the stability of the closed-loop system formed when the control law resulting from an infinite-time performance index is implemented. Throughout, we shall assume a constant system that is stabilizable:

$$\dot{x} = Fx + Gu \tag{3.2-1}$$

We shall also assume a performance index with constant nonnegative definite Q and constant positive definite R:

$$V(x(t_0), u(\cdot), t_0) = \int_{t_0}^{\infty} (u'Ru + x'Qx)\, dt \tag{3.2-2}$$

We recall that the optimal performance index is $x'(t_0)\bar{P}x(t_0)$, where $\bar{P}$ is the limiting solution of a Riccati equation; also, $\bar{P}$ satisfies the algebraic equation

$$\bar{P}F + F'\bar{P} - \bar{P}GR^{-1}G'\bar{P} + Q = 0 \tag{3.2-3}$$

The optimal control is given by

$$u^* = -R^{-1}G'\bar{P}x \tag{3.2-4}$$

and, accordingly, the closed-loop system becomes

$$\dot{x} = (F - GR^{-1}G'\bar{P})x \tag{3.2-5}$$

We ask the question: When is (3.2-5) an asymptotically stable system? Certainly, (3.2-5) is not always stable. Consider the example $\dot{x} = x + u$, with $V = \int_{t_0}^{\infty} u^2\, dt$. Immediately, the optimal control is $u \equiv 0$, and the closed-loop system is $\dot{x} = x$, which is plainly unstable. In this instance, there are two factors contributing to the difficulty.

1. The original open-loop system is unstable.
2. The unstable trajectories do not contribute in any way to the performance index—in a sense, the unstable states are not *observed* by the performance index.

Intuitively, one can see that if 1 and 2 were true in an arbitrary optimization problem, there would be grounds for supposing that the closed-loop system would be unstable.

Accordingly, to ensure asymptotic stability of the closed-loop system, it is necessary to prevent the occurrence of 1 and 2. This motivates the introduction of the following assumption, which will be shown to guarantee stability of the closed-loop system.

Assumption 3.2-1

The pair $[F, D]$ is *detectable*, where D is any matrix such that $DD' = Q$ (see Appendix B).

We can note at once that the question of whether Assumption 3.2-1 holds is determined by Q alone, and not by the particular factorization DD'. To see this, suppose D_1 and D_2 are such that $D_1D_1' = D_2D_2' = Q$. Recall that $[F, D_1]$ detectable tells us $Fw = \lambda_iw$, $D_1'w = 0$, $w \neq 0$, only if $\text{Re}(\lambda_i[F]) < 0$. Since $D_1'w = 0$ implies $D_1D_1'w = 0$, $Qw = 0$, $D_2D_2'w = 0$, which in turn implies $D_2'w = 0$ then $[F, D_1]$ detectable implies $[F, D_2]$ detectable, and the converse. Hence, either $[F, D_1]$ and $[F, D_2]$ are detectable simultaneously, or they are not detectable simultaneously. (Notice incidentally that $[F, D]$ is detectable if and only if $[F, Q]$ is detectable.)

Assumption 3.2-1 essentially ensures that all potentially unstable trajectories will show up in the $x'Qx$ part of the integrand of the performance index. Since the performance index is known a priori to have a finite value, it is plausible that any potentially unstable trajectories will be stabilized by the application of the feedback control.

The actual proof of asymptotic stability can be achieved via the Lemma of Lyapunov (see Appendix A), and properties of detectability. A direct proof is now given.

Denoting $(F - GR^{-1}G'\bar{P})$ as $\bar{F}$, then $\bar{P}$ satisfies

$$\bar{P}\bar{F} + \bar{F}'\bar{P} = -DD' - \bar{P}GR^{-1}G'\bar{P} \tag{3.2-6}$$

Let $\bar{F}w = \lambda w$, $w \neq 0$; then we need to show $\text{Re}[\lambda] < 0$. Now pre- and postmultiplication of the above equation for $\bar{P}$ by w^*, w where here superscript $*$ denotes complex conjugate transpose, gives

$$(\lambda + \lambda^*)w^*\bar{P}w = -w^*DD'w - w^*\bar{P}GR^{-1}G'\bar{P}w$$

If $\lambda + \lambda^* = 0$, then $D'w = 0$, $R^{-1}G'\bar{P}w = 0$, and in turn $Fw = \bar{F}w = \lambda w$, which leads to a contradiction, since $\text{Re}[\lambda] < 0$ under detectability—see Appendix B definitions. Again, if $\lambda + \lambda^* > 0$, then the nonpositivity of the right side implies $w^*\bar{P}w \leq 0$, while the nonnegativity of $\bar{P}$ forces $w^*\bar{P}w \geq 0$, so that in fact $\bar{P}w = 0$. Then, as before, $D'w = 0$, $R^{-1}G'\bar{P}w = 0$, $Fw = \lambda w$ for $\text{Re}[\lambda] > 0$, contradicting detectability.

The practical implications of the stability result should be clear. Normally no one wants an unstable system; here, we have a procedure for guaranteeing that an optimal system is bound to be stable, irrespective of the stability of the open-loop plant. The result is in pleasant contrast to some of the results and techniques of classical control, where frequently the main aim is to achieve stability and questions of optimality occupy an essentially secondary role in the design procedure. The contrast will actually be heightened when we exhibit some of the additional virtues of the optimal regulator solution in later chapters.

The interested reader may wonder to what extent the stability result applies in time-varying, infinite-time problems. There are actually two key issues in the infinite-time, time-varying case. The first is to secure boundedness of the matrices

$\overline{P}(t)$ defining the optimal index, and $\overline{K}(t)$, the control gain. The second is to secure exponential stability of the regulator, since mere asymptotic stability inevitably leads to robustness problems.

It clearly makes sense to restrict attention to $F(t)$, $G(t)$, $Q(t)$, $R(t)$, $R^{-1}(t)$ bounded. Then with $[F(t), G(t)]$ uniformly controllable, $[F(t), D(t)]$ uniformly observable (see Appendix B), it may be shown, see [2], that the desired bounded-ness and exponential stability are achieved. This stability result is not altogether easy to achieve. Actually, it is possible to define concepts of uniform stabilizability and detectability (which are much more subtle concepts in the time-varying case than the time-invariant case), and to establish the boundedness and exponential stability when these weaker concepts replace uniform controllability and observability. Such treatments are more immediate in the discrete-time case [3].

There is one further property of interest. In the finite time problem, $P(t, T)$ is nonnegative for all t, T. It follows that $\overline{P}(t)$ in the infinite time problem is non-negative, as is $\overline{P}$ in the time-invariant problem. When is $\overline{P}$ actually positive definite? This question is answered as follows.

Lemma

Consider the time-invariant regulator problem as defined by Eqs. (3.2-1) through (3.2-5) and the associated remarks. Adopt Assumption 3.2-1. Then $\overline{P}$ is positive definite if and only if $[F, D]$ is completely observable.

Proof.   From the form of the performance index (3.2-2), it is clear that $x'(t_0)\overline{P}x(t_0)$ must be nonnegative for all $x(t_0)$. Suppose for some nonzero x_0 we have $x_0'\overline{P}x_0 = 0$, with observability holding. Now the only way the integral in (3.2-2) will turn out to be zero is if the nonnegative integrand is always zero. This requires the optimal control to be zero for all t; consequently, the system state at time t when the optimal control is applied is simply $\exp[F(t - t_0)]x_0$. Furthermore,

$$0 = x_0'\overline{P}x_0 = \int_{t_0}^{\infty} x'Qx\,dt = \int_{t_0}^{\infty} x_0'\exp[F'(t - t_0)]DD'\exp[F(t - t_0)]x_0\,dt$$

Now, we have a contradiction, because the preceding equation implies $D'\exp[F(t - t_0)]x_0 = 0$ for all t, but we have assumed x_0 to be nonzero and we have assumed observability. So observability implies nonsingular $\overline{P}$.

To establish the converse, suppose there exists a nonzero x_0 such that $D'\exp[F(t - t_0)]x_0 = 0$ for all t. Apply the control $u(t) = 0$ for all t to the open-loop plant. It then follows that the associated performance index is zero. Since the optimal performance index is bounded below by zero, it follows that it, too, is zero—that is, $x_0'\overline{P}x_0 = 0$, which implies finally that $\overline{P}$ is singular.

To illustrate these ideas, we recall briefly the two examples of the previous section. In the first case, we had

$$\dot{x} = fx + gu \qquad V(x(t_0), u(\cdot), t_0) = \int_{t_0}^{\infty} (ru^2 + qx^2)\,dt$$

Obviously, with $g \neq 0$, $q \neq 0$ the stabilizability and detectability conditions are satisfied. Therefore, it is no surprise that the closed-loop system, computed to be $\dot{x} = -(f^2 + g^2q/r)^{1/2}x$, is asymptotically stable.

For the voltage regulator example, we recall that the F matrix is

$$\begin{bmatrix} -0.2 & 0.5 & 0 & 0 & 0 \\ 0 & -0.5 & 1.6 & 0 & 0 \\ 0 & 0 & -\frac{100}{7} & \frac{600}{7} & 0 \\ 0 & 0 & 0 & -25 & 75 \\ 0 & 0 & 0 & 0 & -10 \end{bmatrix}$$

and the weighting matrix Q is

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

A matrix D such that $DD' = Q$ is given by $D' = [1 \quad 0 \quad 0 \quad 0 \quad 0]$. Let us examine the rank of $[D \quad F'D \quad \ldots \quad (F')^4 D]$. This matrix is readily checked to be triangular with nonzero elements all along the diagonal. Hence it has rank 5. Consequently $[F, D]$ is observable, and thus detectable. Hence the closed-loop system is known to be asymptotically stable.

It is natural to ask whether detectability is necessary for stability of the closed-loop system. The answer is yes. Roughly speaking, if there exists an unstable mode that is not observed, then the cheapest control for it is $u \equiv 0$, and obviously that will not stabilize it. Details of the arguments are requested in Problem 3.2-1.

A related question is whether for existence and stability of the closed-loop system (and finite performance index), detectability and stabilizability are both required. (Recall from the last section that if there is no closed-loop stability requirement, stabilizability is sufficient but not necessary for closed-loop system existence.)

The answer is easy: stabilizability is required, since stabilizability is equivalent to the existence of a stabilizing feedback gain K, that is, the opportunity to obtain a stable closed-loop system. In summary

> ***Role of detectability and stabilizability.*** For the infinite-time, time-invariant problem parametrized by $F, G, Q = DD'$ and $R = R' > 0$, there exists a stable closed-loop optimal system with finite performance index if and only if $[F, G]$ is stabilizable and $[F, D]$ is detectable.

This result has its origins in the work of [2] and [4–6].

It is appropriate to comment further on the task of solving the Riccati equation. We noted in the last section that constant F, G, Q, R allow the linear differential equation associated with the Riccati equation to be solved (in terms of a matrix exponential); a solution to the Riccati differential equation can then be obtained and from it, a solution to the steady state equation. Appendix E sets out a number

of other approaches applicable in the time-invariant case (where the detectability and stabilizability conditions are fulfilled). A derivation of some of the background theory is sought in the problems. In particular, there is only one nonnegative definite solution to the steady state equation (see Problem 3.2-2), and there is only one solution to the steady state equation which defines a stabilizing control law $K = -\bar{P}GR^{-1}$ (see Problem 3.2-7).

The Hamiltonian matrix

$$M = \begin{bmatrix} F & -GR^{-1}G' \\ -Q & -F' \end{bmatrix}$$

appearing in the associated linear differential equation has the property that if λ is an eigenvalue, so is $-\lambda$, and no eigenvalue is pure imaginary (see Problem 3.2-8). The steady state solution of the Riccati equation can be described in terms of the eigenvectors of M associated with negative real part eigenvalues (see Problem 3.2-9). The negative real part eigenvalues of M are also the eigenvalues of the closed-loop matrix $F - GR^{-1}G'\bar{P}$ (see Problem 3.2-10). Finally, $\bar{P}$ can be expressed by using a Schur form for M (see Problem 3.2-11).

The results of Problem 3.2-8 and 3.2-10 provide a convenient way to show that $P(t, T)$ approaches $\bar{P}$ with time constants determined by half those of the closed-loop system (see Problem 3.2-12).

Main points of the section.   For the time-invariant, infinite-time linear quadratic problem, stabilizability and detectability ensure existence and asymptotic stability of the closed-loop system. The optimal controller is time-invariant, and the optimal closed-loop system is time-invariant.

Problem 3.2-1.   Suppose for the conventional time-invariant problem defined by matrices F, G, $Q = DD'$ and $R > 0$, the pair $[F, D]$ is not detectable; conclude that the closed-loop system will not be stable. [*Hint:* Choose a nonzero possibly complex w for which $Fw = \lambda w$, Re $\lambda > 0$ and $D'w = 0$. Choose an initial state $x(0) = $ Re w. Show that $x'(0)$ $P(0, T)x(0) = 0$ for all T by observing that $u(t) \equiv 0$ achieves zero performance index. Infer the instability result.]

Problem 3.2-2.   The scalar version of (3.2-3) is a simple quadratic equation, which, in general, has more than one solution. Likewise, when (3.2-3) is a true matrix equation, there is generally more than one solution. Show that there is only one nonnegative definite solution under Assumption 3.2-1 and a stabilizability assumption. [This identifies the limiting solution of the Riccati equation uniquely as *the* nonnegative definite solution of (3.2-3).] [*Hint:* Suppose there are two solutions, $\bar{P}_1$ and $\bar{P}_2$, both nonnegative definite. Define $F_i = F - GR^{-1}G'\bar{P}_i$ and show that both F_i have eigenvalues with negative real parts. Prove that $(\bar{P}_1 - \bar{P}_2)F_1 + F_2'(\bar{P}_1 - \bar{P}_2) = 0$ and use the result that the matrix equation $AX + XB = 0$ has the unique solution $X = 0$, provided $\lambda_i[A] + \lambda_j[B] \neq 0$ for any i and j; see also Appendix A for this result.]

Problem 3.2-3. Consider the two infinite-time performance indices $V_1^\infty = \int_{t_0}^\infty (u'Ru + x'Qx)\, dt$, $V_2^\infty = V_1 + \lim_{T \to \infty} x'(T)Ax(T)$ with $F, G, Q = DD'$, $R = R' > 0$, $A = A' \geq 0$, stabilizability of $[F, G]$ and detectability of $[F, D]$. Show that the optimum values of V_1 and V_2 are the same; conclude that the Riccati equation $-\dot{P} = PF + F'P - PGR^{-1}G'P + Q$ with $P(T, T) = A$ and solution $P(t, T)$ yields $\lim_{T \to \infty} P(t, T) = \bar{P}$ for *any* $A = A' \geq 0$.

Problem 3.2-4. In the standard problem of this section, with the detectability condition strengthened to observability, now $\bar{P}$ is positive definite. Show that $x'\bar{P}x$ is a Lyapunov function associated with the closed-loop system $\dot{x} = (F - GR^{-1}G'\bar{P})x$, under $Q > 0$. Relax the condition on Q to $[F, D]$ completely observable for $DD' = Q$. Use the Lyapunov theorems in Appendix D.

Problem 3.2-5. For the standard problem of this section, show that the closed-loop system is more stable than the open-loop system in that the center of gravity of the closed-loop eigenvalues lies to the left of the center of gravity of open-loop eigenvalues. [*Hint:* For a square matrix A, $\Sigma \lambda_i(A) = \text{trace } A$].

Problem 3.2-6. Consider the driven harmonic oscillator with damping ζ (positive or negative, but $|\zeta| < 1$).

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\omega_0^2 & -2\zeta\omega_0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

Consider also the performance index in which $R = 1$ and $Q = \text{diag}\,[q_1, q_2]$. Solve the steady state Riccati equation and verify that the closed loop is stable. Consider the effect of q_1, $q_2 \to 0$ when $\zeta > 0$ and $\zeta < 0$ on the closed-loop characteristic polynomial. [*Hint:* Write out the steady state Riccati equation term by term, and solve the resulting equations, using the nonegativity property to define which solution out of several possibilities should apply.]

Problem 3.2-7. Problem 3.2-2 sought demonstration that there is only one nonnegative definite solution to the Riccati equation. Use the ideas of Problem 3.2-2 to show there is only one solution $\bar{P}$ such that $K = -\bar{P}GR^{-1}$ is stabilizing.

Problem 3.2-8. Let J, M be the matrices

$$J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \qquad M = \begin{bmatrix} F & -GR^{-1}G' \\ -Q & -F' \end{bmatrix}$$

Show that $J^2 = -I$, $JMJ = M'$ and conclude that if λ_0 is an eigenvalue of M, so is $-\lambda_0$. Show that M has no pure imaginary eigenvalue under the detectability/ stabilizability assumptions. Do this by postulating that $w = [w_1'\ \ w_2']'$ is an eigenvector corresponding to eigenvalue $j\mu$ for real μ, and examine the quantity $[w_2^*\ \ w_1^*]Mw = j\mu(w_2^*w_1 + w_1^*w_2)$.

Problem 3.2-9.   By virtue of Problem 3.2-8, there exists a nonsingular W such that

$$W^{-1}MW = \begin{bmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{bmatrix}$$

where Λ_1, Λ_2 are real Jordan matrices with all eigenvalues possessing negative real and positive real part eigenvalues respectively. With W partitioned into four $n \times n$ blocks, show that $\bar{P} = W_{21}W_{11}^{-1}$. [*Hint:* Express $P(t, T)$ in terms of W_{ij}, exp $\Lambda_1 t$ and exp $\Lambda_2 t$ and evaluate the limit as $T - t \to \infty$.]

Problem 3.2-10.   Using the result of Problem 3.2-9, show that the eigenvalues of Λ_1 are the same as those of $F - GR^{-1}G'\bar{P}$.

Problem 3.2-11.   Suppose that

$$M = U'LU = \begin{bmatrix} U'_{11} & U'_{21} \\ U'_{12} & U'_{22} \end{bmatrix}\begin{bmatrix} L_{11} & L_{12} \\ 0 & L_{22} \end{bmatrix}\begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix}$$

where U is orthogonal, and L is a real Schur form with L_{11}, L_{22} possessing negative real part and positive real part eigenvalues respectively. Verify that $\bar{P} = U_{11}^{-1}U_{12}$ solves the steady state Riccati equation and is stabilizing. [*Hint:* Equate the 2-1 and 2-2 terms in the identity $UM = LU$, and eliminate L_{22} to show that a solution of the steady state equation is $-U_{22}^{-1}U_{21}$, assuming U_{22} is nonsingular. Show that the transpose also satisfies the equation. By considering the 2-2 term, show that $F + GR^{-1}G'(U_{22}^{-1}U_{21})'$ has eigenvalues with negative real parts, so that $-(U_{22}^{-1}U_{21})'$ is the desired solution. Using orthogonality of U, verify $\bar{P} = U_{11}^{-1}U_{12}$.]

Problem 3.2-12.   Show that, under the usual assumptions, $P(t, T) = [A + B(t)][C + D(t)]^{-1}$ where $B(t)$ and $D(t)$ consist of linear combinations of terms exp $[-(\lambda_i + \lambda_j)(t - T)]$ where λ_i, λ_j are eigenvalues of $F - GR^{-1}G'\bar{P}$. (Assume that M has no repeated eigenvalues, for convenience, and use the idea of Problem 3.2-9),

3.3 SUMMARY AND DISCUSSION OF THE REGULATOR PROBLEM RESULTS

In this section, we first summarize all the important regulator problem results hitherto established.

Regulator problem and solution.   Consider the system

$$\dot{x} = F(t)x + G(t)u \qquad x(t_0) \text{ given} \tag{3.3-1}$$

with the entries of F and G assumed continuous. Let the matrices $Q(t)$ and $R(t)$ have continuous entries, be symmetric, and be nonnegative and positive definite, respectively. Let A be a nonnegative definite matrix. Define the performance index

$$V(x(t_0), u(\cdot), t_0) = \int_{t_0}^{T} (u'(t)R(t)u(t) + x'(t)Q(t)x(t)) \, dt$$

$$+ x'(T)Ax(T) \tag{3.3-2}$$

where T is finite. Then the minimum value of the performance index is

$$V^*(x(t_0), t_0) = x'(t_0)P(t_0, T)x(t_0) \tag{3.3-3}$$

where $P(t_0, T)$ is the solution of the Riccati equation

$$-\dot{P} = PF + F'P - PGR^{-1}G'P + Q \tag{3.3-4}$$

with boundary condition $P(T, T) = A$. The matrix $P(t, T)$ exists for all $t \leq T$. The associated optimal control is given by the linear feedback law

$$u^*(t) = -R^{-1}(t)G'(t)P(t, T)x(t) \tag{3.3-5}$$

   Infinite-time regulator problem and solution.   Suppose the preceding hypotheses all hold, save that $A = 0$. Suppose also that the system (3.3-1) is completely controllable for all time. Then

$$\bar{P}(t) = \lim_{T \to \infty} P(t, T) \tag{3.3-6}$$

exists, and the minimum value of the performance index (3.3-2) with T replaced by infinity is $x'(t_0)\bar{P}(t_0)x(t_0)$. The matrix $\bar{P}$ satisfies the Riccati equation (3.3-4), and the optimal control law is

$$u^*(t) = -R^{-1}(t)G'(t)\bar{P}(t)x(t) \tag{3.3-7}$$

   Time-invariant regulator problem and solution.   Suppose that the infinite-time regulator problem is specialized to the time-invariant case with matrices F, G, Q, and R constant. Suppose also that $[F, G]$ is stabilizable. Then

$$\bar{P} = \lim_{T \to \infty} P(t, T) = \lim_{t \to -\infty} P(t, T) \tag{3.3-8}$$

is constant, and satisfies the algebraic equation

$$\bar{P}F + F'\bar{P} - \bar{P}GR^{-1}G'\bar{P} + Q = 0 \tag{3.3-9}$$

The optimal control law is a time-invariant law

$$u^*(t) = -R^{-1}G'\bar{P}x(t) \tag{3.3-10}$$

   Asymptotically stable time-invariant regulator problem and solution. Suppose in addition that $[F, D]$ is detectable, for any $DD' = Q$. Then $\bar{P}$ is nonnegative definite and, see Problem 3.2-2, is the only solution of (3.3-9) with this property. Moreover, the optimal closed-loop system

$$\dot{x} = [F - GR^{-1}G'\bar{P}]x \tag{3.3-11}$$

is asymptotically stable. If the pair $[F, D]$ is completely observable where D is any matrix such that $DD' = Q$, then $\bar{P}$ is positive definite. Furthermore, $x'\bar{P}x$ is a Lyapunov function.

The corresponding discrete-time results are now summarized—with proof of the infinite-time results (not given in the text) relegated to the Problem 3.3-1. See also [7].

Finite time regulator.   Consider the system

$$x(t+1) = F(t)x(t) + G(t)u(t) \qquad x(t_0) \text{ given} \qquad (3.3\text{-}12)$$

Let $Q(t)$ and $R(t)$ be nonnegative definite matrices for all t, with $G'(t-1)Q(t)G(t-1) + R(t)$ nonsingular for all t. Define the performance index

$$V(x(t_0), u(\cdot), t_0) = \sum_{t=t_0+1}^{T} [x'(t)Q(t)x(t) + u'(t-1)R(t)u(t-1)] \qquad (3.3\text{-}13)$$

Then the minimum value of the performance index is

$$V^*(x(t_0), t_0) = x'(t_0)P(t_0, T)x(t_0)$$

where

$$P(t, T) = S(t) - Q(t) \qquad (3.3\text{-}14)$$

with $S(t)$ defined recursively via

$$S(t) = F'(t)\{S(t+1) - S(t+1)G(t)[G'(t)S(t+1)G(t) \\ + R(t+1)]^{-1}G'(t)S(t+1)\}F(t) + Q(t) \qquad S(T) = Q(T) \qquad (3.3\text{-}15)$$

The associated optimal control law is given by

$$u^*(t) = -[G'(t)S(t+1)G(t) + R(t+1)]^{-1}G'(t)S(t+1)F(t)x(t) \qquad (3.3\text{-}16)$$

Infinite-time regulator.   Assuming now that $T \to \infty$, and that the pair $[F(t), G(t)]$ is completely controllable for all t, then

$$\bar{P}(t) = \lim_{T \to \infty} P(t, T) \qquad (3.3\text{-}17)$$

exists, and the optimum value of the performance index (3.3-13), with T replaced by infinity, is $x'(t_0)\bar{P}(t_0)x(t_0)$. The matrix $\bar{P}$ satisfies the recursion relations (3.3-14), (3.3-15) (apart from the boundary condition) and the optimal control law is given by (3.3-16).

Time-invariant regulator.   When F, G, Q, and R are constant, with $[F, G]$ stabilizable, $\bar{P}$ is constant and may be obtained via

$$\bar{P} = \lim_{t \to -\infty} P(t, T) \qquad (3.3\text{-}18)$$

The matrix

$$\bar{S} = \bar{P} + Q \qquad (3.3\text{-}19)$$

satisfies the steady state version of (3.3-15), viz.

$$\bar{S} = F'\{\bar{S} - \bar{S}G[G'\bar{S}G + R]^{-1}G'\bar{S}\}F + Q \qquad (3.3\text{-}20)$$

The optimal control law is also constant, being given by

$$u^* = -[G'\bar{S}G + R]^{-1}G'\bar{S}Fx \qquad (3.3\text{-}21)$$

With D any matrix such that $DD' = Q$, complete detectability of the pair $[F, D]$ guarantees asymptotic stability of the closed-loop system. Under observability of $[F, D]$, then $\bar{P} > 0$ and $x'(k)\bar{P}x(k)$ is a Lyapunov function.

Now that the solution to the regulator problem has been determined, we might ask if this solution is of interest for application to situations other than those in which everything is set up as in the formal statements just summarized.

Frequently, the system states are not available for use in a controller input. In Chapters 7 and 8 we show that if the system is specified in the form

$$\dot{x} = Fx + Gu \qquad y = H'x$$

where only the input u and output y are available for controller inputs, a state estimator can be constructed with inputs u and y and output $\hat{x}$, an estimate of the state x. Implementing the control law $u = -R^{-1}G'P\hat{x}$ is a satisfactory alternative to implementing the control law $u = -R^{-1}G'Px$.

So far in our discussions, we have assumed that the performance index matrices R and Q are specified. Since specification is usually up to the designer, a range of values for these matrices can be considered and by trial and error the most suitable values can be selected. Chapter 6 considers methods to guide this trial and error approach. Of course, if we are interested only in stability, then the closed-loop system $\dot{x} = (F - GR^{-1}G'\bar{P})x$ is always stable, irrespective of the choice of Q and R within their prescribed limits. In other words, we have a general method for stabilizing multiple input linear systems (assuming that state estimators may be constructed).

We are now led to ask whether linear regulators, optimal in the sense previously discussed, have desirable properties for the engineer other than simply "stability." Or we might ask whether there are methods for selecting the index parameters Q and R so that desirable properties, such as good transient response and good sensitivity characteristics, are achieved.

With these questions in mind, we now move on to the next sections and chapter, which discuss various extensions of the regulator problem, and then to the following few chapters, which deal with further properties of regulator systems.

Main points of the section.   For continuous and discrete time, the linear-quadratic problem in finite time yields a linear control law and optimal performance index quadratic in the state. With detectability and stabilizability assumptions, the infinite-time time-invariant problem yields a stabilizing feedback law.

Problem 3.3-1.   For the discrete-time problem with constant F, G, Q, and R, derive the results claimed above under stabilizability and detectability assumptions (Follow the continuous-time theory. Also see [7]).

Problem 3.3-2.   (Compare with Problem 2.4-1). Find the optimal control and performance index for

$$x(t + 1) = x(t) + u(t)$$

[where $x(\cdot)$ and $u(\cdot)$ are scalar quantities] with performance index

$$\sum_{t=1}^{\infty} [2x^2(t) + u^2(t - 1)]$$

Check that the optimal control is stabilizing.

Problem 3.3-3.   (Open problem requiring computer) For the cases studied in Problem 2.4-4, examine the corresponding infinite time Riccati equation solutions and associated feedback gains. For each of the designs examine the closed-loop eigenvalues. This problem is continued as Problem 5.4-7.

Problem 3.3-4

(a) Consider the index

$$\int_{t_0}^{T} [u'(t)R(t)u(t) + x'(t)Q(t)x(t)] \, dt + x'(T)Ax(T)$$

for the system $\dot{x} = F(t)x + G(t)u$ with the usual assumptions. Suppose that a control $u(t) = K'(t)x(t)$ is implemented. Show that the value assumed by the performance index is given by $x'(t_0)P_K(t_0)x(t_0)$ where $P_K(\cdot)$ is the solution of the linear differential equation

$$-\dot{P}_K = P_K(F + GK') + (F + GK')P_K + KRK' + Q$$

$$P_K(T) = A$$

Verify that if $K = -P_K GR^{-1}$, then P_K satisfies the usual Riccati equation.

(b) For the discrete time system $x(t + 1) = F(t)x(t) + G(t)u(t)$ and performance index

$$\sum_{t_0+1}^{T} [x'(t)Q(t)x(t) + u'(t - 1)R(t)u(t - 1)]$$

with the usual assumptions, suppose that the control law $u(t) = K'(t)x(t)$ is implemented. Show that the value of the performance index is given by $x'(t_0)[S_K(t_0) - Q(t_0)]x(t_0)$ where

$$S_K(t) = [F(t) + G(t)K'(t)]'S_K(t + 1)[F(t) + G(t)K'(t)] + K(t)R(t + 1)K'(t)$$

$$+ Q(t) \qquad S_K(T) = Q(T)$$

Verify that if $K(t) = -[G'(t)S_K(t + 1)G(t) + R(t + 1)]^{-1}G'(t)S_K(t + 1)F(t)$, then $S_K(t)$ agrees with the optimal $S(t)$. This fact offers a numerically more attractive procedure for iteratively determining $S(t)$, in that nonnegative matrices are added rather than differenced as part of the solution process.

Problem 3.3-5.   Let P_K be as defined in the previous problem and rewrite the Riccati equation for P as

$$-\dot{P} = P(F + GK') + (F' + KG')P - PGK' - KG'P - PGR^{-1}G'P + Q$$

Show that $Z = P_K - P$ satisfies

$$-\dot{Z} = Z(F + GK') + (F' + KG')Z + (K + PGR^{-1})R(K' + R^{-1}G'P)$$

$$Z(T) = 0$$

Express the solutions of this equation, regarding everything as known except Z as an integral. Conclude that $Z(t_0) \geq 0$ for all t_0, so that $P_K \geq P$. This shows the optimality of P.

3.4 CROSS-PRODUCT TERMS AND SECOND VARIATION THEORY

There are many avenues along which the regulator problem can be extended; a number of them will be explored throughout the text. The approach used in developing the extensions is to reduce by a suitable transformation the given problem to the standard regulator problem.

In this section, we first consider extension of our theory for the case when there are cross-product terms in the quadratic performance index. Such terms can well arise when power into a system is penalized—see Problem 2.1-2, for example. Another very important application of such results is to the case when an open-loop optimal control is in place for a nonlinear plant and/or nonquadratic index, but additional closed-loop regulation is required to maintain as closely as possible the optimal trajectory in the presence of disturbances that cause small perturbations from the trajectory. In this case, under reasonable smoothness assumptions, optimal linearization of the plant and "quadraticization" of the index about the optimal trajectory yield a linear time-varying plant with quadratic index containing cross-product terms. This will be demonstrated later in the section. Meanwhile the case of cross-product terms in the index will be studied.

We consider the determination of an optimal control and associated optimal performance index for the system

$$\dot{x} = F(t)x + G(t)u \qquad x(t_0) \text{ given} \tag{3.4-1}$$

when the performance index contains cross-product terms as

$$V(x(t_0), u(\cdot), t_0) = \int_{t_0}^{T} [u'(t)R(t)u(t) + 2x'(t)S(t)u(t)$$
$$+ x'(t)Q(t)x(t)] \, dt \tag{3.4-2}$$

with R positive definite and the following constraint holding:

$$Q - SR^{-1}S' \geq 0 \tag{3.4-3}$$

(shorthand for $Q - SR^{-1}S'$ is nonnegative definite). If desired, T can be infinite, and F, G, Q, R, and S constant.

To reduce this problem to one covered by the previous theory, we note the following identity, obtained by completing the square:

$$u'Ru + 2x'Su + x'Qx = (u + R^{-1}S'x)'R(u + R^{-1}S'x) + x'(Q - SR^{-1}S')x.$$

Making the definition

$$u_1 = u + R^{-1}S'x \qquad (3.4\text{-}4)$$

the original system (3.4-1) becomes equivalent to

$$\dot{x} = (F - GR^{-1}S')x + Gu_1 \qquad (3.4\text{-}5)$$

and the original performance index is equivalent to

$$V(x(t_0), u_1(\cdot), t_0) = \int_{t_0}^{T} [u_1' Ru_1 + x'(Q - SR^{-1}S')x] \, dt \qquad (3.4\text{-}6)$$

If u and u_1 are related according to (3.4-4), the trajectories of the two systems (3.4-1) and (3.4-5) are the same [provided they start from the same initial state $x(t_0)$]. Furthermore, the values taken by the two performance indices—viz., (3.4-2), which is associated with the system (3.4-1), and (3.4-6), which is associated with the system (3.4-5)—are also the same. Consequently, the following statements hold.

1. The optimal controls u^* and u_1^* for the two optimization problems are related by $u_1^* = u^* + R^{-1}S'x$.
2. The optimal performance indices for the two problems are the same.
3. The closed-loop trajectories (when the optimal controls are implemented) are the same.

Now the optimization problem associated with (3.4-5) and (3.4-6) is certainly solvable [and here we are using the nonnegativity constraint (3.4-3)]. The optimal u_1^* is given by

$$u_1^*(t) = -R^{-1}(t)G'(t)P(t, T)x(t) \qquad (3.4\text{-}7)$$

where

$$-\dot{P} = P(F - GR^{-1}S') + (F' - SR^{-1}G')P \\ - PGR^{-1}G'P + Q - SR^{-1}S' \qquad (3.4\text{-}8)$$

with $P(T, T) = 0$. The optimal index is $x'(t_0)P(t_0, T)x(t_0)$. The optimal control for the optimization problem associated with (3.4-1) and (3.4-2) is thus

$$u^*(t) = -R^{-1}(t)[G'(t)P(t, T) + S'(t)]x(t) \qquad (3.4\text{-}9)$$

and the optimal performance index is again $x'(t_0)P(t_0, T)x(t_0)$.

The assumption (3.4-3) allows ready application of all the preceding theory.

Actually, it may not be fulfilled in certain applications, especially those involving second variation theory. What happens if it is absent?

With reference to (3.4-6), recall that (3.4-3) serves to guarantee an under-bound on $V^*(x(t_0), t_0)$, of zero. Given an overbound too (choose say $u \equiv 0$), the Riccati equation (3.4-8) could have no escape time. Without an underbound on $V^*(x(t_0), t_0)$ there does arise the possibility that for some t_1, $\lim_{t \downarrow t_1} P(t, T)$ is not finite, so that, for some w, $w'P(t, T)w \rightarrow -\infty$ as $t \downarrow t_1$. When t_1 is the closest such point to the left of T, it is called a conjugate point; $V^*(x(t_0), t_0)$ is well defined for all $x(t_0)$ only when $t_0 > t_1$, but may be negative.

Let us now discuss the infinite-time problem, assuming that (3.4-3) holds. To consider the infinite-time problem [i.e., T in (3.4-2) is replaced by infinity], we make the following assumption.

Assumption 3.4-1

The system (3.4-1) is completely controllable at every time t.

To ensure existence of the limit as T approaches infinity of $P(t, T)$, one evidently requires that the system (3.4-5) be completely controllable at every time t. [Then the optimization problem associated with (3.4-5) and (3.4-6) is solvable and yields a solution of the optimization problem associated with (3.4-1) and (3.4-2).] The complete controllability of (3.4-5) is an immediate consequence of Assumption 3.4-1, since controllability is invariant under state variable feedback—see Appendix B.

The control law for the infinite-time problem will be constant if F, G, Q, R, and S are all constant, as may easily been seen. The closed-loop system will be asymptotically stable if the pair $[F - GR^{-1}S', D]$ is completely detectable where D is any matrix such that $DD' = Q - SR^{-1}S'$. (If D were such that $DD' = Q$, note that complete detectability of $[F, D]$ or $[F - GR^{-1}S', D]$ would not necessarily imply asymptotic stability.)

Let us now show how linear quadratic problems with cross-product terms arise when dealing with linearized plants and "quadraticized" indices. We return to the nonlinear optimal control task of Chapter 2, Section 2, that is, the minimization of

$$\bar{V}(\xi(t_0), v(\cdot), t_0) = \int_{t_0}^{T} l(\xi(\tau), v(\tau), \tau) \, d\tau + m(\xi(\tau)) \qquad (3.4\text{-}10)$$

subject to the constraint

$$\dot{\xi} = f(\xi, v, t), \qquad \xi(t_0) \quad \text{given} \qquad (3.4\text{-}11)$$

under suitable smoothness assumptions on $l(\cdot), m(\cdot), f(\cdot)$.

A linearization about an optimal state [control] trajectory $\xi^*(\cdot)[v^*(\cdot)]$, both assumed available as time functions, is readily achieved using Taylor series expansions, and neglecting higher order terms. Thus for $x \triangleq \xi - \xi^*$, $u \triangleq v - v^*$, with both small,

$$f(\xi, v, t) = f(\xi^*, v^*, t) + F(t)x(t) + G(t)u(t)$$

$$F(t) = \frac{\partial f}{\partial \xi}\bigg|_{\substack{\xi(t) = \xi^*(t) \\ v(t) = v^*(t)}}, \qquad G(t) = \frac{\partial f}{\partial v}\bigg|_{\substack{\xi(t) = \xi^*(t) \\ v(t) = v^*(t)}} \qquad (3.4\text{-}12)$$

[Here the ijth component of $F(t)$ is the partial derivative with respect to ξj of the ith component of $f(\cdot)$ evaluated at $\xi^*(t), v^*(t)$. Likewise for $G(t)$.] The linearized "plant" is now (an approximation)

$$x(t) = F(t)x(t) + G(t)u(t) \qquad (3.4\text{-}13)$$

where $x(t)$, $u(t)$ are small departures from the optimal trajectories $\xi^*(t)$, $v^*(t)$ and $F(t)$, $G(t)$ are the first partial derivatives associated with $f(\xi, v, t)$ evaluated along $v^*(\cdot)$, $\xi^*(\cdot)$.

Our concern is with the following issue. Suppose that, for whatever reason, in applying the optimal control to (3.4-11) the state at time t is not $\xi^*(t)$ but is $\xi^*(t) + x(t)$ with $x(t)$ small. What should the optimal control value now become? Intuitively, one would expect a control $v^*(t) + u(t)$ with $u(t)$ small, and related to $x(t)$.

The cost of using $v^*(t) + u(t)$ over $[t_0, T]$ differs from that obtained using $v^*(t)$. The first-order difference is zero, while the second-order difference is (see Appendix C)

$$\delta^2 \overline{V} = \int_{t_0}^{T} [x' \quad u'] \begin{bmatrix} H_{\xi\xi} & H_{\xi v} \\ H_{v\xi} & H_{vv} \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} dt + \tfrac{1}{2} x'(T) m_{\xi\xi} x(T) \qquad (3.4\text{-}14)$$

where

$$H = l(\xi, v, t) + p'f(\xi, v, t) \qquad (3.4\text{-}15)$$

and p is the adjoint variable arising in the determination of $\xi^*(\cdot)$, $v^*(\cdot)$ using the Pontryagin Minimum Principle (see Appendix C). The derivatives in the integral in (3.4-14) are of course evaluated on the optimal trajectory $\xi^*(t)$ with the optimal control $v^*(t)$, and $m_{\xi\xi}$ is evaluated at $\xi^*(T)$.

Obviously, the increment in cost depends on u. It is minimized by taking u as the solution of the linear quadratic problem defined by (3.4-12) and (3.4-14). In relation to our earlier notation,

$$Q(t) = H_{\xi\xi} \qquad R(t) = H_{vv} \qquad S(t) = H_{\xi v} \qquad A = m_{\xi\xi}$$

For these ideas to work, smoothness is essential (and is not present in every optimal control problem). Further, H_{vv} must be positive definite (although v minimizes H, this only guarantees that H_{vv} is nonnegative definite).

Notice that the cross-product term $S(\cdot)$ is not, in general, zero, or even constant. Also we stress that the Q, R, S, A selection depends on the optimal trajectory, assumed calculated off line, by means not explored in this text. It is quite possible that crude approximations to v^*, ξ^*, and also Q, R, S, A can be used effectively in practice. See References [7, 8] for more details.

Discrete-time case.   Analogous "linearizing" and "quadraticizing" procedures as above apply to the discrete-time case. For details see [7].

Main points of the section. More general quadratic performance indices arise in dealing with regulation about a general nonlinear optimal control trajectory. Theory for the case when cross-product terms are involved can be derived from results when they are not present.

Problem 3.4-1. Consider a one-port passive electric network with input current u and voltage y, so that uy represents the instantaneous power flow into the network. Suppose that the network equations are $\dot{x} = Fx + Gu$, $y = H'x + Ju$, with $J > 0$. Suppose the network is initially storing energy, and we seek to extract the maximum energy from the network, that is, we aim to maximize $\int_0^\infty (-uy)\, dt$ or minimize $\int_0^\infty (uy)\, dt$. Show that the associated linear quadratic problem does not have the equivalent of $Q - SR^{-1}S' \geq 0$. Argue that there is no conjugate point for the Riccati equation with boundary condition $P(T, T) = 0$, and that the associated Riccati equation solution is nonpositive definite. [*Hint:* Recognize that a passive network initially storing energy can only deliver a finite amount of energy to the outside world].

Problem 3.4-2. Consider an optimization problem for a linear system $\dot{x} = F(t)x + G(t)u$ and nonquadratic performance index. Suppose that the optimal trajectory $x^*(\cdot)$ and control $u^*(\cdot)$ have been obtained for a certain initial condition, but that the associated costate trajectory (adjoint variable trajectory) $p(\cdot)$ is not known. Show that there is sufficient data to formulate the second variation problem, assuming adequate smoothness.

3.5 REGULATOR WITH A PRESCRIBED DEGREE OF STABILITY

As we show in this section, it is possible to define a modified regulator problem which achieves a closed-loop system with a prescribed degree of stability α. That is, for some prescribed $\alpha > 0$, the states $x(t)$ approach zero at least as fast as $e^{-\alpha t}$ in the continuous time case. We will focus on the time-invariant case, when the optimal controller is constant and achieves closed-loop eigenvalues with real parts less than $-\alpha$. Of course, the larger is α, the more stable is the closed-loop system. A high degree of closed-loop stability may only be achieved at excessive control energy cost, or controller complexity cost, so that the selection of α must be a considered one, as discussed later in the section. The results of this section first appeared in [9].

Modified regulator problem. Consider the system

$$\dot{x} = Fx + Gu \qquad x(t_0) \text{ given} \qquad (3.5\text{-}1)$$

where F and G are constant and the pair $[(F + \alpha I), G]$ is completely stabilizable. Consider also the associated performance index

$$V(x(t_0), u(\cdot), t_0) = \int_{t_0}^{\infty} e^{2\alpha t}(u'Ru + x'Qx)\, dt \qquad (3.5\text{-}2)$$

where R and Q are constant, symmetric, and respectively positive definite and nonnegative definite. Let α be a nonegative constant (which will turn out to be the minimum degree of stability of the closed-loop system). With D any matrix such that $DD' = Q$, let $[F + \alpha I, D]$ be completely detectable. Define the minimization problem as the task of finding the minimum value of the performance index (3.5-2) and the associated optimal control.

The strategy we adopt in solving this modified problem is to introduce transformations that convert the problem to an infinite-time regulator problem of the type considered earlier in the chapter. Accordingly, we make the definitions

$$\hat{x}(t) = e^{\alpha t}x(t) \qquad (3.5\text{-}3)$$

$$\hat{u}(t) = e^{\alpha t}u(t). \qquad (3.5\text{-}4)$$

Just as $x(\cdot)$ and $u(\cdot)$ may be related [via Eq. (3.5-1)], so $\hat{x}(\cdot)$ and $\hat{u}(\cdot)$ may be related. Observe that

$$
\begin{aligned}
\dot{\hat{x}} &= \frac{d}{dt}(e^{\alpha t}x(t)) = \alpha e^{\alpha t}x(t) + e^{\alpha t}\dot{x}(t) \\
&= \alpha\hat{x} + e^{\alpha t}Fx + e^{\alpha t}Gu \qquad (3.5\text{-}5) \\
&= (F + \alpha I)\hat{x} + G\hat{u}
\end{aligned}
$$

Thus, given the relations (3.5-3) and (3.5-4), the system equation (3.5-1) implies the system equation (3.3-5). The converse is clearly true, too. Corresponding initial conditions for the two systems (3.5-1) and (3.5-5) are given by setting $t = t_0$ in (3.5-3)—i.e. $\hat{x}(t_0) = e^{\alpha t_0} x(t_0)$.

The integrand in (3.5-2) may also be written in terms of $\hat{u}$ and $\hat{x}$:

$$e^{2\alpha t}(u'Ru + x'Qx) = \hat{u}'R\hat{u} + \hat{x}'Q\hat{x}.$$

Consequently, we may associate with the system (3.5-5) the performance index

$$\hat{V}(\hat{x}(t_0), \hat{u}(\cdot), t_0) = \int_{t_0}^{\infty} (\hat{u}'R\hat{u} + \hat{x}'Q\hat{x})\, dt \qquad (3.5\text{-}6)$$

Moreover, there is a strong connection between the mimimization problem associated with the equation pair (3.5-1), (3.5-2), and the pair (3.5-5), (3.5-6). Suppose $u^*(t)$ is the value of the optimal control at time t when the initial state is $x(t_0)$. Then the value of the optimal control at time t for the second problem is $\hat{u}^*(t) = e^{\alpha t}u^*(t)$, and the resulting value of the state at time t is given by $\hat{x}(t) = e^{\alpha t}x(t)$, provided $\hat{x}(t_0) = e^{\alpha t_0}x(t_0)$. Also, the minimum performance index is the same for each problem.

Moreover, if the optimal control for the second problem can be expressed in feedback form as

$$\hat{u}^*(t) = k(\hat{x}(t), t) \tag{3.5-7}$$

then the optimal control for the first problem may also be expressed in feedback form; thus

$$u^*(t) = e^{-\alpha t} \hat{u}^*(t) = e^{-\alpha t} k(e^{\alpha t} x(t), t) \tag{3.5-8}$$

[We know that the control law (3.5-7) should be a linear one; and, indeed, we shall shortly note the specific law; the point to observe here is that a *feedback* control law for the second problem readily yields a *feedback* control law for the first problem, irrespective of the notion of linearity.]

Our temporary task is now to study the system (3.5-5), and to select a control $\hat{u}^*(\cdot)$ that minimizes the performance index (3.5-6), where R and Q have the constraints imposed at the beginning of the section.

To guarantee existence of an optimal control, via the theory of an earlier section, we require that $[F + \alpha I, G]$ be completely stabilizable. One of the problems seeks a proof that a sufficient condition is *that $[F, G]$ is completely controllable.*

Given this complete stabilizability constraint, let us apply the material of Sec. 3.3. Let $P(t, T)$ be the solution at time t of the equation

$$-\dot{P} = P(F + \alpha I) + (F' + \alpha I)P - PGR^{-1}G'P + Q \tag{3.5-9}$$

with boundary condition $P(T, T) = 0$. Then

$$\bar{P} = \lim_{t \to -\infty} P(t, T) \tag{3.5-10}$$

exists as a constant matrix, satisfying the steady state version of (3.5-9):

$$\bar{P}(F + \alpha I) + (F' + \alpha I)\bar{P} - \bar{P}GR^{-1}G'\bar{P} + Q = 0 \tag{3.5-11}$$

Then the optimal control becomes

$$\hat{u}^*(t) = -R^{-1}G'\bar{P}\hat{x}(t) \tag{3.5-12}$$

and the closed-loop system is

$$\dot{\hat{x}} = (F + \alpha I - GR^{-1}G'\bar{P})\hat{x} \tag{3.5-13}$$

We recall from the results of an earlier section that a necessary and sufficient condition ensuring asymptotic stability of (3.5-13) is that $[F + \alpha I, D]$ should be completely detectable, where D is any matrix such that $DD' = Q$. (A sufficient condition is that $[F, D]$ is completely observable.)

We can now apply these results to the original optimization problem. Equations (3.5-7) and (3.5-8) show us that

$$u^*(t) = -e^{-\alpha t}R^{-1}G'\bar{P}e^{\alpha t}x(t) = -R^{-1}G'\bar{P}x(t) \tag{3.5-14}$$

This is the desired constant control law; note that it has the same structure as the control law of (3.5-12).

To demonstrate the degree of stability, we have from (3.5-3) that $x(t) = e^{-\alpha t}\hat{x}(t)$. Since the closed-loop system (3.5-13) is asymptotically stable by virtue

of the complete detectability assumption on $[F + \alpha I, D]$, we know that $\hat{x}(t)$ approaches zero as t approaches infinity, and, consequently, that $x(t)$ approaches zero at least as fast as $e^{-\alpha t}$ when t approaches infinity.

The minimum value achieved by (3.5-2) is the same as the minimum value achieved by (3.5-6). As was shown in the previous chapter, the optimal performance index for (3.5-6) is expressible in terms of $\bar{P}$ as $\hat{x}'(t_0)\bar{P}\hat{x}(t_0)$. Consequently, the minimum value achieved by (3.5-2) is $x'(t_0)e^{-2\alpha t_0}\bar{P}x(t_0)$. Let us now summarize the results.

> ***Solution of the regulator problem with prescribed degree of stability.***   The optimal performance index for the modified regulator problem stated at the start of this section is $x'(t_0)e^{-2\alpha t_0}\bar{P}x(t_0)$, where $\bar{P}$ is defined as the limiting solution of the Riccati equation (3.5-9) with boundary condition $P(T, T) = 0$. The matrix $\bar{P}$ also satisfies the algebraic equation (3.5-11). The associated optimal control is given by the constant linear feedback law (3.5-14), and the closed-loop system has degree of stability of at least α.

One might ask if there are positive functions $f(t)$ other than $e^{\alpha t}$ such that minimizing the index $V = \int_{t_0}^{\infty} f(t)(u'Ru + x'Qx)\,dt$ under (3.5-1) leads to a linear constant control law. It is not difficult to show that (apart from trivial situations such as $Q = 0$) essentially the only possible $f(t)$ are those already considered—that is, those of the form $e^{\alpha t}$—with α real.

One might well ask also if it is possible to construct a performance index with α equal to zero such that the control law resulting from the minimization is the same as that obtained from the preceding problem when α is nonzero. The answer is yes (see Problem 3.5-2). In other words, there are sets of pairs of matrices R and Q such that the associated regulator problem with zero α leads to a closed-loop system with degree of stability α. However, it does not appear possible to give an explicit formula for writing down these matrices without first solving a regulator problem with α nonzero.

Let us now return to the question implicit in an earlier discussion. When is the cost of achieving a degree of stability α large? As a preliminary, consider the simpler question, when is the cost of solving a linear quadratic problem likely to be large? We know that a necessary condition for solvability is that the pair $[F, G]$ be stabilizable, so that all unstable modes are controllable. But what if such modes are barely controllable, that is, what if it takes a large amount of control energy to bring to zero an initial state $x(t_0)$ with $\|x(t_0)\| = 1$ and $Fx(t_0) = \lambda x(t_0)$, Re $\lambda > 0$? Then any stabilizing control gain K will necessarily be large, control values will be large, and the optimal cost itself will be large.

Now for the problem of this section, the above argument needs to be modified. Any eigenvalue of the original open-loop system to the right of Re $s = -\alpha$ becomes an eigenvalue of the modified system of (3.5-5) to the right of Re $s = 0$. Since controlling the original system is equivalent in a sense to control of the modified system, it follows that any barely controllable mode of the original system

in Re $s \geq -\alpha$ will give a large value of performance index, control gain, and control signal.

What will happen if we choose $\alpha < 0$? As one might expect, this tends to destabilize the closed-loop system. Recall the example of Section 2.3 when the index $\int_{t_0}^{T} (2e^{-t}u^2 + \frac{1}{2}e^{-t}x^2) \, dt$ is minimized subject to $\dot{x} = \frac{1}{2}x + u$. Here $\alpha = -\frac{1}{2}$ is negative and $Q = \frac{1}{2}$, $R = 2$, $F = \frac{1}{2}$, $G = 1$. The optimal control, recall is

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

Notice that as $T \to \infty$, $u(t) \approx -\frac{1}{2}x(t)$, $\dot{x}(t) \approx 0$ and there is only a marginal stability result.

Consider the general scalar example when $\int_{t_0}^{\infty} e^{2\alpha t}(ru^2 + qx^2) \, dt$ is minimized subject to $\dot{x} = fx + gu$ with $q \neq 0$, $g \neq 0$, $\alpha \geq 0$. The results for the case $\alpha = 0$ given in Section 3.1 can be applied here with f replaced by $\alpha + f$. Thus now

$$\bar{P}_2 = \frac{(\alpha + f) + [(\alpha + f)^2 + g^2 q/r]^{1/2}}{r^{-1}g^2} > 0$$

$$u^* = -\frac{(\alpha + f) + [(\alpha + f)^2 + g^2 q/r]^{1/2}x}{g}$$

$$\dot{x} = -\{\alpha + [(\alpha + f)^2 + g^2 q/r]^{1/2}\}x$$

Notice that as α increases $\bar{P}_2$ increases, the controller gain increases, and the closed-loop bandwidth increases.

By way of another example, consider an idealized angular position control system where the position of the rotation shaft is controlled by the torque applied, with no friction in the system. The equation of motion is

$$J\ddot{\theta} = T$$

where θ is the angular position, T is the applied torque, and J is the moment of inertia of the rotating parts. In state-space form, this becomes

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}x + \begin{bmatrix} 0 \\ a \end{bmatrix}u$$

where $x_1 = \theta$, $x_2 = \dot{\theta}$, $u = T$, and $a = 1/J$. As a performance index guaranteeing a degree of stability α, we choose

$$\int_0^\infty e^{2\alpha t}(u^2 + x_1^2) \, dt$$

The appropriate algebraic equation satisfied by $\bar{P}$ is

$$\bar{P}\begin{bmatrix} \alpha & 1 \\ 0 & \alpha \end{bmatrix} + \begin{bmatrix} \alpha & 0 \\ 1 & \alpha \end{bmatrix}\bar{P} - \bar{P}\begin{bmatrix} 0 & 0 \\ 0 & a^2 \end{bmatrix}\bar{P} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = 0$$

One possible way to solve this equation is to write down simultaneous equations for the entries of $\bar{P}$; thus,

$$2\alpha \bar{p}_{11} - \bar{p}_{12}^2 a^2 + 1 = 0$$

$$\bar{p}_{11} + 2\alpha \bar{p}_{12} - \bar{p}_{12}\bar{p}_{22}\, a^2 = 0$$

$$2\bar{p}_{12} + 2\alpha \bar{p}_{22} - \bar{p}_{22}^2\, a^2 = 0$$

Substitution for $\bar{p}_{11}$ from the first equation, $\bar{p}_{12}$ from the third equation, into the second equation gives a fourth-order equation, yielding solutions that can be used to find $\bar{P} > 0$. For the case $\alpha = 1$, a positive definite $\bar{P}$ is verified to be given from

$$\bar{p}_{11} = \frac{1}{a^2}\left[2 + 2\sqrt{1 + a^2} + \frac{1}{2}(2 + 2\sqrt{1 + a^2})^{3/2}\right]$$

$$\bar{p}_{12} = \frac{1}{a^2}[1 + \sqrt{1 + a^2} + \sqrt{2 + 2\sqrt{1 + a^2}}]$$

$$\bar{p}_{22} = \frac{1}{a^2}[2 + \sqrt{2 + 2\sqrt{1 + a^2}}]$$

The optimal control law is

$$u^* = -g'\bar{P}x$$

$$= \begin{bmatrix} -\dfrac{1}{a}(1 + \sqrt{1 + a^2} + \sqrt{2 + 2\sqrt{1 + a^2}}) \\[2mm] -\dfrac{1}{a}(2 + \sqrt{2 + 2\sqrt{1 + a^2}}) \end{bmatrix}' \begin{bmatrix} x_1 \\[2mm] x_2 \end{bmatrix}$$

This is implementable with proportional plus derivative (in this case, tacho, or angular velocity) feedback. The closed-loop system equation is

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ -(1 + \sqrt{1 + a^2} + \sqrt{2 + 2\sqrt{1 + a^2}}) & -(2 + \sqrt{2 + 2\sqrt{1 + a^2}}) \end{bmatrix} x$$

for which the characteristic polynomial is

$$s^2 + (2 + \sqrt{2 + 2\sqrt{1 + a^2}})s + (1 + \sqrt{1 + a^2} + \sqrt{2 + 2\sqrt{1 + a^2}})$$

It is readily checked that the roots of this polynomial are complex for all a; therefore, the real part of the closed-loop poles is

$$-1 - \tfrac{1}{2}\sqrt{2 + 2\sqrt{1 + a^2}} < -1$$

Thus, the requisite degree of stability is achieved.

Discrete-time results.   Analogous results accrue for the discrete-time case. The relevant index is

$$V(x(t_0), u(\cdot), t_0) = \sum_{t = t_0 + 1}^{\infty} \lambda^{-2t}[x'(t)Qx(t) + u'(t-1)Ru(t-1)] \qquad (3.5\text{-}15)$$

for some $0 < \lambda \leq 1$. Now F in the standard Riccati equation is replaced by $\lambda^{-1}F$, to yield control laws that achieve a degree of stability λ—that is, all closed-loop eigenvalues are less than λ in magnitude. Further details are left to the reader in one of the problems.

Main points of the section. Introducing a weighting term into the performance index allows us to achieve a prescribed degree of stability in the closed-loop system. This must be applied with caution in those cases where stable, lightly damped open-loop system modes are present.

Problem 3.5-1. Consider the system (with constant F and G)

$$\dot{x} = Fx + Gu \qquad x(t_0) \text{ given}$$

and the associated performance index

$$\int_{t_0}^{\infty} e^{2\alpha t} [(u'u)^5 + (x'Qx)^5] \, dt$$

where Q is a constant nonnegative definite matrix. Find a related linear system and performance index where the integrand in the performance index is not specifically dependent on time (although, of course, it depends on u and x). Show that if an optimal feedback law exists for this related system and performance index, it is a constant law, and that from it an optimal feedback law may be determined for the original system and performance index.

Problem 3.5-2. Consider the system

$$\dot{x} = Fx + Gu \qquad x(t_0) \text{ given}$$

where F and G are constant and $[F, G]$ is completely controllable. Show that associated with any performance index of the form

$$\int_{t_0}^{\infty} e^{2\alpha t}(u'Ru + x'Qx) \, dt$$

where R is constant and positive definite, Q is constant and nonnegative definite, and α is positive, there is a performance index

$$\int_{t_0}^{\infty} (u'Ru + x'\hat{Q}x) \, dt$$

where $\hat{Q}$ is constant and nonnegative definite, such that the optimal controls associated with minimizing these indices are the same. [*Hint:* Define $\hat{Q}$, using the solution of the first minimization problem.]

Problem 3.5-3. Consider the second-order example of the section for the case $\alpha = 1$, $a^2 = 3$. Derive an expression for $\partial \bar{P}/\partial \alpha$. Use this to estimate the effect on the control law, and closed-loop poles when $\alpha = 2$.

Problem 3.5-4. Consider the modified regulator problem as stated at the beginning of the section, and let $\bar{P}$ be defined as in Eqs. (3.5-9) through (3.5-11). Suppose that $[F, D]$ is completely observable. Show that the degree of stability result follows by using the Lyapunov function $V = x'\bar{P}x$.

Problem 3.5-5.   Prove that complete controllability of $[F, G]$ implies and is implied by complete controllability of $[F + \alpha I, G]$. Conclude that complete controllability of $[F, G]$ implies $[F + \alpha I, G]$ is stabilizable.

Problem 3.5-6.   Imagine two optimization problems of the type considered in this chapter with the same F, G, Q, and R but with two different α—viz., α_1 and α_2, with $\alpha_1 > \alpha_2$. Show that $\bar{P}_{\alpha 1} - \bar{P}_{\alpha 2}$ is positive definite. [*Hint:* Refer to Problem 3.5-3.]

Problem 3.5-7.   Recall the worked example in the text where

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ a \end{bmatrix} u$$

and the performance index is $\int_0^\infty e^{2\alpha t}(u^2 + x_1^2)\ dt$. Show that the closed-loop characteristic polynomial is $(s + \alpha)^2 + [2\alpha^2 + 2\sqrt{\alpha^4 + a^2}]^{1/2}(s + \alpha) + \sqrt{\alpha^4 + a^2}$. [*Hint:* Using Appendix E, write down the Hamiltonian matrix for the modified problem and evaluate its characteristic polynomial. Factor this characteristic polynomial as $p(s)p(-s)$ where $p(s)$ is stable. Argue that $p(s)$ is the characteristic polynomial for the closed loop of the modified system, and conclude the result.]

Problem 3.5-8.   Using the performance index (3.5-15), develop the results of this section for the discrete-time case.

REFERENCES

[1] P. Sannuti and P. V. Kokotovic, "Near Optimum Design of Linear Systems by a Singular Perturbation Method," *IEEE Trans. Auto. Control,* Vol. AC-14, No. 1 (February 1969), pp. 15–21.

[2] R. E. Kalman, "Contributions to the Theory of Optimal Control," *Bol. Soc. Matem. Mex.,* (1960), pp. 102–119.

[3] B. D. O. Anderson and J. B. Moore, "Detectability and Stabilizability of Time-varying Discrete-time Linear Systems," *SIAM J. of Control and Optimization,* Vol. 18, No. 1 (January 1981), pp. 20–32.

[4] R. E. Kalman, "When is a Linear Control System Optimal?", *Trans. ASME Ser. D: J. Basic Eng.,* Vol. 86 (March 1964), pp. 1–10.

[5] W. M. Wonham, "On a Matrix Riccati Equation of Stochastic Control," *SIAM J. Control,* Vol. 6, No. 4 (1968), pp. 681–697.

[6] K. Mårtensson, "On the Matrix Riccati Equation," *Information Sci.,* Vol. 3 (1971), pp. 17–49.

[7] F. L. Lewis, *Optimal Control.* New York: John Wiley and Sons, 1986.

[8] A. E. Bryson, Jr, and Y. C. Ho, *Applied Optimal Control.* New York: Hemisphere, 1975.

[9] B. D. O. Anderson and J. B. Moore, "Linear System Optimization with Prescribed Degree of Stability," *Proc. IEE,* Vol. 116, No. 12 (December 1969), pp. 2083–2087.

4

Tracking Systems

4.1 THE PROBLEM OF ACHIEVING
A DESIRED TRAJECTORY

In previous chapters, the regulator problem—viz., the problem of returning a
system to its zero state in some optimal fashion—is considered. This problem is, in
fact, a special case of a wider class of problems where it is required that the outputs
of a system follow or track a desired trajectory in some optimal sense. For the
regulator, the desired trajectory is, of course, simply the zero state. In this chapter,
we apply regulator theory and give extensions to solve the wider class of control
problems that involves achieving a desired trajectory.

The regulator theory developed so far results in linear nondynamic (propor-
tional) state feedback controllers, although we have foreshadowed the use of state
estimation and state estimate feedback resulting in dynamic feedback controllers.
In the trajectory following theory of this chapter, based on the earlier regulator
results, the controllers consist of state (or state estimate) proportional feedback
controllers together with feed-forward controllers involving processing of the de-
sired trajectory. Of course, more general regulators/trackers involving dynamic
state (or state estimate) feedback such as proportional plus integral state (or state
estimate) feedback may be more useful in certain applications. For example, a
particularly common form of servo problem involves having the output of a plant
track a step input (set-point control). Such trajectory problems are often encoun-
tered in classical control, and integral control is often a feature of the controller

design. The controller design of this chapter will, however, make limited contact with this idea; later in the book, particularly in our discussion of frequency shaping in Chapter 9, we will return to this classical idea. See also Problem 4.3-5 of this chapter.

It is convenient to refer to trajectory following problems by one of three technical terms, the particular term used depending on the nature of the desired trajectory. If the plant outputs are to follow a *class* of desired trajectories, for example, all polynomials up to a certain order, the problem is referred to as a *servo* (servomechanism) *problem*; if the desired trajectory is a *particular* prescribed function of time, the problem is called a *tracking problem*. When the outputs of the plant are to follow the response of another plant (or model), the problem is referred to as the *model-following problem*.

The remainder of this section is devoted to a discussion of considerations common to all three of these problems, with particular attention being given to the selection of a performance index.

We recall that in selecting a performance index for a regulator, cost terms are constructed for the control energy and the energy associated with the states. More specifically, for the linear system

$$\dot{x} = Fx + Gu \qquad x(t_0) \text{ given} \tag{4.1-1}$$

the principal performance index adopted throughout the book is the quadratic index

$$V(x(t_0), u(\cdot), T) = \int_{t_0}^{T} (u'Ru + x'Qx)\, dt \tag{4.1-2}$$

where R is some positive definite matrix and Q is some nonnegative definite matrix (the matrices being of appropriate dimensions). The quadratic nature of the cost terms ensures that the optimal law is linear, and the constraints on the matrices Q and R ensure that the control law leads to a finite control.

When one is attempting to control the system (4.1-1) such that its output $y(\cdot)$ given by

$$y = H'x \tag{4.1-3}$$

tracks a desired trajectory $\bar{y}(\cdot)$, there clearly should be a cost term in the performance index involving the error $(y - \bar{y})$. A performance index that comes to mind immediately as a natural extension of the index (4.1-2) is the following

$$V(x(t_0), u(\cdot), T) = \int_{t_0}^{T} [u'Ru + (y - \bar{y})'Q(y - \bar{y})]\, dt \tag{4.1-4}$$

where Q is nonnegative definite and R is positive definite. For ease of presentation, we will neglect in this section any terminal cost term $[y(T) - \bar{y}(T)]'\bar{A}[y(T) - \bar{y}(T)]$, $\bar{A} = \bar{A}' \geq 0$. Once again, we have quadratic terms that, as the next sections show, give rise to linear control laws.

Attempting to minimize $y - \bar{y}$ amounts to attempting to constrain $H'x$ (in

general to a nonzero value). If H has rank m, this imposes m constraints on x. It is clearly legitimate to aim for $n - m$ further constraints on x without creating a conflict of objectives. The right way to do this is as follows: generalize the performance index (4.1-4) to

$$V(x(t_0), u(\cdot), t_0) = \int_{t_0}^T [u'Ru + \bar{y}'Q_1\bar{y} + (y - \bar{y})'Q_2(y - \bar{y})]\, dt \qquad (4.1\text{-}5)$$

where Q_1 and Q_2 are nonnegative definite symmetric matrices and

$$\bar{y} = \bar{H}'x \qquad \bar{H}' = I - LH' \qquad L = H(H'H)^{-1} \qquad (4.1\text{-}6)$$

Notice that $H'\bar{y} = 0$.

It is now immediate that the index (4.1-5) may be written in the convenient form

$$V(x(t_0), u(\cdot), t_0) = \int_{t_0}^T [u'Ru + (x - \bar{x})'Q(x - \bar{x})]\, dt \qquad (4.1\text{-}7)$$

where

$$Q = \bar{H}Q_1\bar{H}' + HQ_2H', \qquad \bar{x} = L\bar{y} \qquad (4.1\text{-}8)$$

The interpretation of the terms in the preceding index is straightforward enough. It appears that the cost terms of the index (4.1-5) involving the state and the error between the system output and the desired output are replaced by a single term involving the error between the state and some special state trajectory $\bar{x}$, related to the desired output trajectory $\bar{y}$. We have from (4.1-8) that $H'\bar{x} = \bar{y}$, and thus, if by some means $\bar{x}$ were to become the state trajectory x, then the system output y would be the desired output trajectory $\bar{y}$. What characterizes the trajectory $\bar{x}$ is the important property, that its component in the null space of H' is zero. The specified properties of $\bar{x}$ suggest that it be referred to as the *desired state trajectory*.

For any particular application of the performance index just developed, selections have to be made of the matrices Q_1 and Q_2 and R. Also, a selection may have to be made of the terminal time T. It may be necessary to try a range of values of these quantities and to select the particular one that is most appropriate for any given situation. A case of particular interest is the limiting case as the terminal time T approaches infinity. For this case, when all the matrices are time invariant, part, if not all, of the optimal controller becomes time invariant. However, difficulties may arise for this case, because it may not be possible for the plant with a finite control law to track the desired trajectories so that the error approaches zero as time becomes infinite. In particular, if dim $\bar{y} >$ dim u, one could not expect generally to ever get perfect tracking of $\bar{y}$ by y, even asymptotically. Moreover, even if it is possible to track in this sense by using a finite control law, unless the control approaches zero as time approaches infinity, other difficulties arise due to the fact that the performance index would be infinite for all controls and therefore attempts at its minimization would be meaningless.

The next section considers finite terminal-time servo-, tracking-, and model-

following problems. The final section considers the limiting situation as the integration interval becomes infinite. Of particular interest in this section are the cases where the desired trajectory is either a step, ramp, or parabolic function. Key references for the material that follows are [1] through [4].

Main points of the section.   The quadratic performance index (4.1-7) is a natural one in conjunction with plants (4.1-1) so as to achieve state trajectories x that track closely a specified trajectory $\bar{x}$. When the objective is for $y = H'x$ to track closely $\bar{y}$, then again such an index is convenient where $\bar{x}$ and Q are specified as in (4.1-8).

Problem 4.1-1.   Verify that under (4.1-6) the second and third terms in the index (4.1-5) are not conflicting. [*Hint:* First decompose x into the sum of two orthogonal components, one in the range space of H, and show that the second and third terms depend separately on these two orthogonal components].

Problem 4.1-2.   Verify that the index (4.1-7) under (4.1-6) and (4.1-8) is equivalent to the index (4.1-5). [*Hint:* Show first that $\bar{x}_2$, the component of $\bar{x}$ in the null space of H', is zero, so that $\bar{x}'\bar{H}Q_1\bar{H}'\bar{x} = 0$.]

4.2 FINITE-TIME RESULTS

The servo problem.   As stated in the previous section, the servo problem is the task of controlling a system so that the system output follows a reference signal, where all that is known about the signal is that it belongs to a known class of signals, such as step changes. We consider a useful servo problem.

Optimal servo problem.   Suppose we are given the n-dimensional linear system having state equations

$$\dot{x} = Fx + Gu \qquad x(t_0) \text{ given} \tag{4.2-1}$$

$$y = H'x \tag{4.2-2}$$

where the m entries of y are generically linearly independent or, equivalently, the matrix H has rank m. Suppose we are also given an m-vector incoming reference signal $\bar{y}$, which is the output of the known p-dimensional linear reference model

$$\dot{z} = Az \tag{4.2-3}$$

$$\bar{y} = C'z \tag{4.2-4}$$

for some initial state $z(t_0)$. Without loss of generality, the pair $[A, C]$ is completely observable. The optimal servo problem is to find the optimal

control u^* for the system (4.2-1), such that the output y tracks the incoming signal $\bar{y}$, minimizing the index

$$V(x(t_0), u(\cdot), t_0) = \int_{t_0}^{T} \{u'Ru + (x - \bar{x})'Q(x - \bar{x})\}\, dt$$

$$\bar{x} = L\bar{y} \tag{4.2-5}$$

where Q is nonnegative definite symmetric, and R is positive definite symmetric. Here Q, L are to be identified as in (4.1-6) and (4.1-8). The selections for Q, L in the index (4.2-5) can be guided by the discussions of the previous section to ensure appropriate penalties for control cost, tracking error cost, and state excitation costs. (As usual, all the various matrices are assumed to have continuous entries.)

Observe that we are requiring that our desired trajectory $\bar{y}$ be derived from a linear differential equation. This, of course, rules out trajectories $\bar{y}$, which have discontinuities for $t > t_0$. We also note that the special case when C is a vector and A is given by

$$A = \begin{bmatrix} 0 & 1 & \cdot & \cdot & \cdot & 0 \\ \cdot & 0 & 1 & & & \cdot \\ \cdot & & & \cdot & \cdot & \cdot \\ \cdot & & & & \cdot & \cdot \\ \cdot & & & & \cdot & 1 \\ 0 & \cdot & \cdot & \cdot & \cdot & 0 \end{bmatrix}$$

leads to the class of $\bar{y}$ consisting of all polynomials of degree $(p - 1)$.

Throughout the book, various minimization problems are solved by first applying a transformation to convert the minimization problem to a standard regulator problem. The standard regulator results are then interpreted, using the transformations to give a solution to the original minimization problem. This will also be our method of attack here. To convert the preceding servo problem to a regulator problem, we require the following assumption, the relaxation of which will be discussed subsequently.

Assumption 4.2-1

The reference model state z is directly measurable.

We now define a new variable

$$\hat{x} = \begin{bmatrix} x \\ z \end{bmatrix} \tag{4.2-6}$$

and new matrices

$$\hat{F} = \begin{bmatrix} F & 0 \\ 0 & A \end{bmatrix} \qquad \hat{G} = \begin{bmatrix} G \\ 0 \end{bmatrix}$$

$$\hat{Q} = \begin{bmatrix} Q & -QLC' \\ -CL'Q & CL'QLC' \end{bmatrix} \tag{4.2-7}$$

These variables and matrices are so constructed that when applied to the problem of minimizing (4.2-5) with the relationships (4.2-1) and (4.2-3) holding, we have the standard regulator problem requiring minimization of the quadratic index

$$V(\hat{x}(t_0), u(\cdot), t_0) = \int_{t_0}^{T} (u'Ru + \hat{x}'\hat{Q}\hat{x}) \, dt \qquad (4.2\text{-}8)$$

with relationship

$$\dot{\hat{x}} = \hat{F}\hat{x} + \hat{G}u \qquad \hat{x}(t_0) \text{ given} \qquad (4.2\text{-}9)$$

holding. This result is readily checked.

Applying the regulator theory of Chapter 3, to the minimization problem, (4.2-8) and (4.2-9), gives immediately that the optimal control u^* is

$$u^* = -R^{-1}\hat{G}'\hat{P}\hat{x} \qquad (4.2\text{-}10)$$

where $\hat{P}(\cdot)$ is the solution of the Riccati equation

$$-\dot{\hat{P}} = \hat{P}\hat{F} + \hat{F}'\hat{P} - \hat{P}\hat{G}R^{-1}\hat{G}'\hat{P} + \hat{Q} \qquad \hat{P}(T) = 0 \qquad (4.2\text{-}11)$$

The minimum index is

$$V^*(\hat{x}(t_0), t_0) = \hat{x}'(t_0)\hat{P}(t_0)\hat{x}(t_0) \qquad (4.2\text{-}12)$$

We now interpret these results in terms of the variables and matrices of the original problem, using the definitions (4.2-6) and (4.2-7). First, we partition $\hat{P}$ as

$$\hat{P} = \begin{bmatrix} P & P_{12} \\ P_{12}' & P_{22} \end{bmatrix} \qquad (4.2\text{-}13)$$

where P is an $n \times n$ matrix. Substituting (4.2-7) and (4.2-13) into (4.2-10) gives the optimal control u^* as

$$u^* = K'x + K_1'z \qquad (4.2\text{-}14)$$

where

$$K' = -R^{-1}G'P \qquad (4.2\text{-}15)$$

$$K_1' = -R^{-1}G'P_{12} \qquad (4.2\text{-}16)$$

The Riccati equation (4.2-11) becomes now

$$-\dot{P} = PF + F'P - PGR^{-1}G'P + Q \qquad (4.2\text{-}17)$$

$$-\dot{P}_{12} = P_{12}A + F'P_{12} - PGR^{-1}G'P_{12} - QLC' \qquad (4.2\text{-}18)$$

$$-\dot{P}_{22} = P_{22}A + A'P_{22} - P_{12}'GR^{-1}G'P_{12} + CL'QLC' \qquad (4.2\text{-}19)$$

with boundary conditions $P(T) = 0$, $P_{12}(T) = 0$, and $P_{22}(T) = 0$. The minimum index is

$$V^*(x(t_0), t_0) = x'(t_0)P(t_0)x(t_0) + 2x'(t_0)P_{12}(t_0)z(t_0)$$
$$+ z'(t_0)P_{22}(t_0)z(t_0) \qquad (4.2\text{-}20)$$
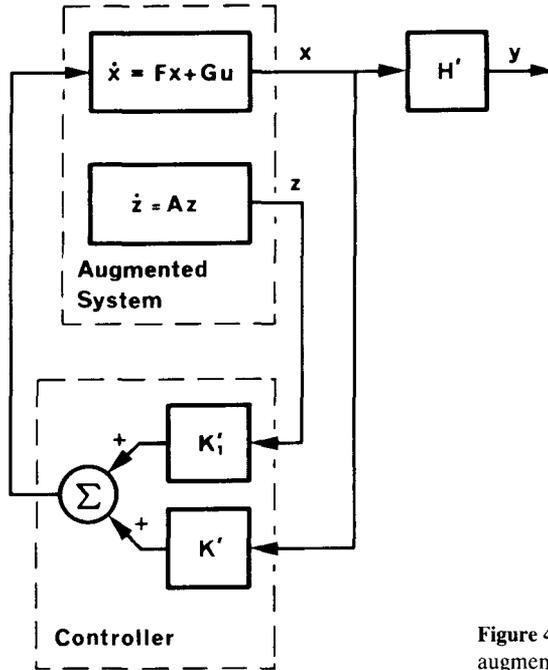
Figure 4.2-1 Regulator control of augmented system.

Figure 4.2-1 shows the augmented system (4.2-9) separated into its component systems (4.2-1) and (4.2-3) and controlled by use of linear state-variable feedback, as for a standard regulator. Figure 4.2-2 shows the same system redrawn as the solution to the servo problem. We observe that it has the form of a regulator designed by minimizing the index

$$V(x(t_0), u(\cdot), t_0) = \int_{t_0}^{T} (u'Ru + x'Qx)\, dt$$

for the system (4.2-1) using regulator theory. There is the addition of an external input, which is the state z of the linear system (4.2-3) and (4.2-4). *The feedback part of the control is independent of A, C, and $z(t_0)$.*

The optimal servo in the finite time case is what is known as a two-degrees-of-
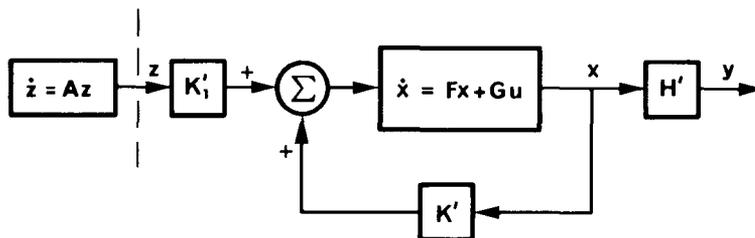
Figure 4.2-2   Rearrangement of regulator system of Fig. 4.2-1.
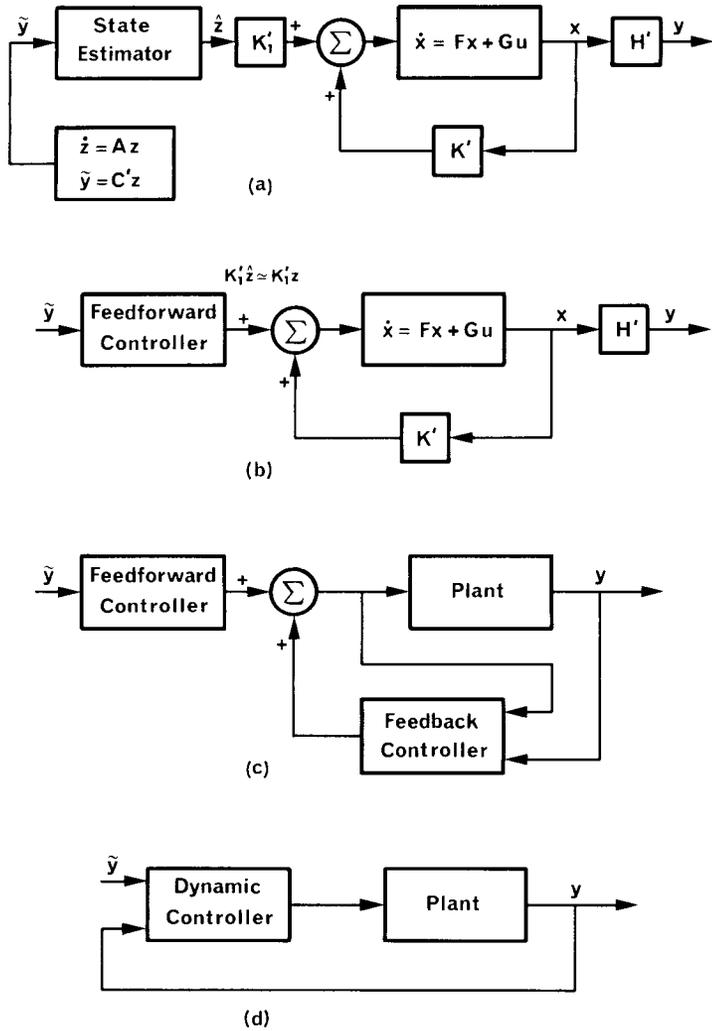
Figure 4.2-3   Four forms for the optimal servo-system.

freedom controller with time-varying feed-forward and feedback gains. In contrast the optimal regulator is a one-degree-of-freedom controller with but a feedback time-varying gain.

The results to this point depend on Assumption 4.2-1, namely, that the state z is directly measurable. Certainly, if z is available, the servo problem is solved. However, often in practice, only an incoming signal $\tilde{y}$ is at hand. For this case, a state estimator may be constructed with $\tilde{y}$ as input and, of course, an estimate $\hat{z}$ of z as output. With the pair $[A, C]$ appearing in (4.2-3) and (4.2-4) completely observable, the estimator may be constructed by using the results of Chapter 7. If A and C are

Command Signal   **Desired Trajectory**

$$\dot{z}_2 = A_2 z_2$$

$$r = C_2' z_2$$

r

$$\dot{z}_1 = A_1 z_1 + B_1 r$$

$$\tilde{y} = C_1' z_1$$

$\tilde{y}$

Command Generator   **Model**
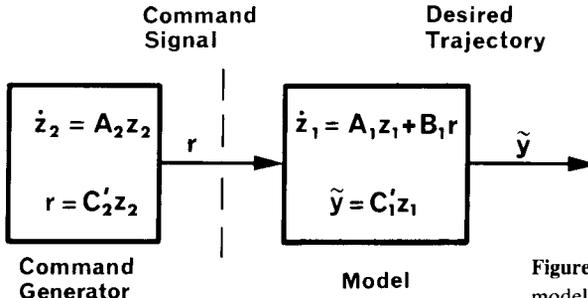
Figure 4.2-4 Desired trajectory for model-following problem.

constant, the estimator can achieve $\hat{z}$ approaching z arbitrarily fast. The resulting system is shown in Fig. 4.2-3(a). It is redrawn in Fig. 4.2-3(b) to illustrate that the state estimator for $\hat{z}$ and control law K_1' can be combined to give a (dynamic) feedforward controller.

Likewise, if the state x of the plant is not directly measurable, the memoryless linear state feedback may be replaced by a dynamic controller, which estimates the state x and then forms the appropriate linear transformation of this estimate, always assuming complete observability of $[F, H]$ (see Chapter 7). Figure 4.2-3(c) shows this arrangement. Figure 4.2-3(d) shows one further possibility, where the estimation of x and z is carried out simultaneously in the one estimator; this arrangement may yield a reduction in the dimension of the linear system comprising the controller. We shall now summarize the optimal servo problem solution.

> **Solution to the finite-time optimal servo problem.** For the systems (4.2-1), (4.2-2), and (4.2-3), (4.2-4) and performance index (4.2-5), the optimal control u^* is given by (4.2-14) through (4.2-19). The minimum index is given in (4.2-20). With Assumption 4.2-1 holding, the form of the optimal controller is indicated in Fig. 4.2-2(a). If only an estimate $\hat{z}$ is available, then an approximation to the optimal controller is as indicated in Fig. 4.2-3(b). The closer the estimate $\hat{z}$ is to z, the better is the approximation. Further possibilities are shown in Figs. 4.2-3(c) and 4.2-3(d), where estimation of x is required.

We remark that at no stage in the above analysis did we use the fact that Q, L possessed the special structure of (4.1-8). Of course, it is this special structure which gives the performance index (4.2-5) meaning for the servo problem so the benefit of relaxing the constraints on Q and L is not clear.

Model-following (servo) problem. This problem is a mild generalization of the servo problem, and is stated as follows.

> *Optimal model-following (servo) problem.* Find a control u* for the linear system (4.2-1) (4.2-2) which minimizes the index (4.2-5) where Q is nonnegative definite symmetric, R is positive definite symmetric, and $\bar{y}$ is the response of a linear system or model

$$\dot{z}_1 = A_1 z_1 + B_1 r \qquad z_1(t_0) \text{ given}$$

$$\tilde{y} = C_1' z_1 \tag{4.2-21}$$

to command inputs r, which, in turn, belong to the class of zero input responses of the system

$$\dot{z}_2 = A_2 z_2 \qquad z_2(t_0) \text{ given}$$

$$r = C_2' z_2 \tag{4.2-22}$$

as indicated in Fig. 4.2-4.

The two systems, (4.2-21) and (4.2-22), together form a linear system

$$z = Az \tag{4.2-3}$$

$$\tilde{y} = C'z \tag{4.2-4}$$

where $z = [z_1' \quad z_2']'$ and the matrices A and C' are given from

$$A = \begin{bmatrix} A_1 & B_1 C_2' \\ 0 & A_2 \end{bmatrix} \quad C' = [C_1' \quad 0] \tag{4.2-23}$$

For the case when z_1 and z_2 are available, the equations for the solution to the model-following problem are identical to those for the servo problem, with A and C given by (4.2-23). In case of nonavailability of z_2, state estimation is required, again in the same manner as in the servo problem.

The tracking problem.   It may be that the desired trajectory $\tilde{y}(t)$ for all t in the range $t_0 \le t \le T$ is known a priori. Such a tracking problem arises, for example, in the altitude control of a terrain-following aircraft, where there is knowledge of the future terrain. To address the tracking problem, we shall set up temporarily a servo problem, and then show that with knowledge of $\tilde{y}(t)$ for all t, it is not necessary to have available the servo model state or its estimate. This represents a considerable saving if $\tilde{y}$ is the output of a high-order system.

We now define the optimal tracking problem and give its solution.

Optimal tracking problem.   Suppose we are given the n-dimensional linear system having state equations (4.2-1) and (4.2-2), where the m entries of y are linearly independent. Suppose we are also given an m vector $\tilde{y}(t)$ for all t in the range $t_0 \le t \le T$ for some times t_0 and T with $t_0 < T$. The optimal tracking problem is to find the optimal control u^* for the system (4.2-1), such that the output y tracks the signal $\tilde{y}$, minimizing the index (4.2-5), where Q is nonnegative definite symmetric and R is positive definite symmetric. In general, L and Q are as in (4.1-6) and (4.1-8).

We first make the following temporary assumption.

Temporary Assumption 4.2-2

The vector $\tilde{y}(t)$ for all t in the range $t_0 \leq t \leq T$ is the output of a linear finite dimensional system

$$\dot{z} = Az \qquad (4.2\text{-}3)$$

$$\tilde{y} = C'z \qquad (4.2\text{-}4)$$

with the pair $[A, C]$ *not* necessarily assumed to be completely observable.

With this assumption holding, the optimal control u^* is given using the optimal servo results (4.2-14) through (4.2-19) in terms of $b \stackrel{\Delta}{=} P_{12}z$ as

$$u^* = K'x + u_{\text{ext}} \qquad (4.2\text{-}24)$$

where

$$u_{\text{ext}} = -R^{-1}G'b \qquad (4.2\text{-}25)$$

The matrix K is calculated as before from (4.2-15) and (4.2-17). See Fig. 4.2-5. Moreover, the minimum index V^*, given from (4.2-20), may be written by using b and $c \stackrel{\Delta}{=} z'P_{22}z$ as

$$V^*(x(t_0), t_0) = x'(t_0)P(t_0)x(t_0) + 2x'(t_0)b(t_0) + c(t_0) \qquad (4.2\text{-}26)$$

The values of the matrices P_{12} and P_{22} and the vector z cannot be determined independently unless the matrices A and C are known. However, the products $b = P_{12}z$ and $c = z'P_{22}z$ can be determined directly from $\tilde{y}(\cdot)$, as follows.

Differentiating the product $(P_{12}z)$ and applying (4.2-18) for P_{12} and (4.2-3) for z, we get

$$-\frac{d}{dt}(P_{12}z) = -\dot{P}_{12}z - P_{12}\dot{z}$$

$$= F'P_{12}z + P_{12}Az - PGR^{-1}G'P_{12}z - QLC'z - P_{12}Az$$

$$= (F - GR^{-1}G'P)'(P_{12}z) - QL\tilde{y}$$

with the boundary condition $P_{12}(T)z(T) = 0$, following from $P_{12}(T) = 0$. This means that with $\tilde{y}(t)$ known for all t in the range $t_0 \leq t \leq T$, the term b can be calculated from the linear differential equation

$$-\dot{b} = (F - GR^{-1}G'P)'b - Q\tilde{x} \qquad b(T) = 0 \qquad (4.2\text{-}27)$$
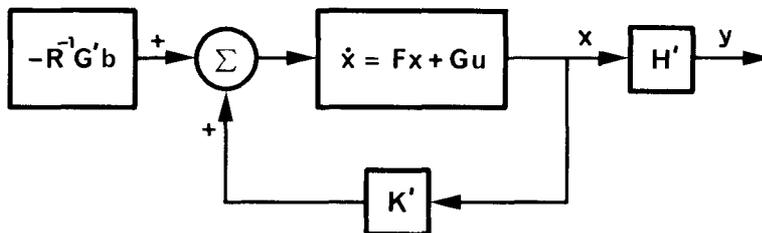
Figure 4.2-5 Optimal tracking system.

The optimal control law (4.2-24) and (4.2-25) is therefore realizable without recourse to using z, or an estimate $\hat{z}$ of z: This equation is solved, backward in time, to determine $b(t)$, which is then used in the optimal control law implementation. Matrices A and C do not play any role in determining u^*.

An equation for $c(\cdot)$ is determined by differentiating $c = z'P_{22}z$ as follows:

$$\frac{d}{dt}(z'P_{22}z) = z'\dot{P}_{22}z + 2z'P_{22}\dot{z}$$

$$= z'P_{12}'GR^{-1}G'P_{12}z - z'CL'QLC'z$$

Using the identifications $b = P_{12}z$ and $c = z'P_{22}z$, we have that $c(\cdot)$ is the solution of the differential equation

$$\dot{c} = b'GR^{-1}G'b - \tilde{x}'Q\tilde{x} \qquad c(T) = 0 \qquad (4.2\text{-}28)$$

We observe by using (4.2-27) and (4.2-28) that the matrices A and C do not play any role in determining V^* from (4.2-26).

Since the differential equations for $b(\cdot)$ and $c(\cdot)$ can, in fact, be solved without Temporary Assumption 4.2-2, we indicate in outline a procedure for verifying that the control given by (4.2-24) and (4.2-25) is optimal without Assumption 4.2-2. With u^* so defined, but not assumed optimal, and with $V^*(x(t_0), t_0)$ defined as in (4.2-26)—again, of course, not assumed optimal—Problem 4.2-1 asks for the establishing of the following identity:

$$V(x(t_0), u(\cdot), t_0) = \int_{t_0}^{T} (u - u^*)'R(u - u^*)\,dt + V^*(x(t_0), t_0) \qquad (4.2\text{-}29)$$

Here, $u(\cdot)$ is an arbitrary control. Optimality of u^* and V^* is then immediate. The preceding results are now summarized.

Solution to the finite-time tracking problem.   For the system (4.2-1) and (4.2-2), and performance index (4.2-5), with the desired trajectory $\tilde{x} = H(H'H)^{-1}\tilde{y}(t)$ available for all t in the range $t_0 \le t \le T$, the optimal control u^* is given from

$$u^* = -R^{-1}G'(Px + b) = K'x - R^{-1}G'b$$

where $P(\cdot)$ is the solution of

$$-\dot{P} = PF + F'P - PGR^{-1}G'P + Q \qquad P(T) = 0$$

and $b(\cdot)$ is the solution of

$$-\dot{b} = (F - GR^{-1}G'P)'b - Q\tilde{x} \qquad b(T) = 0 \qquad (4.2\text{-}27)$$

The minimum index is

$$V^*(x(t_0), t_0) = x'(t_0)P(t_0)x(t_0) + 2x'(t_0)b(t_0) + c(t_0) \qquad (4.2\text{-}26)$$

where $c(t_0)$ is determined from

$$\dot{c} = b'GR^{-1}G'b - \tilde{x}'Q\tilde{x} \qquad c(T) = 0 \qquad (4.2\text{-}28)$$

The optimal controller is as indicated in Fig. 4.2-5.

The section problems ask for alternative derivations of the optimal tracker using the Hamilton–Jacobi approach, and (for those familiar with the notion) the Minimum Principle.

In the above optimal tracking solution, it is assumed that the states x are measurable. If not, then as in the regulator problem they may be replaced by state estimates, under observability of F, H—see Chapter 7.

It is also assumed in the above optimal solution that the desired trajectory $\bar{y}$ is known precisely over the interval $[t_0, T]$. In practical applications, it may be that at any particular time $t_1 \in [t_0, T]$, the future desired trajectory is known only over $[t_1, t_1 + \Delta]$, where Δ is fixed. (Consider aircraft control with a terrain-following radar: Δ corresponds to the "look ahead" time.) As discussed in more detail in the next section, $b(t_1)$, needed for the optimal control at t_1, will often depend mainly on values of $\bar{y}(t)$ for values of t near t_1, so that knowledge of $\bar{y}(t)$ over $[t_1, t_1 + \Delta]$ may be adequate for computing $b(t_1)$. The boundary condition $b(T) = 0$ could be replaced by $b(t_1 + \Delta) = 0$, or alternatively, one could retain $b(T) = 0$ and set $\bar{x}(t) = \bar{x}(t_1 + \Delta)$ for $t \in [t_1 + \Delta, T]$.

The tracking results presented so far do not penalize the final state. Problem 4.2-1 seeks generalization to the case when there is a terminal cost $[x'(T) - \bar{x}'(T)]D[x(T) - \bar{x}(T)]$. Of course, we would anticipate (and correctly so) that the only likely change would be to the boundary conditions on the Riccati equation for $P(\cdot)$, namely from $P(T) = 0$ to $P(T) = D$, and the equations for $b(\cdot)$, $c(\cdot)$. A special case is *terminal state control* when we must exactly achieve a final state $\bar{x}(T)$; then, as can be shown rigorously, $P^{-1}(T) = 0$ is the appropriate boundary condition. Clearly for this case, at least in the vicinity of T, it is more reasonable to work with a Riccati equation for $P^{-1}(t)$. Such an equation is readily derived. Since differentiating $PP^{-1} = I$ yields $dP^{-1}/dt = -P^{-1} \, dP/dt P^{-1}$, then pre- and postmultiplication of the Riccati equation for P by P^{-1} yields

$$dP^{-1}/dt = FP^{-1} + P^{-1}F + P^{-1}QP^{-1} - GR^{-1}G'$$

Of course, in the case when $Q = 0$, then this collapses to a linear equation. Note however, that the value P is still needed to obtain the control law.

Model-following tracking problem. This problem arises when the command signal r is known a priori. It may be solved by direct application of the tracking problem results; the solution is left to the reader (see Problem 4.2-5). There is another approach to solving such a problem which is of interest.

Observe that an augmented system may be defined along the lines previously studied with

$$\hat{x} = \begin{bmatrix} x \\ z_1 \end{bmatrix}, \qquad \hat{F} = \begin{bmatrix} F & 0 \\ 0 & A_1 \end{bmatrix}, \qquad \hat{G} = \begin{bmatrix} G \\ 0 \end{bmatrix}, \qquad \hat{\Gamma} = \begin{bmatrix} 0 \\ B_1 \end{bmatrix} r$$

$$\hat{Q} = \begin{bmatrix} Q & -QLC' \\ -CL'Q & CL'QLC' \end{bmatrix} \qquad \text{(4.2-30)}$$

The optimization task is to minimize

$$V(\hat{x}(t_0), u(\cdot), t_0) = \int_{t_0}^{T} (u'Ru + \hat{x}'\hat{Q}\hat{x})\, dt \tag{4.2-31}$$

subject to

$$\dot{\hat{x}} = \hat{F}\hat{x} + \hat{G}u + \hat{\Gamma}r \tag{4.2-32}$$

where r is known in the interval $[t_0, T]$. Such a task is linear quadratic regulation in the presence of known plant disturbances r, and is of interest in its own right. Problem 4.2-6 asks for a formal solution to this optimization task.

There is a further type of model-following problem, for which we offer no clean linear-quadratic solution. It arises when the command inputs r are completely unknown. There is then no a priori knowledge of the desired trajectory $\bar{y}$. All one can do is design for a particular r, for example, a step, and later check that for those r which are near a step, some sort of adequate model-following still occurs.

Discrete-time tracking.   The discrete-time tracking problem is similar to the continuous-time problem. Suppose the plant is

$$x(t + 1) = F(t)x(t) + G(t)u(t) \tag{4.2-33}$$

and take as the performance index

$$V(x(t_0), u(\cdot), t_0) = \sum_{t=t_0+1}^{T} \{[x(t) - \bar{x}(t)]'Q(t)[x(t) - \bar{x}(t)]$$
$$+ u'(t - 1)R(t)u(t - 1)\} \tag{4.2-34}$$

[Here $\bar{x}(\cdot)$ is a reference trajectory, $R(t)$ and $Q(t)$ are positive and nonnegative definite symmetric]. Let $P(t, T)$, $S(t)$ be as for the regulator problem that is, where $\bar{x}(t) \equiv 0$. Then the optimal control is

$$u^*(t) = -[G'(t)S(t + 1)G(t) + R(t + 1)]^{-1}G'(t)[S(t + 1)F(t)x(t) + b(t + 1)] \tag{4.2-35}$$

where

$$b(t) = [F'(t) + K(t)G'(t)]b(t + 1) - Q(t)\bar{x}(t) \qquad b(T) = 0 \tag{4.2-36}$$

The optimal performance index is

$$V^*[x(t), t] = x'(t)P(t)x(t) + 2x'(t)b(t) + c(t) - \bar{x}'(t)Q(t)\bar{x}(t) \tag{4.2-37}$$

with

$$c(t) = c(t + 1) - b'(t + 1)G(t)[G'(t)S(t + 1)G(t)$$
$$+ R(t + 1)]^{-1}G'(t)b(t + 1) + \bar{x}(t)Q(t)\bar{x}(t) \tag{4.2-38}$$

The derivation can be achieved by modification of the argument used in Chapter 2 for the regulator problem.

Main points of the section.   The optimal servo system can be derived by using regulator theory. There results a two-degrees-of-freedom controller involving a standard optimal feedback regulator and a feed-forward controller. In the finite time case, the gains are time-varying. When the states of the plant and/or desired trajectory generating system are not directly measurable, these can be estimated, leading to dynamics in the feedback and feed-forward controllers respectively. There may be virtue in a realization of the two controllers as a single controller with inputs $\bar{y}$, y and output the optimal control u^*.

The optimal tracking controller design requires a standard feedback regulator design involving the backwards solution of a Riccati equation, and an external signal that results from the backwards solution of a linear differential equation.

One model-following problem can be reorganized as a servo problem for an augmented desired trajectory signal model. A second model-following task can be organized as a standard tracking problem.

Problem 4.2-1.   Show that the index (4.2-5), may be written in the form of (4.2-29), where u^* is given from (4.2-24) and (4.2-25) and V^* is given from (4.2-26). [*Hint:* Show that $(u - u^*)'R(u - u^*) = u'Ru + (x - \bar{x})'Q(x - \bar{x}) + d/dt[x'Px + 2x'b + c]$.]

Problem 4.2-2.   Derive the tracking problem results directly, using the Hamilton–Jacobi theory of Chapter 2 (i.e., without using the regulator theory results). [*Hint:* Try a solution to the Hamilton–Jacobi equation of the form $x'Px + 2x'b + c$.] For those who have studied the Minimum Principle, see Appendix C, a further problem is to derive the results that way [*Hint:* Set $p(t) = P(t)x(t) + b(t)$.]

Problem 4.2-3.   Consider the servo problem of the section. Derive a solution using the coordinate basis

$$\hat{x} = \begin{bmatrix} x - LC'z \\ z \end{bmatrix} \equiv \begin{bmatrix} x - \bar{x} \\ z \end{bmatrix}$$

instead of that of (4.2-6). This means that we expect a control law given directly as $u^* = K'(x - \bar{x}) + K_1'z$ for some K', K_1', which is attractive from the implementation viewpoint. Use the following definition in formulating replacements for (4.2-7):

$$M = FLC' + LC'A + \frac{d(LC')}{dt}$$

[*Hint:* Find a state-variable equation linking $\hat{x}$ and u, and express the integrand in the performance index (4.2-5) in the form $u'Ru + \hat{x}'\hat{Q}\hat{x}$ for some $\hat{Q}$. Then follow the derivation of the text, with obvious changes.]

Problem 4.2-4.   (Error form of the optimal tracker). It is sometimes convenient to describe the solution of the optimal tracker in terms of a feedback signal

depending on $x - \tilde{x}$ and a feed-forward signal depending on $\tilde{x}$. When $d\tilde{x}/dt$ is available, it turns out that this is possible. Prove that, for the usual optimal tracker problem with no terminal weighting,

$$u^* = K'(x - \tilde{x}) - R^{-1}G'\overline{b}$$

with

$$-\dot{\overline{b}} = (F + GK')'\overline{b} + P\left(F\tilde{x} - \frac{d\tilde{x}}{dt}\right)$$

$$\overline{b}(T) = 0$$

and

$$V^* = [x(t_0) - \tilde{x}(t_0)]'P(t_0)[x(t_0) - \tilde{x}(t_0)] + 2[x(t_0) - \tilde{x}(t_0)]'\overline{b}(t_0) + \overline{c}(t_0)$$

with

$$\dot{\overline{c}} = \overline{b}'GR^{-1}G'\overline{b} - 2\overline{b}'(F\tilde{x} - \dot{x}) \qquad \overline{c}(T) = 0$$

[*Hint:* Show that $\overline{b} = b + G'P\tilde{x}$, and use the equation for b. Proceed similarly for $\overline{c}$, or proceed from the servo solution as given in Problem 4.2-3 along the lines of the text derivation of the tracking results.]

Problem 4.2-5.   Describe how the tracking problem results will yield solutions to the model-following problems as stated in the section, for the case when the command signal r is known a priori and construction of a state estimator for the command signal generator is out of the question. Assume known the initial state of the model.

Problem 4.2-6.   (i) For the linear quadratic regulation task (4.2-31), (4.2-32), where known disturbances $r \in [t_0, T]$ are present, solve for the optimal control. Use either the Hamilton–Jacobi or Minimum Principle method and show that

$$u^* = -R^{-1}\hat{G}'(\hat{P}\hat{x} + \hat{b})$$

where

$$-\dot{\hat{P}} = \hat{P}\hat{F} + \hat{F}'\hat{P} - \hat{P}\hat{G}R^{-1}\hat{G}'\hat{P} + \hat{Q}$$

$$-\dot{\hat{b}} = [\hat{F} - \hat{G}R^{-1}\hat{G}'\hat{P}]'\hat{b} - \hat{P}\hat{\Gamma}r$$

Notice the similarity to the optimal tracking solution.

(ii) Using the definitions (4.2-30) for $\hat{F}$, $\hat{G}$, etc. which apply for the optimal model-following (tracking) problem, show that the optimal control for this case is

$$u^* = -R^{-1}G'(Px + P_{12}z_1 + b_1)$$

where P, P_{12}, b_1 come from the partitioning

$$\hat{P} = \begin{bmatrix} P & P_{12} \\ P_{12}' & P_{22} \end{bmatrix}, \qquad \hat{b} = \begin{bmatrix} b_1 \\ 0 \end{bmatrix}$$

$$-\dot{b}_1 = (F - GR^{-1}G'P)'b_1 - PB_1r$$

Remark: The techniques of this problem can be used to give optimal tracking results in the presence of known disturbances.

4.3 INFINITE-TIME RESULTS

In this section, we restrict attention to plants and performance indices with time-invariant parameters, and we extend the finite-time results of the previous section to the infinite-time case. This is done as earlier for the regulator, by letting the terminal time T become infinite. As one might expect, the state (or state estimate) feedback part of the controller becomes time-invariant.

Crucial issues to be considered include signal and performance index boundedness and steady state tracking error. We shall begin our treatment with the tracking problem.

Infinite-time tracking problem.   Let us return to the statement of the optimal tracking problem in the last section, but with the obvious changes. We are given

$$\dot{x} = Fx + Gu \qquad (4.3\text{-}1)$$

$$y = H'x \qquad (4.3\text{-}2)$$

with H of rank m. We are given an m-vector function $\bar{y}(t)$ for $t \in (t_0, \infty)$. We suppose that $[F, G]$ is stabilizable, and we seek the optimal control u^* minimizing

$$\lim_{T \to \infty} V(x(t_0), u(\cdot), T) = \lim_{T \to \infty} \int_{t_0}^{T} [u'Ru + (x - \bar{x})'Q(x - \bar{x})]\, dt \qquad (4.3\text{-}3)$$

Here

$$\bar{x} = L\bar{y} \qquad (4.3\text{-}4)$$

for some matrix L, usually given by

$$L = H(H'H)^{-1} \qquad (4.3\text{-}5)$$

and Q is usually of the form

$$Q = [I - LH']'Q_1[I - LH'] + HQ_2H' \qquad (4.3\text{-}6)$$

for nonnegative Q_1 and Q_2. We suppose that $[F, D]$ is detectable, where $DD' = Q$.

One way to proceed is to write down the finite time problem solution and then let $T \to \infty$. Obviously, by virtue of stabilizability $P(t, T) \to \bar{P}$. Importantly also, we

shall show that $b(t)$ for all finite t approaches a finite quantity $\bar{b}(t)$. From (4.2-24) and the fact that $K = -\bar{P}GR^{-1}$ in the limiting case, we obtain, as $T \to \infty$

$$\bar{b}(t) = -\int_t^\infty \exp\left[(F + GK')'(\tau - t)\right]Q\tilde{x}(\tau)\,d\tau \tag{4.3-7}$$

Note that differentiation of (4.3-7) yields

$$\dot{\bar{b}} = -(F + GK')'\bar{b} + Q\tilde{x} \tag{4.3-8}$$

Because $(F + GK')'$ has all eigenvalues with negative real parts, under the stabilizability and detectability assumptions, it is easy to establish that bounded $\tilde{x}(\cdot)$ implies bounded $\bar{b}(\cdot)$: to see this, suppose that

$$\|\exp(F + GK')t\| \le \alpha \exp(-\beta t) \qquad \alpha, \beta > 0$$

and

$$\|Q\tilde{x}(s)\| \le \gamma$$

Then

$$\|\bar{b}(t)\| \le \alpha\gamma \int_t^\infty \exp\left[-\beta(s - t)\right] ds = \frac{\alpha\gamma}{\beta}$$

Even though P, b remain well defined as $T \to \infty$, in general $c(t)$ does not. Recall that

$$\dot{c} = bGR^{-1}G'b - \tilde{x}'Q\tilde{x} \qquad c(T) = 0$$

When $T \to \infty$, $c(t)$ is defined by an integration over the interval (t, ∞) of a quantity which in general is not zero. This means that the limiting "optimal" performance index is infinite, and so the label of the control $u = K'x - R^{-1}G'\bar{b}$ as optimal is in part a misnomer. It should be noted that the infinite index is unavoidable; apart from some special cases, which are considered later, it is impossible to secure simultaneously $u \to 0$ and $x - \tilde{x} \to 0$ as $T \to \infty$ [which would be necessary for a finite index; see (4.3-3)].

A step-function tracking example.   By way of example, consider a single-input, single-output plant

$$\dot{x} = Fx + gu, \qquad y = h'x$$

with $\tilde{y}$ a unit step function. Suppose further that $Q_1 = 0$, $Q_2 = 1$, $R = \rho$, and the regulation feedback law is $u = k'x$ where $k' = -\rho^{-1}g'\bar{P}$. Now $Q = hh'$, $Q\tilde{x} = h\tilde{y} = h$, and so (4.3-8) yields

$$\bar{b}(t) = \left[(F + gk')'\right]^{-1}h$$

for all t. Thus the optimal control is

$$u = -\rho^{-1}g'\bar{P}x - \rho^{-1}h'(F + gk')^{-1}g$$

With this control, the plant becomes

$$\dot{x} = (F + gk')x - g\rho^{-1}h'(F + gk')^{-1}g$$

and the limiting state becomes

$$x(\infty) = (F + gk')^{-1}g\rho^{-1}h'(F + gk')^{-1}g$$

The limiting output is

$$y(\infty) = \rho^{-1}[h'(F + gk')^{-1}g]^2$$

and the steady state control is

$$u(\infty) = \rho^{-1}[k'(F + gk')^{-1}g - 1]h'(F + gk')^{-1}g$$

Of course, the "optimal" performance index is in general infinite.

Approximately optimal trackers. Returning to the general tracking task, various approximately optimal trackers can be constructed. The formula (4.3-7) for $\bar{b}(t)$ shows that $\bar{b}(t)$ *in effect* depends on $\bar{x}(s)$ only for $s \in [t, t + \Delta]$, where Δ is five times the dominant time constant associated with the eigenvalues of $F + GK'$. So one only really need look a finite time in the future, and use

$$\bar{b}(t) = -\int_t^{t+\Delta} \exp\left[(F + GK')'(\tau - t)\right]Q\bar{x}(\tau)\,d\tau \tag{4.3-9}$$

or define $\bar{b}(t)$ as the solution at time t of

$$\frac{db(\tau, t + \Delta)}{d\tau} = -(F + GK')'b(\tau, t + \Delta) + Q\bar{x}(\tau) \qquad b(t + \Delta, t + \Delta) = 0 \tag{4.3-10}$$

If $\bar{x}(\cdot)$ is slowly varying, then one can use the approximation

$$\bar{b}(t) = [(F + GK')']^{-1}\bar{x}(t) \tag{4.3-11}$$

obtained by setting $\dot{\bar{b}} = 0$ and solving the differential equation for b. Note that this becomes an exact solution of (4.3-7) or (4.3-8) if $\bar{x}(t)$ is constant.

Example with integrator as plant. To understand better the error involved, and from this understanding to deduce another approximation again, consider the simple scalar example

$$\dot{x} = u \qquad y = x$$

with

$$V(x(0), u(\cdot)) = \int_0^\infty [(y - \bar{y})^2 + \tau^2 u^2]\,dt$$

and

$$\bar{y} = 0 \qquad 0 \leqslant t < t_1$$

$$= a(t - t_1) \qquad t_1 \leqslant t \leqslant t_2$$

$$= a(t_2 - t_1) \qquad t_2 \leqslant t < \infty$$

Solving the tracking problem leads to $P = \tau$,

$$u = -\tau^{-1}x - \tau^{-2}\bar{b}$$

with

$$\dot{\bar{b}} = \tau^{-1}\bar{b} + \tilde{y} \qquad b(\infty) = 0$$

Setting $\dot{\bar{b}} = 0$ leads to a suboptimal control $u = -\tau^{-1}x + \tau^{-1}\tilde{y}$.

Integrating the differential equation for $\bar{b}(\cdot)$ leads to a more complex control. The resulting tracking performance is depicted in Figure 4.3-1. Notice the anticipating characteristic of the optimal response due to prior knowledge of $\tilde{y}$. The optimal and suboptimal feedforward controls are depicted in Figure 4.3-2. Again notice the anticipation in the optimal command signal. The case $x(0) = 0$ is depicted throughout.

These characteristics suggest a further approximation, viz. to set

$$\bar{b}(t) = -\tau\tilde{y}(t + \tau) \qquad (4.3\text{-}12)$$

and achieve a refined suboptimal control

$$u = -\tau^{-1}x + \tau^{-1}\tilde{y}(t + \tau) \qquad (4.3\text{-}13)$$

(The effect in Figures 4.3-1 and 4.3-2 would be of course to shift the suboptimal y and suboptimal feedforward commands left by τ units.) There is a good justification for this. The transfer function linking $\tilde{y}$ to $\bar{b}$ is $(j\omega - \tau^{-1})^{-1}$, which has a phase of $\varphi = \tan^{-1}(\omega\tau)$. The group delay is $-d\varphi/d\omega = -\tau(\omega^2\tau^2 + 1)^{-1}$ so that for low frequency signals, the mapping from $\tilde{y}$ to $\bar{b}$ advances the signals by $\tau(\omega^2\tau^2 + 1)^{-1} \simeq \tau$. The gain of the transfer function for small ω is $-\tau$. Hence we get, approximately, $\bar{b}(t) = -\tau\tilde{y}(t + \tau)$.
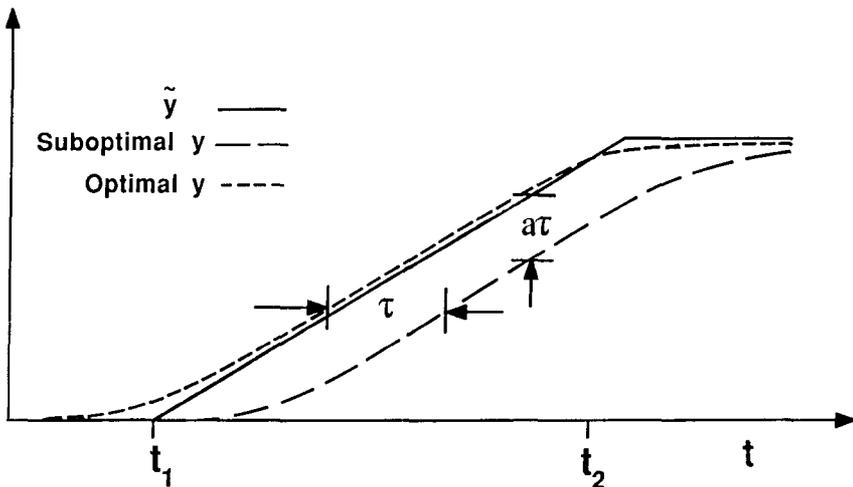
Figure 4.3-1   Tracking performance of two control laws.
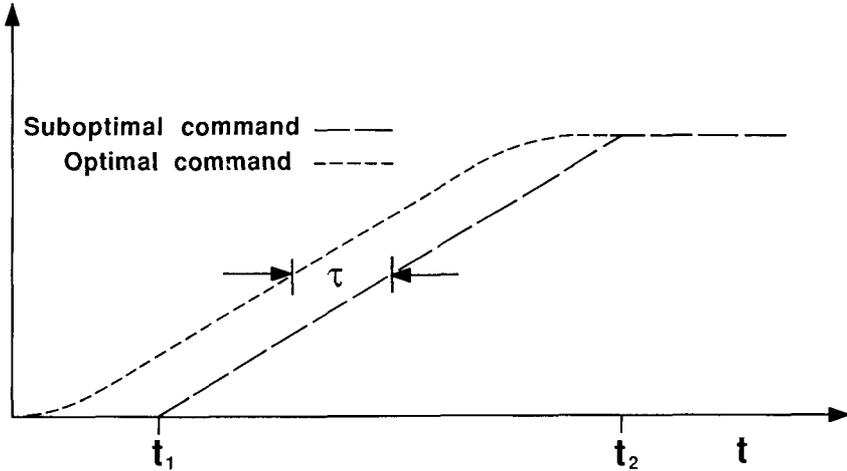
Figure 4.3-2   Optimal and suboptimal feedforward controls.

Equation (4.3-8) can be handled in the same way if $F + GK'$ has real eigenvalues and is diagonalizable, by changing the coordinate basis so that $(F + GK')$ is diagonal. Then the new (4.3-8) is a collection of uncoupled first-order equations of the type

$$\dot{\beta}_i = \lambda_i \beta_i + \mu_i(t)$$

and the approximate solution, valid for low frequency $\mu_i(t)$, is

$$\beta_i(t) = -\lambda_i^{-1} \mu_i(t + \lambda_i^{-1}) \tag{4.3-14}$$

When $(F + GK')$ has complex eigenvalues, such simplification is not always possible. We sum up these ideas as follows:

> *Solution to infinite-time tracking problem.*   With time-invariant F, G, H, Q, R, and L and with detectability and stabilizability assumptions holding, the optimal control for (4.3-1) and (4.3-3) becomes $u = K'x + \bar{b}$ where $K = -\bar{P}GR^{-1}$, $\bar{P}$ being the nonnegative definite and stabilizing solution of the steady-state Riccati equation, and $\bar{b}$ being given by (4.3-7). If $\bar{y}(\cdot)$ is bounded, so are $\bar{b}(\cdot)$, $u(\cdot)$, and $x(\cdot)$. Approximations to $\bar{b}$ are available—the finite horizon approximation (4.3-9) of (4.3-7), the evaluation (4.3-11), obtained by setting $\bar{b} = 0$ in the defining differential equation, and, in case $(F + GK')$ has all real eigenvalues and is diagonalizable, the gain-and-advance approximation exemplified by (4.3-12). The last two approximations rely on the reference trajectory's being slowly varying.

Infinite-time servo problem.   In the last section, we derived the tracking results from the servo results. Here, we shall use the reverse procedure. We

simply observe that a servo problem is a tracking problem where $\bar{y}$ is generated according to

$$\dot{z} = Az \qquad (4.3\text{-}15)$$

$$\bar{y} = C'z \qquad (4.3\text{-}16)$$

Let us distinguish these cases.

1. Re $\lambda_i(A) < 0$ for all i. Then $\bar{y}(t)$ and $\bar{x}(t)$ decay to zero exponentially fast. As a result, so does $\bar{b}(t)$, as defined by (4.3-7). (Problem 4.3-1 requests checking of this fact.) It follows that with the feed-forward part of the optimal control decaying to zero exponentially fast, so also does $x(t)$, and the optimal index is finite.

2. Re $\lambda_i(A) \leq 0$ with Re $\lambda_i(A) = 0$ for some i, and A has no repeated pure imaginary eigenvalues. It follows that $\bar{y}$ is bounded, but does not decay, in general, to zero. This is akin to the usual tracking situation.

3. Re $\lambda_i(A) > 0$ for some i or A has repeated pure imaginary eigenvalues. Obviously, $\bar{y}$ is unbounded. The quantity $\bar{b}(t)$ defined in (4.3-7) may not even exist. However, under the constraint

$$\text{Re } \lambda_i(F + GK') + \text{Re } \lambda_j(A) < 0 \qquad \forall_i \text{ and } j \qquad (4.3\text{-}17)$$

the integrand (for fixed t) in (4.3-7) is guaranteed exponentially decaying and $\bar{b}(t)$ exists, but will not be bounded. Accordingly $u(t)$ is unbounded. [Since the feedback control stabilizes the plant and the output of the plant is desired to follow an unstable trajectory, it is not surprising that the "optimal" (feed-forward part of) u is unbounded.]

Actually, there is a further possibility, to which we now turn.

Internal models and zero-error servo behavior.   We now examine the possibility of obtaining zero asymptotic error *and* finite value of the performance index using infinite-time servo results. Assuming that Q and L have the typical specialized structure, it is clear that to achieve our objective, we must have $y - \bar{y} \to 0$ and also $u \to 0$. If now $\bar{y}$ is derived as in a typical servo problem, it consists of a linear combination of exponentials. The ones of interest are those which do not decay to zero as $t \to \infty$. If $y - \bar{y} \to 0$, then y must equally contain these exponentials. And if y contains these exponentials while $u \to 0$, then they must be modes of the open-loop plant. This means that there should be a coordinate basis for z and x such that

$$A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} \qquad F = \begin{bmatrix} A_1 & F_{12} \\ 0 & F_{22} \end{bmatrix} \qquad (4.3\text{-}18)$$

where Re $\lambda_i(A_1) \geq 0$ for all i, Re $\lambda_i(A_2) < 0$. But also, in order to secure $y - \bar{y} \to 0$ when $u \to 0$, we must have the possibility of matching zero-input responses at the

output of the plant. This means that the coordinate basis must also ensure that for some $\alpha \neq 0$

$$C' = [C_1' \quad C_2'] \qquad H' = [\alpha C_1' \quad H_2'] \tag{4.3-19}$$

Then given $z_1(t_0)$, there exists α, $x_1(t_0)$ for which the zero-input responses match as

$$C_1' \exp[A_1(t - t_0)] z_1(t_0) = \alpha C_1' \exp[A_1(t - t_0)] x_1(t_0)$$

In summary, we have given an argument for the

> ***Internal Model Principle.***   Consider a servo problem in which Re $\lambda_i(A) \geq 0$. In order that there hold $y - \bar{y} \to 0$ and $u \to 0$ for all servo initial conditions, it is necessary that there be a coordinate basis for the servo and plant state so that (4.3-18) and (4.3-19) hold.

The reason for the name is obvious: the plant must internally model the non-asymptotically stable part of the servo model. More general statements of this principle can be made by allowing relaxation of the requirement that $u \to 0$ and allowing dynamic feedback controllers. Then the internal model must be in the plant-controller open loop. This motivates the study of proportional plus integral feedback for tracking constant inputs in Chapter 9 on frequency shaped designs. See also Problem 4.3-5.

We shall now show that under the assumptions in the above internal model principle, the servo problem with

$$V(x(t_0), u(\cdot)) = \int_{t_0}^{\infty} [u'Ru + (y - \bar{y})'Q_2(y - \bar{y})] \, dt \tag{4.3-20}$$

is readily reducible to a regulator problem, and as a consequence yields $u \to 0$, $y - \bar{y} \to 0$ and V^* finite. First, the two system equations

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} A_1 & F_{12} \\ 0 & F_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} G_1 \\ G_2 \end{bmatrix} u, \qquad \begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

can be combined as

$$\dot{w} = \begin{bmatrix} A_1 & F_{12} & 0 \\ 0 & F_{22} & 0 \\ 0 & 0 & A_2 \end{bmatrix} w + \begin{bmatrix} G_1 \\ G_2 \\ 0 \end{bmatrix} u \qquad w = \begin{bmatrix} x_1 - z_1 \\ x_2 \\ z_2 \end{bmatrix} \tag{4.3-21}$$

$$y - \bar{y} = [\alpha C_1' \quad H_2' \quad -C_2']w$$

and this has a stabilizability property if $[F, G]$ is stabilizable. Denote the new state variable in (4.3-21) by w. Observe that the performance index (4.3-20) is then quadratic in u and w. If detectability holds for the regulator problem obtained with $\bar{y} = 0$ in (4.3-20), it holds for the new situation involving state w. Hence, under optimality, $u \to 0$, $w \to 0$, and V^* is finite, as we claimed.

Example.   In the first example in this section, we considered a tracking problem for a scalar plant with $\bar{y}$ a unit step function. Let us return to that example, postulating now that the F matrix is singular. Thus the plant can reproduce at its output with zero excitation a signal y that would cancel $\bar{y}$, a unit step function.

Because the plant transfer function has a pole at the origin and because all signals are bounded, it ought to turn out that $u(\infty) = 0$. Let us check that this is so. Observe that

$$1 - k'(F + gk')^{-1} g = \det [1 - k'(F + gk')^{-1} g]$$

$$= \det [1 - (F + gk')^{-1} gk']$$

$$= \det [[F + gk']^{-1} [F + gk' - gk']]$$

$$= \frac{\det F}{\det [F + gk']}$$

Hence $1 - k'(F + gk')^{-1} g = 0$. The formula for the steady state control then shows that $u(\infty) = 0$. It should also turn out that $y(\infty) = 1$. Let us verify that this is so. The Riccati equation yields

$$\bar{P}F + F'\bar{P} - \bar{P}g\rho^{-1}g'\bar{P} + hh' = 0$$

or

$$\bar{P}(F + gk') + (F' + kg')\bar{P} + k\rho k' + hh' = 0$$

Premultiply by $\rho^{-1}g'(F' + kg')^{-1}$ and postmultiply by $(F + gk')^{-1}g$. There results, since $k' = -\rho^{-1}g'\bar{P}$,

$$-2k'(F + gk')^{-1}g + [k'(F + gk')^{-1}g]^2 + \rho^{-1}[h'(F + gk')^{-1}g]^2 = 0$$

Recalling that $k'(F + gk')^{-1}g = 1$; there follows

$$y(\infty) = \rho^{-1}[h'(F + gk')^{-1}g]^2 = 1$$

Of course, we could have set up this problem ab initio as a servo problem, and converted it to a regulator problem, and obtained these results in a far more transparent way.

Model-following: Step commands.

Sometimes for a two-degrees-of-freedom time-invariant controller design, a natural framework to adopt is that of an infinite-time model-following task where the plant is forced to track the step response to a specified model; that is, the external reference input r is a unit step. It may be, for instance, that a flexible wing aircraft (the plant) is required to perform (after connection of feed-forward and feedback control) in the same manner as a rigid body aircraft (the reference model). Then we build on the finite-time model-following results of the previous section to achieve time-invariant controller designs.

In the notation of the finite-time, model-following problem, let us work with an augmented plant consisting of the original plant with input u, state x, output y and the model with input r, state z_1, and output $\bar{y}$. Thus consider

$$\dot{\hat{x}} = \hat{F}\hat{x} + \hat{G}u + \hat{\Gamma}r \tag{4.3-22}$$

$$y = \hat{H}'_1\hat{x} \qquad \bar{y} = \hat{H}'_2\hat{x} \qquad (y - \bar{y}) = (\hat{H}_1 - \hat{H}_2)'\hat{x}$$

where

$$\hat{x} = \begin{bmatrix} x \\ z_1 \end{bmatrix} \qquad \hat{F} = \begin{bmatrix} F & 0 \\ 0 & A_1 \end{bmatrix} \qquad \hat{G} = \begin{bmatrix} G \\ 0 \end{bmatrix} \qquad \hat{\Gamma} = \begin{bmatrix} 0 \\ B_1 \end{bmatrix}$$

$$\hat{H}'_1 = [H' \quad 0] \qquad \hat{H}'_2 = [0 \quad C'_1] \tag{4.3-23}$$

The plant control u is to be selected so that the original plant output y tracks the model plant output $\bar{y}$ in the presence of the known step function external input r. The model is assumed to be asymptotically stable.

The quadratic index should penalize a term $u'Ru$ and a term $(y - \bar{y})'Q(y - \bar{y})$, or equivalently, $\hat{x}'(\hat{H}_1 - \hat{H}_2)Q(\hat{H}_1 - \hat{H}_2)'\hat{x}$ for some $R > 0$, $Q > 0$. Thus with $\hat{Q} = (\hat{H}_1 - \hat{H}_2)Q(\hat{H}_1 - \hat{H}_2)'$, consider the optimization of

$$V = \int_{t_0}^{\infty} [u'Ru + \hat{x}'\hat{Q}\hat{x}]\, dt \tag{4.3-24}$$

over $u(\cdot)$ subject to

$$\dot{\hat{x}} = \hat{F}\hat{x} + \hat{G}u + \hat{\Gamma}r \tag{4.3-22}$$

with r a known constant.

We take the limiting solution of the finite-time results explored in Problem 4.2-6 as

$$u^* = -R^{-1}\hat{G}'(\hat{P}\hat{x} + \hat{b})$$

where

$$\hat{P}\hat{F} + \hat{F}'\hat{P} - \hat{P}\hat{G}R^{-1}\hat{G}'\hat{P} + \hat{Q} = 0 \qquad \hat{P} \geq 0$$

and $\hat{b}$ is the limiting solution of

$$-\dot{\hat{b}} = [\hat{F} - \hat{G}R^{-1}\hat{G}'\hat{P}]'\hat{b} - \hat{P}\hat{\Gamma}r \qquad \hat{b}(\infty) = 0$$

namely, since r is constant,

$$\hat{b} = [(\hat{F} - \hat{G}R^{-1}\hat{G}'\hat{P})']^{-1}\hat{P}\hat{\Gamma}r$$

Now with $\hat{P}$, $\hat{b}$ partitioned as

$$\hat{P} = \begin{bmatrix} P & P_{12} \\ P'_{12} & P_{22} \end{bmatrix}, \qquad \hat{b} = \begin{bmatrix} b_1 \\ * \end{bmatrix}$$

further manipulations show that P satisfies the usual steady state Riccati equation

for the original plant, and P_{12} is the solution of a linear matrix equation; the stability of A and of $F - GR^{-1}G'P$ guarantees existence of P_{12}. Also,

$$u^* = -R^{-1}G'(Px + P_{12}z_1 + b_1)$$

with

$$b_1 = (F' - PGR^{-1}G')^{-1}PB_1r$$

Thus the optimal control law has the simple form

$$u^* = K'x + K'_{12}z_1 + K'_1r = K'x + u_f$$

$$K' = -R^{-1}G'P, \qquad K'_{12} = -R^{-1}G'P_{12} \tag{4.3-25}$$

$$K'_1 = -R^{-1}G'(F' - PGR^{-1}G')^{-1}PB_1$$

The feed-forward controller to the original plant generating u_f, driven by r, and incorporating the model is

$$\dot{z}_1 = A_1z_1 + B_1r$$
$$u_f = K'_{12}z_1 + K'_1r \tag{4.3-26}$$

We stress that such a feed-forward controller does not include anticipatory characteristics for variable reference signals r, being optimal only for $r = $ constant. Any adjustment to the feed-forward controller to give improved transient response will usually involve some lead-lag network to replace the constant gain K_1. Such design adjustments involve classical rather than optimal techniques, and will not be discussed further here.

It is known from classical design theory that improved step performance response can be achieved by using proportional plus integral feedback, rather than just proportional feedback as here. Then zero steady state error can be achieved by
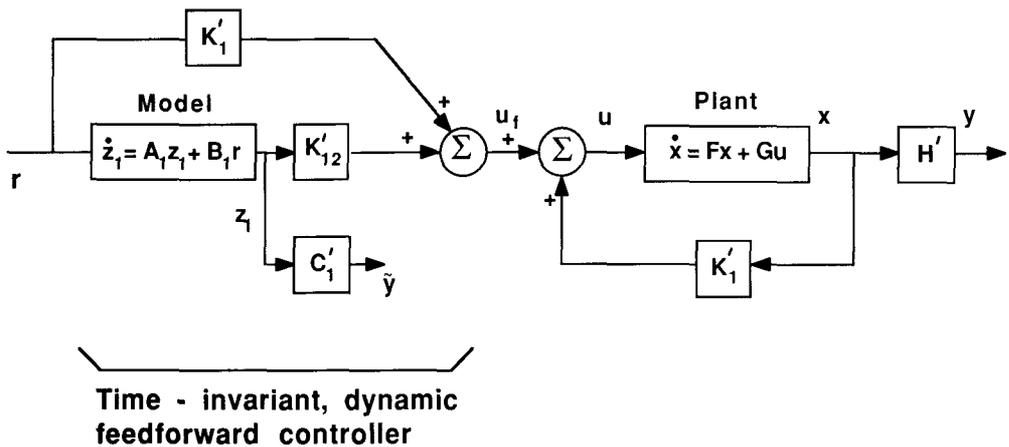
Figure 4.3-3 Optimal model-following.

virtue of the internal model (integrator) in the feedback loop. This aspect is developed in Chapter 9 on frequency shaping. See also one approach in Problem 4.3-5.

Main points of the section.   With time-invariant parameters, and assumption of detectability and stabilizability, the infinite-time optimal tracking problem (regarded as the limit of a finite-time problem) has a solution that yields bounded $u(\cdot)$ and $y(\cdot)$ when $\bar{y}(\cdot)$ is bounded, but generally yields an infinite performance index. Several approximations to the feed-forward control are available, all of which involve limiting the look-ahead requirement. Servo problems are conveniently divided into those with asymptotically stable, neutrally stable, and unstable servo model: the latter may have no solution. When the plant itse!f contains a copy of the nonasymptotically stable part of the servo model, asymptotically zero tracking error is secured, with a finite optimal performance index.

Problem 4.3-1.   Suppose that $\tilde{x}(t)$ decays to zero exponentially fast and $F + GK'$ has all eigenvalues with negative real parts. Let

$$\bar{b}(t) = -\int_t^\infty \exp\left[(F + GK')'(\tau - t)\right]Q\tilde{x}(\tau)\,d\tau$$

Show that $\|\bar{b}(t)\|$ decays to zero exponentially fast.

Problem 4.3-2.   Suppose we are given a time-invariant completely stabilizable and detectable plant $\dot{x} = Fx + Gu$, $y = H'x$ and a servo problem in which $\dot{z} = Az$, $\bar{y} = C'z$ and Re $\lambda_i(A) < 0$ for all i. Suppose that the performance index is $\int_{t_0}^\infty [u'Ru + (y - \bar{y})'Q(y - \bar{y})]\,dt$ with R, Q positive definite to achieve model-following with the plant output y tracking the model output $\bar{y}$. Show that this model-following (servo) problem can be solved as a regulator problem.

Problem 4.3-3.   Consider a servo problem in which Re $\lambda_i(A) < \alpha$ for all i and some positive α, while Re $\lambda_i(A) > 0$ for some i. Show how a performance index can be chosen which will ensure that the optimal control is finite at all times.

Problem 4.3-4.   In this section, a number of optimization problems have been identified where the optimal index becomes infinite. Suppose the performance index is varied to

$$\lim_{T \to \infty} \left[\frac{1}{T}\int_{t_0}^T [u'Ru + (x - \tilde{x})'Q(x - \tilde{x})]\,dt\right]$$

For the servo problem with $\bar{y}$ a step function, show that this index remains bounded.

Problem 4.3-5.   Suppose a plant does not have a pole at the origin, and it is desired that the plant output should follow a unit step $\bar{y}$ with zero steady state error. The classical approach is to use a controller incorporating an integrator. One approach to achieving this is to first augment the plant at its input with an integra-

tor. Consider the following servo/tracking problem, which generalizes this approach to the multivariable case. The system is

$$\begin{bmatrix} \dot{x} \\ \dot{u} \end{bmatrix} = \begin{bmatrix} F & G \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} + \begin{bmatrix} 0 \\ I \end{bmatrix} \dot{u} \qquad y = H'x$$

with $\{F, G, H\}$ controllable and observable. (Here, $\dot{u}$ is a new external input, driving an integrator connected in each line at the input of the plant.) The index is

$$V(x(t_0), u(t_0), \dot{u}(\cdot)) = \lim_{T \to \infty} \int_{t_0}^{T} [\dot{u}'R\dot{u} + (y - \bar{y})'Q(y - \bar{y})] \, dt$$

with Q positive definite and $\bar{y}$ is a constant m-vector. Assume that $H'F^{-1}G$ is $m \times m$ square and nonsingular. Show that a regulator problem is obtained which leads to $y - \bar{y} \to 0$. Show that the control for the plant is defined by an equation of the form $\dot{u} = L_1 u + L_2 x + L_3 \bar{y}$. [Hint: Let $W = (H'F^{-1}G)^{-1}$ and examine the state vector

$$w = \begin{bmatrix} x - F^{-1}GW\bar{y} \\ u + W\bar{y} \end{bmatrix}$$

4.4   A PRACTICAL DESIGN EXAMPLE

Before concluding Part I, on the basic theory of the optimal regulator, and before moving on to the next part, which explores properties of the optimal controllers, we will study an actual engineering design using much of the theory so far.

An underlying assumption of Part I is that the states of the plants are measurable in a noise-free environment. Of course, if a satisfactory engineering design in terms of performance and robustness cannot be achieved with a full-state design, then there is no point in proceeding to incorporate state estimation, or working with any output feedback controller employing less than full-state information. Clearly, the actuators are inadequate and must be upgraded or the plant modified to be "more controllable."

The application we study here is drawn from [5] and [6]. The example illustrates the design of a controller for the lateral motion of a B-26 aircraft. The main idea is to provide control to cause the actual aircraft to perform similarly to a model; the way this qualitative notion is translated into quantitative terms will soon become clear. A servo approach will be used.

The general equations governing the aircraft motion are the linear set

$$\begin{bmatrix} \dot{\phi} \\ \ddot{\phi} \\ \dot{\beta} \\ \dot{r} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & L_p & L_\beta & L_r \\ \dfrac{g}{V} & 0 & Y_\beta & -1 \\ \dfrac{N_{\dot{\beta}}g}{V} & N_p & N_\beta + N_{\dot{\beta}}Y_\beta & N_r - N_{\dot{\beta}} \end{bmatrix} \begin{bmatrix} \phi \\ \dot{\phi} \\ \beta \\ r \end{bmatrix}$$

$$+ \begin{bmatrix} 0 & 0 \\ L_{\delta_r} & L_{\delta_a} \\ Y_{\delta_r} & 0 \\ N_{\delta_r} + N_{\dot\beta} Y_{\delta_r} & N_{\delta_a} \end{bmatrix} \begin{bmatrix} \delta_r \\ \delta_a \end{bmatrix}$$

In these equations, ϕ denotes the bank angle, β the sideslip angle, r the yaw rate, δ_r, the rudder deflection, and δ_a the aileron deflection. Of course, we identify x with $[\phi \quad \dot\phi \quad \beta \quad r]'$, and so forth. The quantities L_p, and so forth, are fixed parameters associated with the aircraft. For the B-26, numerical values for the F and G matrices become

$$F = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & -2.93 & -4.75 & -0.78 \\ 0.086 & 0 & -0.11 & -1.0 \\ 0 & -0.042 & 2.59 & -0.39 \end{bmatrix}$$

$$G = \begin{bmatrix} 0 & 1 \\ 0 & -3.91 \\ 0.035 & 0 \\ -2.53 & 31 \end{bmatrix}$$

However, the dynamics represented by this arrangement are unsatisfactory. In particular, the zero-input responses are preferred to be like those of the model

$$\dot z = Az \qquad A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & -1 & -73.14 & 3.18 \\ 0.086 & 0 & -0.11 & -1 \\ 0.0086 & 0.086 & 8.95 & -0.49 \end{bmatrix}$$

This model is derived by varying those parameters in the F matrix corresponding to aircraft parameters which could be physically varied. For this reason, the first and third rows of A are the same as the first and third row of F. The eigenvalues of A are

$$-1.065, \qquad +0.00275, \qquad -0.288 \pm j2.94$$

Although one is unstable, the associated time constant is so large that the effect of this unstable mode can be cancelled by appropriate use of an external nonzero input.

To achieve actual performance resembling that of the model, we pose the quantitative problem of minimizing

$$\int_0^\infty [u'u + (x - z)'Q(x - z)]\, dt$$

Thus, we are considering a servo problem that differs from the standard one only by the specializations $H = C = L = I$.

As we know, the optimal control becomes of the form

$$u = K'x + K_1'z$$

where K and K_1 are determined by the methods given previously. In practice, the control

$$u = K'x + K_1'z + u_{ext}$$

would be used, where u_{ext} denotes an externally imposed control. Likewise, in practice, the model equation $\dot{z} = Az$ would be replaced by

$$\dot{z} = Az + Gu_{ext}$$

At this stage, we are faced with the problem of selecting the matrix Q in the performance index. From inspection of the performance index, it is immediately clear that the larger Q is, the better will be the following of the model by the plant. This is also suggested by the fact that large Q leads to some poles of the closed-loop system

$$\dot{x} = (F + GK')x$$

being well in the left half-plane—that is, leads to the plant with feedback around it tending to respond fast to that component of the input, $K_1'z$, which arises from the model.

On the other hand, we recall that the larger Q is taken, the larger are likely to be entries of K, and, for the aircraft considered, it is necessary to restrict the *magnitude* of the entries of K to be less than those of

$$K_{max} = \begin{bmatrix} 5 & 5 \\ 0.5 & 2 \\ 5 & 20 \\ 5 & 1 \end{bmatrix}$$

To begin with, a Q matrix of the form pI can be tried. Either trial and error, or an approximate expression for the characteristic polynomial of $F + GK'$ obtained in [6], suggests that $Q = 5I$ is appropriate. This leads to eigenvalues of $F + GK'$ which have the values

$$-0.99, \quad -1.3, \quad -5.13, \quad -9.14$$

Larger Q leads to more negative values for the last two eigenvalues, whereas the first two do not vary a great deal as Q is increased. Comparison with the model eigenvalues suggests the possibility of further improvement, in view of the fact that the ratio of the nondominant eigenvalue of $F + GK'$ nearest to the origin to the eigenvalue of the model most remote from the origin is about 5. On the other hand, the gain matrix K associated with $Q = 5I$ is

$$\begin{bmatrix} -2.53 & -2.21 \\ -0.185 & -1.83 \\ 1.58 & 0.7 \\ -2.34 & -0.01 \end{bmatrix}$$

and at least one of the entries (the 2–2 one) is near its maximum permissible value in magnitude.

This suggests that some of the diagonal entries of Q should be varied. To discover which, one can use two techniques

1. One can plot root loci for the eigenvalues of $F + GK'$, obtained by varying one q_{ii}. Variations causing most movement of the eigenvalues leftward and simultaneously retaining the constraints on K can be determined.
2. One can examine the error between the model state $z(t)$ and the plant state $x(t)$ obtained for several initial conditions, $z(0) = x(0) = [1 \quad 0 \quad 0 \quad 0]'$, $[0 \quad 1 \quad 0 \quad 0]'$, $[0 \quad 0 \quad 1 \quad 0]'$ and $[0 \quad 0 \quad 0 \quad 1]'$, for example, using the design resulting from $Q = 5I$. One can then adjust those diagonal entries of Q which weight those components of $(z - x)$ most in error.

Case 2 leads to the greatest errors being observed in $(z_2 - x_2)$ and $(z_4 - x_4)$. This suggests adjustment of q_{22} and/or q_{44}, and this is confirmed from Case 1. However, adjustment of q_{22} causes the 2–2 entry of K to exceed its maximum value. On the other hand, adjustment of q_{44}, from 5 to 20, proves satisfactory. The new eigenvalues of $F + GK'$ become

$$-0.908, \qquad -0.66, \qquad -9.09, \qquad -11.2$$

and the gain matrix K is

$$\begin{bmatrix} -0.201 & -2.23 \\ -0.185 & -1.83 \\ 1.42 & 0.164 \\ -4.42 & -0.264 \end{bmatrix}$$

For completeness, we state the feed-forward gain matrix K_1 associated with the model states. For $Q = \text{diag}\,[5, \quad 5, \quad 5, \quad 20]$, this is

$$K_1 = \begin{bmatrix} 0.101 & 2.045 \\ 0.344 & 2.172 \\ -2.153 & -1.54 \\ 5.61 & 2.42 \end{bmatrix}$$

Although the model is unstable, the sum of any eigenvalue òf the model matrix A and the matrix $F + GK'$ is negative, which as we know guarantees the existence of K_1.

Main points of the section.   At this stage in the text, we have powerful tools available for practical controller design. These must not be used naively, but with insights from classical control design. The rather ad hoc method of performance index selection and the necessity to have all states measurable, motivates the next parts of the text, which deal with state estimation, and systematic design methods using the tools of Part I.

Problem 4.4-1.   (Requires computer solution). Confirm the results of this section, using standard software packages.

REFERENCES

[1] M. Athans and P. L. Falb, *Optimal Control*. New York: McGraw-Hill, 1966.

[2] E. Kreindler, "On the Linear Optimal Servo Problem," *Intern. J. Control*, Vol. 9, No. 4 (1969), pp. 465–472.

[3] R. E. Kalman, "The Theory of Optimal Control and the Calculus of Variations," *Mathematical Optimization Techniques*, R. Bellman, ed. Los Angeles, University of California Press, 1963, Chapter 16.

[4] F. L. Lewis, *Optimal Control*. New York: John Wiley and Sons, 1986.

[5] J. S. Tyler, "The Characteristics of Model Following Systems as Synthesized by Optimal Control," *IEEE Trans. Auto. Control*, Vol. AC-9, No. 4 (October 1964), pp. 485–498.

[6] J. S. Tyler and F. B. Tuteur, "The Use of a Quadratic Performance Index to Design Multivariable Control Systems," *IEEE Trans. Auto. Control*, Vol. AC-11, No. 1 (January 1966), pp. 84–92.

5

Properties of Regulator Systems with a Classical Control Interpretation

5.1 THE REGULATOR FROM AN ENGINEERING VIEWPOINT

We have earlier intimated a desire to point out what might be termed the "engineering significance" of the regulator. Until now, we have exposed a mathematical theory for obtaining feedback laws for linear systems. These feedback laws minimize performance indices that reflect the costs of control and of having a nonzero state. In this sense, they may have engineering significance. Furthermore, we have indicated in some detail for time-invariant systems a technique whereby the closed-loop system will be asymptotically stable, and will even possess a prescribed degree of stability. This, too, has obvious engineering significance.

Again, there is engineering significance in the fact that, in distinction to most classical design procedures, the techniques are applicable to multiple-input systems, and to time-varying systems. (We have tended to avoid discussion of the latter because of the additional complexity required in, for example, assumptions guaranteeing stability of the closed-loop system. However, virtually all the results presented hitherto and those to follow are applicable in some way to this class of system.)

But there still remains a number of unanswered questions concerning the engineering significance of the results. For example, we might well wonder to what extent it is reasonable to think in terms of state feedback when the states of a system are not directly measurable. All the preceding, and most of the following, theory is

101

built upon the assumption that the system states are available; quite clearly, if this theory is to be justified, we shall have to indicate some technique for dealing with a situation where no direct measurement is possible. We shall discuss such techniques in a subsequent chapter. Meanwhile, we shall continue with the assumption that the system states are available.

In classical control, the notions of gain margin and phase margin play an important role giving quantitative measures of robustness to uncertainties or changes at the plant input. Thus, engineering system specifications will often place lower bounds on these quantities, since it has been found, essentially empirically, that if these quantities are too small, actual system performance, as distinct from nominal system performance, will be degraded in some way. For example, if for a system with a small amount of time delay a controller is designed neglecting the time delay, and if the phase margin of the closed loop is small, there may well be oscillations in the actual closed loop. The natural question now arises as to what may be said about the gain margin and phase margin (if these quantities can, in fact, be defined) of an optimal regulator.

Of course, at first glance, there can be no parallel between the dynamic feedback of the output of a system, as occurs in classical control, and the memory-less feedback of states, as in the optimal regulator. But both schemes have associated with them a closed loop. Figure 5.1-1 shows the classical feedback arrangement for a system with transfer function $h'(sI - F)^{-1}g$, where the output is fed back through a dynamic controller with transfer function $\beta(s)$. Figure 5.1-2 shows a
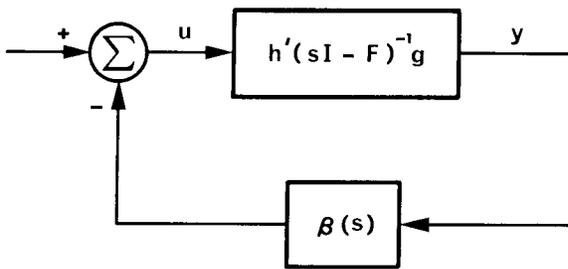
Figure 5.1-1   Classical feedback arrangement with dynamic controller driven by system.
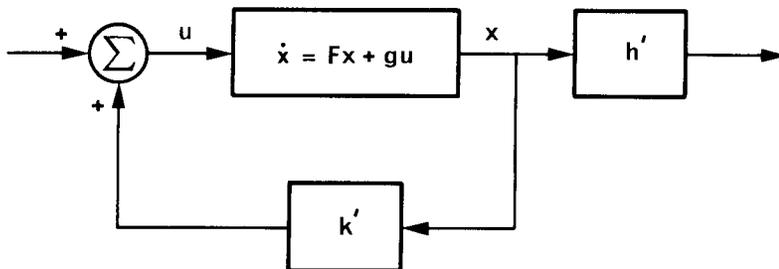
Figure 5.1-2   Modern feedback arrangement with memoryless controller driven by system states.

system with transfer function $h'(sI - F)^{-1}g$ but with memoryless state-variable feedback. Here a closed loop is formed; however, it does not include the output of the open-loop system, merely the states. This closed loop is shown in Fig. 5.1-3.

Now it is clear how to give interpretations of the classical variety to the optimal feedback system. The optimal feedback system is like a classical situation where unity negative feedback is applied around a (single-input, single-output) system with transfer function $-k'(sI - F)^{-1}g$. Thus, the gain margin of the optimal regulator may be determined from a Nyquist, or some other, plot of $W(j\omega) = -k'(j\omega I - F)^{-1}g$ in the usual manner.

We may recall the attention given in classical control design procedures to the question of obtaining satisfactory transient response. Thus, to obtain for a second-order system a fast response to a step input, without excessive overshoot, it is suggested that the poles of the closed-loop system should have a damping ratio of about 0.7. For a higher-order system, the same sort of response can be achieved if two dominant poles of 0.7 damping ratio are used. We shall discuss how such effects can also be achieved by using an optimal regulator; the key idea revolves around appropriate selection of the weighting matrices (Q and R) appearing in the performance index definition.

A common design procedure for systems containing a nonlinearity is to replace the nonlinearity by an equivalent linear element, and to design and analyze with this replacement. One then needs to know to what extent the true system performance will vary from the approximating system performance. As will be seen, a number of results involving the regulator can be obtained, giving comparative information of the sort wanted.

We return to the question of control loop robustness; it is important to discover how well an optimal regulator will perform with variations in the parameters of the forward part of the closed-loop system. One of the common aims of classical control (and particularly that specialization of classical control, feedback amplifier design) is to insert feedback so that the input-output performance of the closed-loop system becomes less sensitive to variations in the forward part of the system. In other words, one seeks to desensitize the performance to certain parameter variations.

The quantitative discussion of many of the ideas just touched upon depends on
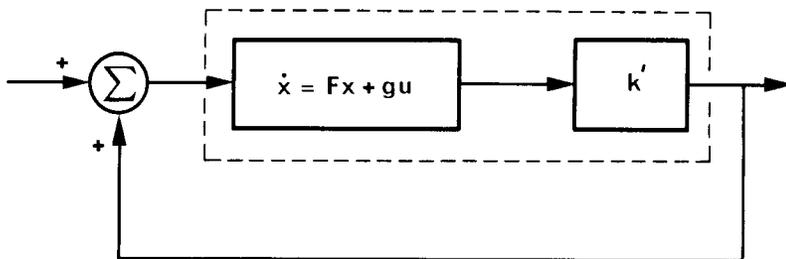
Figure 5.1-3   The closed-loop part of a feedback system using modern feedback arrangement.

the application of one of several basic formulas, which are derived in the next section. Then we pass on to the real meat of the regulator ideas in this and the next chapter.

5.2 RETURN DIFFERENCE EQUALITY AND RELATED FORMULAS

To fix ideas for the remainder of this chapter, we shall restrict attention to closed-loop systems that are completely stabilizable, time-invariant, and asymptotically stable. Thus, we shall take as our fundamental open-loop system

$$\dot{x} = Fx + Gu \tag{5.2-1}$$

with $[F, G]$ completely stabilizable. As the performance index, we take

$$V(x(t_0), u(\cdot), t_0) = \int_{t_0}^{\infty} (u'Ru + x'Qx)\, dt \tag{5.2-2}$$

with the usual constraints on Q and R, including that $[F, D]$ be completely detectable for any D such that $DD' = Q$. Let K be the optimal control law. Then we have:

Return difference equality and consequences.   The following identity holds:

$$R + G'(-j\omega I - F')^{-1}Q(j\omega I - F)^{-1}G$$
$$= [I - G'(-j\omega I - F')^{-1}K]\,R\,[I - K'(j\omega I - F)^{-1}G] \tag{5.2-3}$$

and the following are consequences of this identity:

$$[I + G'(-j\omega I - F' - KG')^{-1}K]\,R\,[I + K'(j\omega I - F - GK')^{-1}G]$$
$$= R - G'(-j\omega I - F' - KG')^{-1}Q(j\omega I - F - GK')^{-1}G \tag{5.2-4}$$

$$[I - G'(-j\omega I - F')^{-1}K]\,R\,[I - K'(j\omega I - F)^{-1}G] \geq R \tag{5.2-5}$$

$$[I + G'(-j\omega I - F' - KG')^{-1}K]\,R\,[I + K'(j\omega I - F - GK')^{-1}G] \leq R \tag{5.2-6}$$

The name "return difference equality" is given to (5.2-3), since the quantity $I - K'(j\omega I - F)^{-1}G$, at least in the scalar case, has long been known as the return difference, when the plant and controller combination is organized as shown in Fig. 5.1-3.

We shall prove (5.2-3). The steady state Riccati equation yields

$$\bar{P}F + F'\bar{P} - \bar{P}GR^{-1}G'\bar{P} + Q = 0$$

or

$$\bar{P}(j\omega I - F) + (-j\omega I - F')\bar{P} + KRK' = Q$$

Multiply on the left by $G'(-j\omega I - F')^{-1}$ and on the right by $(j\omega I - F)^{-1}G$. Use the fact that $\bar{P}G = -KR$. There results

$$G'(-j\omega I - F')^{-1}KR + RK'(j\omega I - F)^{-1}G + G'(-j\omega I - F')^{-1}KRK'(j\omega I - F)^{-1}G$$
$$= G'(-j\omega I - F')^{-1}Q(j\omega I - F)^{-1}G$$

from which (5.2-3) follows easily.

The proof of (5.2-4) is dealt with in Problem 5.2-1. The inequalities (5.2-5) and (5.2-6) are simply consequences of (5.2-3) and (5.2-4), once it is recognized that $A^*QA \geq 0$ when $Q \geq 0$ and A is arbitrary.

Equation (5.2-3) is also a form of spectral factorization. Spectral factorization is concerned with the following problem. Given a matrix $\Phi(j\omega)$ which is positive definite hermitian on the $j\omega$-axis, find a transfer function matrix $W(j\omega)$ with

$$\Phi(j\omega) = W'(-j\omega)W(j\omega) \tag{5.2-7}$$

for all ω. Often, $W(j\omega)$ and/or $W^{-1}(j\omega)$ may be restricted to being stable. In fact, the factorization (5.2-7) is unique to within left multiplication of $W(j\omega)$ by an orthogonal matrix if both $W(j\omega)$ and $W^{-1}(j\omega)$ are stable. In our case, we have

$$\Phi(j\omega) = R + G'(-j\omega I - F')^{-1}Q(j\omega I - F)^{-1}G \tag{5.2-8}$$

which is certainly positive definite hermitian. Indeed, $\Phi(j\omega) \geq R > 0$. Also, we have

$$W(j\omega) = R^{1/2}[I - K'(j\omega I - F)^{-1}G] \tag{5.2-9}$$

This is not necessarily stable. However,

$$W^{-1}(j\omega) = [I + K'(j\omega I - F - GK')^{-1}G]R^{-1/2} \tag{5.2-10}$$

as may be checked—see Appendix B or Problem 5.2-1—and thus is stable.

In the single-input case, $K'(j\omega I - F)^{-1}G$ becomes a scalar transfer function. Then (5.2-3) becomes (using lowercase letters to emphasize that certain quantities are scalar or vector rather than, as normal, matrices):

$$r + g'(-j\omega I - F')^{-1}Q(j\omega I - F)^{-1}g = r|1 - k'(j\omega I - F)^{-1}g|^2 \tag{5.2-11}$$

and in case $Q = hh'$, there holds

$$r + |h'(j\omega I - F)^{-1}g|^2 = r|1 - k'(j\omega I - F)^{-1}g|^2 \tag{5.2-12}$$

The inequality (5.2-5) is simply

$$|1 - k'(j\omega I - F)^{-1}g|^2 \geq 1 \tag{5.2-13}$$

The return difference is thus lower bounded by 1 for all ω. In case $R = I$, the inequality (5.2-5) tells us that in the vector input case, the *singular values* of the return difference are lower-bounded by 1. [Recall (see Appendix A) that the singular values of a matrix A are the quantities $[\lambda_i(A^*A)]^{1/2}$. When $A^*A \geq I$, $\lambda_i(A^*A) \geq 1$ for all i].

Positive real and bounded real transfer function matrices.   Positive real and bounded real transfer function matrices are important in network theory [1], and stability theory [2]. A real rational matrix $Z(s)$ is positive real if all poles lie in Re $[s] \leq 0$ and

$$Z(s) + Z'(s^*) \geq 0$$

for all s in Re $[s] > 0$. A real rational $S(s)$ is bounded real if all poles lie in Re $[s] < 0$ and

$$I - S'(-j\omega)S(j\omega) \geq 0$$

From (5.2-6), it is virtually immediate that

$$S(j\omega) = R^{1/2}[I + K'(j\omega I - F - GK')^{-1}G]R^{-1/2} \tag{5.2-14}$$

is bounded real. It turns out that the following matrix is positive real:

$$Z(s) = -RK'(sI - F - \tfrac{1}{2}GK')^{-1}G \tag{5.2-15}$$

(see Problem 5.2-2). Actually, it is a standard network theoretic result that if S is bounded real, $(I - S)(I + S)^{-1}$ is positive real. The $Z(s)$ of (5.2-15) is in fact given as $2(I - S)(I + S)^{-1}$. These properties will be exploited subsequently.

The closed-loop eigenvalues.   The return difference equality allows a characterization of the system closed-loop poles, that is, the eigenvalues of $F + GK'$, in terms of F, G, Q, and R. The result is as follows:

Closed-loop characteristic polynomial characterization.   The quantity, parametrized by F, G, Q, R,

$$\alpha(s) = \det[R + G'(-sI - F')^{-1}Q(sI - F)^{-1}G]\det(sI - F)\det(-sI - F') \tag{5.2-16}$$

is a polynomial that is even in s. It can be factored as

$$\alpha(s) = \beta(s)\beta(-s) \tag{5.2-17}$$

where $\beta(s)$ has all roots in Re $[s] < 0$. Then $\beta(s)$ is (a scalar multiple of) the closed-loop characteristic polynomial $\det[sI - F - GK']$.

Why is this so? Observe first that

$$\det[I - K'(sI - F)^{-1}G] = \det[I - (sI - F)^{-1}GK']$$

$$= \det(sI - F)^{-1}\det(sI - F - GK')$$

$$= \frac{\det(sI - F - GK')}{\det(sI - F)}$$

Taking determinants in (5.2-3) with s replacing $j\omega$ and using the above definition of $\alpha(s)$ yields

$$\alpha(s) = \det\,(-sI - F' - KG')\,\det R\,\det\,(sI - F - GK')$$

Obviously, $\beta(s) = (\det R)^{1/2}\,\det\,(sI - F - GK')$.

Scalar plants: the closed-loop eigenvalues and feedback gain.

For scalar input plants, the result becomes especially simple. Define $p_0(s) = \det\,(sI - F)$ and $p_c(s) = \det\,(sI - F - GK')$. Then

$$p_0(-s)p_0(s)g'(-sI - F')^{-1}Q(sI - F)^{-1}g = q(s) \tag{5.2-18}$$

for some even polynomial $q(\cdot)$, nonnegative on the $j\omega$-axis, and (5.2-3) becomes

$$p_0(-s)p_0(s) + r^{-1}q(s) = p_c(-s)p_c(s) \tag{5.2-19}$$

Note also that for scalar plants, if one knows F, g, and $p_c(s)$, then from

$$1 - k'(sI - F)^{-1}g = \frac{p_c(s)}{p_0(s)} \tag{5.2-20}$$

it is trivial to find k. So we have for scalar plants another way of determining k: first, one obtains the closed-loop characteristic polynomial using (5.2-19)—note again that $p_c(s)$ has all roots in $\mathrm{Re}\,(s) < 0$—and hence one obtains k via (5.2-20). Actually, this construction can be generalized to multiple-input plants, and we outline the ideas for the interested reader.

Polynomial matrix fraction descriptions.†

For some purposes, it can be convenient to represent a real rational transfer function matrix as a fraction of two polynomial matrices, for example, $W(s) = B(s)P^{-1}(s)$, generalizing the obvious scalar description. Notions of coprimeness can be defined. Also, if $[F, G]$ is completely controllable, one can find a $P_0(s)$ such that when H varies in $H'(sI - F)^{-1}G$, the only variation in the associated fraction is in the numerator, that is, there is a bijective mapping $H \leftrightarrow B_H(s)$ so that

$$H'(sI - F)^{-1}G = B_H(s)P_0^{-1}(s) \tag{5.2-21}$$

Let $Q = DD'$. Then the return difference equality (5.2-3) states, with $j\omega$ replaced by s,

$$R + [P_0^{-1}(-s)]'B_D'(-s)B_D(s)P_0^{-1}(s)$$
$$= \{I - [P_0^{-1}(-s)]'B_K'(-s)\}R[I - B_K(s)P_0^{-1}(s)]$$

which is equivalent to

$$P_0'(-s)RP_0(s) + B_D'(-s)B_D(s)$$
$$= [P_0'(-s) - B_K'(-s)]R[P_0(s) - B_K(s)] \tag{5.2-22}$$

† This subsection may be omitted if desired.

Since F, G, Q, and R are part of the problem data, the left side of (5.2-22) is known. The right side is a polynomial matrix generalization of the right side of the scalar identity (5.2-19). Also, $P_0(s) - B_K(s) \triangleq P_c(s)$ has determinant which is det $(sI - F - GK')$, and so is stable. It turns out that stability of this polynomial, plus the knowledge that $B_K(s)P_0^{-1}(s)$ is strictly proper, is enough to factor the left side of (5.2-22) and determine $B_K(s)$ uniquely. Using the bijection behind (5.2-21), we find that this gives K.

The zeros of the transfer function matrix $K'(sI - F)^{-1}G$ will not be the same as those of the state variable realization $\{F, G, K\}$ if the latter is nonminimal. (The distinction is made in the appendix). Let us observe that if $[F, G]$ is stabilizable and $[F, D]$ is detectable, then $[F, K]$ is necessarily detectable, so all nonminimal modes are stable, for the following reason. If $K'w = 0$ and $(\lambda I - F)w = 0$, then $(\lambda I - F - GK')w = 0$, so that Re $\lambda < 0$, since $F + GK'$ is necessarily stable. Hence the conclusion extends to zeros of the state-variable realization.

Minimum phase property of loop gain.   The zeros of the loop gain $K'(sI - F)^{-1}G$ necessarily lie in Re $[s] \leq 0$. To see this, consider first the single-input case, and suppose (to secure a contradiction) there is a zero s_0 with Re $[s_0] > 0$. Now s_0 must also be a zero of $k'(sI - F - gk')^{-1}g$, since zeros are unaffected by state-variable feedback. It follows that the bounded real function $S(s) = 1 + k'(sI - F - gk')^{-1}g$ obeys $S(s_0) = 1$. Now $S(s)$ is analytic in Re $[s] \geq 0$ and bounded in magnitude by 1 on the $j\omega$-axis. Therefore, by the maximum modular theorem of complex analysis, $S(s) = 1$ for all s, that is, $k'(sI - F)^{-1}g \equiv 0$, which is nonsense (unless $Q = 0$). A variant on this argument can be developed for the multivariable case, but of course the use of the maximum modulus theorem is more involved. Notice that the zeros do not necessarily lie in Re $[s] < 0$, but only in Re $[s] \leq 0$. However, if $G'(j\omega I - F')^{-1}Q(j\omega I - F)^{-1}G > 0$ for all ω, then $I - S'(j\omega)S(j\omega) > 0$ for all ω, and it is impossible to have $S(j\omega_0)v = v$ for some ω_0 and $v \neq 0$; so it is impossible to have $K'(j\omega_0 I - F - GK')^{-1}Gv_0 = 0$; that is, it is impossible to have $j\omega_0$ as a zero of $K'(sI - F)^{-1}G$. Another way of seeing the minimum phase property is available for those who are familiar with the properties of positive real transfer functions and matrices. Such objects necessarily have all zeros in Re $[s] \leq 0$. Now (5.2-15) is positive real and has the same zeros as $K'(sI - F)^{-1}G$. Positive real functions can have $j\omega$-axis zeros; they are necessarily simple. This fact naturally carries over to the zeros of the loop gain.

Main points of the section.   The Return Difference Equality relates the return difference $I - K'(sI - F)^{-1}G$ to an expression formed from F, G, Q, and R. Several related equalities or inequalities can be obtained. For single input systems, one of these states that $|1 - k'(j\omega I - F)^{-1}g| \geq 1$ for all ω. The equality is a form of spectral factorization and from it, the positive realness and bounded realness of two different transfer function matrices can be established. It also yields a polynomial spectral factorization formula, which yields the closed-loop

characteristic polynomial, and for single-input plants, the gain vector k can be found therefrom.

Problem 5.2-1.   Establish the identity (5.2-4). [*Hint:* Establish first that

$$H'(j\omega I - F)^{-1}G[I - K'(j\omega I - F)^{-1}G]^{-1}$$
$$= H'(j\omega I - F - GK')^{-1}G$$

and that

$$[I - K'(j\omega I - F)^{-1}G]^{-1} = I + K'(j\omega I - F - GK')^{-1}G]$$

Problem 5.2-2.   In the notation of this section, show that

$$Z(s) = -RK'(sI - F - \tfrac{1}{2}GK')^{-1}G$$

is positive real. *Hint:* Show that the steady state Riccati equation can be written as

$$\bar{P}(F + \tfrac{1}{2}GK') + (F' + \tfrac{1}{2}KG')\bar{P} = -Q$$

Rewrite this as

$$\bar{P}(sI - F - \tfrac{1}{2}GK') + (s^*I - F' - \tfrac{1}{2}KG')\bar{P} = Q + 2\,\mathrm{Re}[s]\,\bar{P}$$

Premultiply and postmultiply by certain quantities to achieve $Z(s) + Z'(s^*)$ on the left side and a nonnegative matrix on the right.

Problem 5.2-3.   In the notation of this section, suppose that $Q > 0$ and F has no pure imaginary eigenvalues. For a single input system, show that for all finite ω, $|1 - k'(j\omega I - F)^{-1}g| > 1$.

Problem 5.2-4.   Consider a collection of performance indices with the same Q, R but indexed by $\alpha \geq 0$:

$$V(x(t_0), u(\cdot), \alpha) = \int_{t_0}^{\infty} e^{2\alpha t}(u'Ru + x'Qx)\, dt$$

Let $\bar{P}_\alpha$, K_α denote the corresponding steady state Riccati equation solution and optimal gain. Establish a return difference equality, involving F, G, Q, R, $\bar{P}_\alpha$, and $K'_\alpha(sI - F)^{-1}G$. Show that if G is a vector, there holds, whenever $\alpha_1 < \alpha_2$,

$$|1 - k'_{\alpha_1}(j\omega I - F)^{-1}g| < |1 - k'_{\alpha_2}(j\omega I - F)^{-1}g|$$

for all finite ω. [*Hint:* An intermediate result is $\bar{P}_{\alpha_1} < \bar{P}_{\alpha_2}$.]

Problem 5.2-5.   Consider the system $\dot{x} = Fx + Gu$ with performance index

$$V = \int_{t_0}^{\infty} (u'Ru + 2x'Su + x'Qx)\, dt$$

Assume that an optimum exists and that the optimal control law $u = K'x$ is stabilizing. The steady state Riccati equation turns out to be

$$\bar{P}(F - GR^{-1}S') + (F' - SR^{-1}G')\bar{P} - \bar{P}GR^{-1}G'\bar{P} + (Q - SR^{-1}S') = 0$$

with $K = -(\bar{P}G + S)R^{-1}$. Establish the return difference equality

$$R + G'(-sI - F')^{-1}S + S'(sI - F)^{-1}G + G'(-sI - F')^{-1}Q(sI - F)^{-1}G$$

$$= [I - G'(-sI - F')^{-1}K]R[I - K'(sI - F)^{-1}G]$$

Problem 5.2-6.   The Hamiltonian matrix

$$M = \begin{bmatrix} F & -GR^{-1}G' \\ -Q & -F' \end{bmatrix}$$

is used to develop properties of the matrix $\bar{P}$, or as a basis for computing $\bar{P}$. Show by direct manipulation involving M that

$$\det(sI - M) = \det(sI - F)\det(sI + F')$$

$$\det[R + G'(-sI - F')^{-1}Q(sI - F)^{-1}G]\det R^{-1}$$

5.3 SOME CLASSICAL CONTROL IDEAS: SENSITIVITY, COMPLEMENTARY SENSITIVITY, AND ROBUSTNESS

In the previous section, a classical control concept, the return difference, appeared. In this section, we shall digress from the discussion of optimal systems to review some classical control ideas. For the sake of generality, we shall discuss multivariable systems. A good reference is [3].

We refer first to Fig. 5.3-1, assumed to depict a multivariable loop. The quantities r, e, u, y, d, and n are respectively the external (reference) input, the measured tracking error, the plant input, the plant output, the disturbance signal
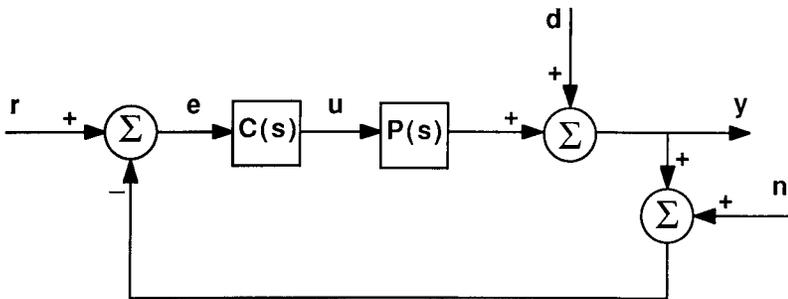
Figure 5.3-1   Classical control loop including disturbance and noise.

(referred to the plant output), and measurement or sensor noise. The quantity $r - y$ is usually termed the tracking error. To distinguish it from the measured tracking error $e = r - (y + n)$, we can term $\bar{e} = r - y$ the noise free tracking error.

The following equations are easily established:

$$y = PC(I + PC)^{-1}(r - n) + (I + PC)^{-1}d \tag{5.3-1}$$

$$e = r - y - n = (I + PC)^{-1}(r - d) - (I + PC)^{-1}n \tag{5.3-2}$$

$$u = C(I + PC)^{-1}(r - n - d) \tag{5.3-3}$$

We assume that the closed loop is stable. This means that there is no unstable pole-zero cancellation in forming the product PC, and that each transfer function matrix in (5.3-1) to (5.3-3) is stable. Alternatively, if we were to introduce a further external input v, adding on to the plant input so that $u = v + Ce$, then stability would correspond to all the transfer functions from r, v (also d and n) to u, e (and y) being stable. For a good discussion of such ideas, see [2], and Problem 5.3-1.

The quantities

$$S = (I + PC)^{-1} \tag{5.3-4}$$

and

$$T = PC(I + PC)^{-1} \tag{5.3-5}$$

are known as the sensitivity function and complementary sensitivity functions. Of course, S^{-1} is the return difference.

Notice that

$$S + T = I \tag{5.3-6}$$

We now make some key observations.

1. For good tracking, i.e. $\|r - y\|$ small when d and n are zero, (5.3-2) shows that $(I + PC)^{-1}$ should be small. More specifically, at any frequency ω, for good tracking we need

$$\bar{\sigma}[S(j\omega)] \ll 1 \tag{5.3-7}$$

(Here, $\bar{\sigma}$ denotes largest singular value; see Appendix A. In the scalar case, $\bar{\sigma}[S] = |S|$.) Of course, (5.3-7) is like the large loop gain condition of classical control.

2. For good disturbance suppression, that is, d affects y to the least extent possible, (5.3-1) shows gain that $(I + PC)^{-1}$ should be small; that is, (5.3-7) should hold.

3. For good noise suppression, that is, n affects y to the least extent possible, (5.3-1) shows that $PC(I + PC)^{-1}$ should be small; that is,

$$\bar{\sigma}[T(j\omega)] \ll 1 \tag{5.3-8}$$

The inconsistency between the objectives of good tracking and disturbance

rejection on the one hand, and good noise suppression on the other, is manifest, given (5.3-6) through (5.3-8).

Let us note a further potential difficulty with enforcing (5.3-7), at least in a certain frequency range.

4. Under (5.3-7),

$$u \simeq P^{-1}(r - n - d) \tag{5.3-9}$$

and so if (5.3-7) holds at frequencies outside the bandwidth of the plant, that is, where $\bar{\sigma}[P]$ is small or $\underline{\sigma}[P^{-1}] = [\bar{\sigma}[P]]^{-1}$ is large, then u will be large (and may cause plant saturation).

Both S and T play a role in considering the effects of plant variation. Consider the open-loop control arrangement of Fig. 5.3-2, where $\check{C}(s)$ is so chosen that the same transfer function matrix from r to y is secured. Thus $\check{C} = C(I + PC)^{-1}$. Suppose further that the plant depends on a parameter μ which can vary without any corresponding change in $\check{C}$. Variations of μ in the open-loop arrangement have a direct effect on the output y. In contrast, if μ varies in the closed-loop setup, y is varied, and the variations are fed back. They may inject a signal that compensates for the variation in μ—historically, this was one of the earlier aims of using feedback, and perhaps the main aim in electronic amplifier design where the "plant" included a vacuum tube with highly variable gain. For the open-loop design, there is no such compensation. Let us compare the two possibilities. With $r(\cdot)$ fixed for the first setup, we have

$$y_c = P(j\omega; \mu)C(j\omega)[I + P(j\omega; \mu)C(j\omega)]^{-1}r$$

and

$$\frac{\partial y_c}{\partial \mu} = [I + P(j\omega; \mu)C(j\omega)]^{-1}\frac{\partial P(j\omega; \mu)}{\partial \mu}C(j\omega)[I + P(j\omega; \mu)C(j\omega)]^{-1}r$$

For the second setup, we have

$$y_0 = P(j\omega; \mu)C(j\omega)[I + P(j\omega; \mu_{nom})C(j\omega)]^{-1}r$$

$$\frac{\partial y_0}{\partial \mu} = \frac{\partial P(j\omega; \mu)}{\partial \mu}C(j\omega)[I + P(j\omega; \mu_{nom})C(j\omega)]^{-1}r$$

Hence if the derivatives are evaluated at μ_{nom},

$$\frac{\partial y_c}{\partial \mu} = [I + P(j\omega; \mu_{nom})C(j\omega)]^{-1}\frac{\partial y_0}{\partial \mu}$$

$$= S(j\omega)\frac{\partial y_0}{\partial \mu} \tag{5.3-10}$$

Thus we can state our next observation; see [4, 5]:
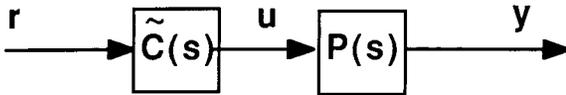
Figure 5.3-2 Open-loop control equivalent to the closed-loop arrangement of Figure 5.3-1.

5. The sensitivity to structured plant parameter variations of a closed-loop control loop is much less than that of the equivalent open-loop control when the sensitivity function is small, that is, when (5.3-7) holds. (Hence the attempt in classical design to keep the loop gain high to suppress the effect of plant parameter variations.)

There is a second type of plant variation we need to consider also, and that is unstructured multiplicative variation in the plant, which is typically associated with high-frequency uncertainty. Thus suppose that $P(j\omega)$ can be perturbed to

$$P_L(j\omega) = [I + L(j\omega)]P(j\omega) \tag{5.3-11}$$

with

$$\overline{\sigma}(L(j\omega)) < l(j\omega) \tag{5.3-12}$$

for some scalar $l(\omega)$, and suppose further that $P_L(j\omega)$ has the same number of unstable modes as $P(j\omega)$. This type of variation is in some ways more useful than an additive variation where $P_L(j\omega) = P(j\omega) + L(j\omega)$. It is preserved when a precompensator is ahead of the plant. It captures more easily the neglect of high frequency dynamics. Thus suppose that $P_L(j\omega)$ is really $\alpha(j\omega + \alpha)^{-1}P(j\omega)$ for some large α. Then $L = -j\omega(j\omega + \alpha)^{-1}$ and $l(\omega) = 1 + \epsilon$ for arbitrary positive ϵ. If one overestimates the roll-off rate in $P(j\omega)$, $l(j\omega)$ can be unbounded. Again, if there is phase uncertainty (as is frequent) at high frequencies, this can be captured by $L(j\omega)$ in a way which usually requires sup $l(j\omega)$ to be at least 2. The model even captures sensor failure, where $L(j\omega)$ can be a diagonal $[-1, 0, \ldots, 0]$ and thus $l(\omega) = 1 + \epsilon$ for arbitrary positive ϵ.

Now let us ask the question: what conditions on P, C, and L or l will ensure retention of closed-loop stability? Consider Figure 5.3-3a, which is a rearrangement of the perturbed plant with controller. The transfer function matrix from X to Y is $-C(I + PC)^{-1}$. It is possible to argue then that if the loop gain in the equivalent Figure 5.3-3b is less than 1, stability is retained. The stability condition is then

$$\overline{\sigma}(LPC(I + PC)^{-1}) < 1 \tag{5.3-13}$$

By invoking (5.3-12) and the fact that $\overline{\sigma}(AB) \le \overline{\sigma}(A)\overline{\sigma}(B)$, a sufficient condition for (5.3-13) becomes

$$\overline{\sigma}(T) = \overline{\sigma}[PC(I + PC)^{-1}] < l^{-1}(j\omega) \tag{5.3-14}$$

In fact, this condition is effectively both necessary and sufficient. If (5.3-14) holds, then for all L satisfying (5.3-12), perturbations of $P(j\omega)$ via (5.3-11) which preserve the unstable pole count will not destroy preexisting stability, while if the reverse

(a)

(b)
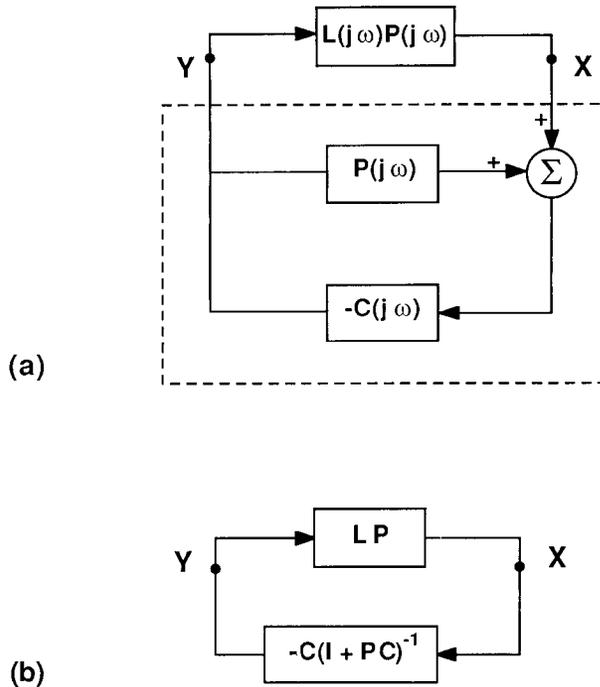
Figure 5.3-3 Redrawing of perturbed plant and controller combination.

inequality holds in (5.3-14), there exists a particular $L(j\omega)$ perturbing $P(j\omega)$ but preserving the unstable pole count such that preexisting stability is changed to instability via the perturbation. (For a derivation, see [3].) Inequality (5.3-14) also has the simple reformulation

$$l(j\omega) < \underline{\sigma}[(I + PC)(PC)^{-1}] = \underline{\sigma}[I + (PC)^{-1}] \tag{5.3-15}$$

Reference [6] makes a general argument for the association of stability robustness and sensitivity improvement conditions. Evidently, the complementary sensitivity function is associated here with stability robustness. We can thus make the observation:

6. For the retention of stability in the face of unstructured muliplicative plant variations, or structured variations modeled in like manner, it is desirable [and essential if all variations subject to (5.3-12) are possible] to have a complementary sensitivity function bounded as in (5.3-14), or an inverse-loop gain function bounded as in (5.3-15).

It is well-known from classical control that to avoid the effects of plant uncertainty at high frequencies, the loop gain must be kept small. Inequality (5.3-15) is a multivariable generalization.

We have already stated that it is virtually essential to satisfy the bound on $\overline{\sigma}(T)$, which normally means keeping $\overline{\sigma}(T)$ small in certain frequency ranges. The other reason for keeping $\overline{\sigma}(T)$ small, viz. the desirability of minimizing the effects of measurement noise, may be rendered nugatory by the use of good sensors, but there is no such escape from (5.3-14).

So the fundamental limitation on securing good performance for a control system, as indicated by $\overline{\sigma}(S) \ll 1$, is the necessity to limit $\overline{\sigma}(T)$ in certain frequency bands. Generally, this leads to the requirement that $\overline{\sigma}(S)$ should be low in the passband and $\overline{\sigma}(T)$ low in the stopband, and in turn that $\overline{\sigma}(PC)$ is large in the passband and $\underline{\sigma}(PC)^{-1}$ large in the stopband. It is regarded as good practice for all the singular values of S or of $S^{-1} = I + PC$ to have roughly the same cross-over frequency.

We have covered above some of the main issues in classical design which involve the sensitivity and complementary sensitivity functions. We have not discussed a range of results (gain–phase constraints) which in the scalar case have the effect of constraining average values of S, or its logarithm, and mathematically highlight the greater difficulties of controlling nonminimum phase plants; see, for example, [7]. Extensions to the multivariable case of some of these results have begun to appear [8].

Our discussion of multiplicative uncertainty focused on uncertainty at the plant output. One can also consider input uncertainty, with $P_L(j\omega) = P(j\omega)[I + L(j\omega)]$. In this case, (5.3-15) is replaced by

$$l(j\omega) < \underline{\sigma}[I + (CP)^{-1}]$$

Actuator failure corresponds to $l(j\omega) > 1$. Of course, it is only in the multivariable case that there is a distinction between having $\underline{\sigma}[I + (CP)^{-1}]$ and $\underline{\sigma}[I + (PC)^{-1}]$.

Main points of the section.   Two key transfer function matrices, the sensitivity function S and complementary sensitivity function T, can be defined for closed-loop systems. They sum to the identity matrix. Table 5.1 sums up some key properties.

Generally, one needs $\overline{\sigma}(S)$ small in the passband of the plant, and $\overline{\sigma}(T)$ small in the stopband.

TABLE 1   THE ROLE OF SENSITIVITY AND COMPLEMENTARY
SENSITIVITY IN CLASSICAL CONTROL

Property Desired	S or T Constraint
Tracking	$\overline{\sigma}(S)$ small
Disturbance suppression	$\overline{\sigma}(S)$ small
Noise suppression	$\overline{\sigma}(T)$ small
Control magnitude limitation	$\overline{\sigma}(S)$ not small when $\overline{\sigma}(P)$ is small
Sensitivity to structured plant parameter variations	$\overline{\sigma}(S)$ small
Sensitivity to unstructured multiplicative uncertainty	$\overline{\sigma}(T)$ small

Problem 5.3-1.   Consider Fig. 5.3-1 with the insertion of a further additive input v at the plant input, so that $u = v + Ce$. Verify that all transfer function matrices from r, v to e, u are stable if and only if the following transfer function matrices are stable:

$$(I + PC)^{-1}, \qquad (I + CP)^{-1}, \qquad C(I + PC)^{-1}, \qquad (I + PC)^{-1}P$$

or equivalently,

$$\begin{bmatrix} I & -P \\ C & I \end{bmatrix}^{-1} \text{ is stable}$$

Problem 5.3-2.   Let $\bar{\sigma}(A)$, $\underline{\sigma}(A)$ denote maximum and minimum singular values. Show that

$$[\underline{\sigma}(PC) + 1]^{-1} \leq \bar{\sigma}(S) \leq [\underline{\sigma}(PC) - 1]^{-1}$$

$$[\bar{\sigma}(S)]^{-1} - 1 \leq \underline{\sigma}(PC) \leq [\bar{\sigma}(S)]^{-1} + 1$$

These inequalities relate small sensitivity functions to large loop gain. [*Hint:* Establish first that $\underline{\sigma}(A) - 1 \leq \underline{\sigma}(A + I) \leq \underline{\sigma}(A) + 1$ by using the characterization $\underline{\sigma}(M) = \min_{\|x\| = 1} \|Mx\|$, together with the triangle inequality. The identity $\underline{\sigma}(A) = [\bar{\sigma}(A^{-1})]^{-1}$ may be helpful also.]

5.4   GAIN MARGIN, PHASE MARGIN, AND TIME-DELAY TOLERANCE

In this section, we shall examine certain properties of the closed-loop scheme depicted in Fig. 5.4-1, where K arises from a linear quadratic design with no cross-product terms in the performance index. We shall be especially interested in the gain margin, phase margin, and time delay tolerance of the scheme for the scalar input case. Where corresponding conclusions apply for the multivariable case, we shall seek to make them. We shall also relate our conclusions to the ideas of the previous section.

Single-input systems.   We recall that the gain margin of a closed-loop system is the amount by which the loop gain can be changed until the system becomes unstable. If the loop gain can be increased without bound—that is, instability is not encountered, no matter how large the loop gain becomes—then the closed-loop system is said to possess an infinite gain margin.

Of course, no real system has infinite gain margin. Such parasitic effects as
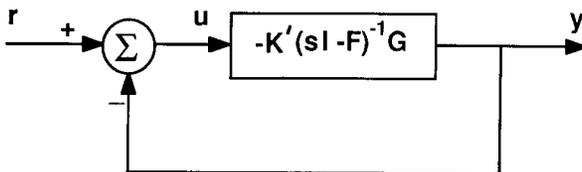
Figure 5.4-1   Closed-loop optimal scheme redrawn as unity negative feedback scheme.

stray capacitance, time delay, and the like will always prevent infinite gain margin from being a physical reality. Some mathematical models of systems may, however, have an infinite gain margin. Clearly, if these models are accurate representations of the physical picture—save, perhaps, for their representation of parasitic effects—it could validly be concluded that the physical system had a very large gain margin.

We shall now show that the optimally designed regulator possesses the infinite gain margin property, as well as a "downside" gain margin, by noting a characteristic feature of the Nyquist diagram of the open-loop gain of the regulator. The scheme of Fig. 5.4-1 is arranged to have unity *negative* feedback, so that we may apply the Nyquist diagram ideas immediately. The associated Nyquist plot is a curve in the complex plane, obtained from the complex values of $-k'(j\omega I - F)^{-1}g$ as ω varies through the real numbers from minus to plus infinity. Now the Nyquist plot of $-k'(j\omega I - F)^{-1}g$ is constrained to avoid a certain region of the complex plane, because the return difference equality implies (in the case of no cross-product terms in the quadratic index)

$$|1 - k'(j\omega I - F)^{-1}g| \geq 1 \qquad (5.4\text{-}1)$$

which is to say that the distance of any point on the Nyquist plot from the point $-1 + j0$ is at least unity. In other words, the plot of $-k'(j\omega I - F)^{-1}g$ avoids a circle of unit radius centered at $-1 + j0$.

See Fig. 5.4-2 for examples of various plots. (The transfer functions are
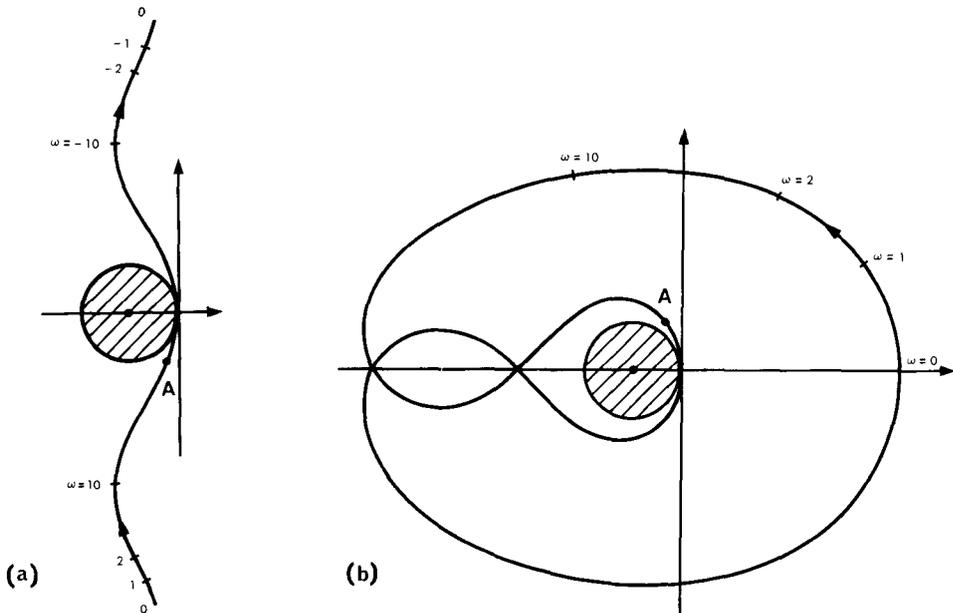
Figure 5.4-2   Nyquist plots of $-k'(j\omega I - F)^{-1}g$ avoiding a unit critical disc center $(-1, 0)$. Points A are at unity distance from the origin.

irrelevant.) The arrow marks the direction of increasing ω. Note that the plots end at the origin, which is always the case when the open-loop transfer function, expressed as a numerator polynomial divided by a denominator polynomial, has the numerator degree less than the denominator degree (i.e., the transfer function is strictly proper).

There is yet a further constraint on the Nyquist plot, which is caused by the fact that the closed-loop system is known to be asymptotically stable. This restricts the number of *counterclockwise encirclements* of the point $-1 + j0$ by the Nyquist plot to being precisely the number of poles of the transfer function $-k'(sI - F)^{-1}g$ lying in Re $[s] \geq 0$. We understand that if a pole lies on Re $[s] = 0$, the Nyquist diagram is constructed by making a small semicircular indentation into the region Re $[s] < 0$ around this pole, and plotting the complex numbers $-k'(sI - F)^{-1}g$ as s moves around this semicircular contour.

Thus, if $-k'(sI - F)^{-1}g$ has no poles in Re $[s] \geq 0$, a diagram such as that in Fig. 5.4-2(b) may be obtained. Figure 5.4-3 illustrates a case where $-k'(sI - F)^{-1}g$ has two poles in Re $[s] > 0$.

The key observation we now make is that *the number of encirclements of the point $-1 + j0$ is the same as the number of encirclements of any other point inside the circle of unit radius and center $-1 + j0$.* The best way to see this seems to be by visual
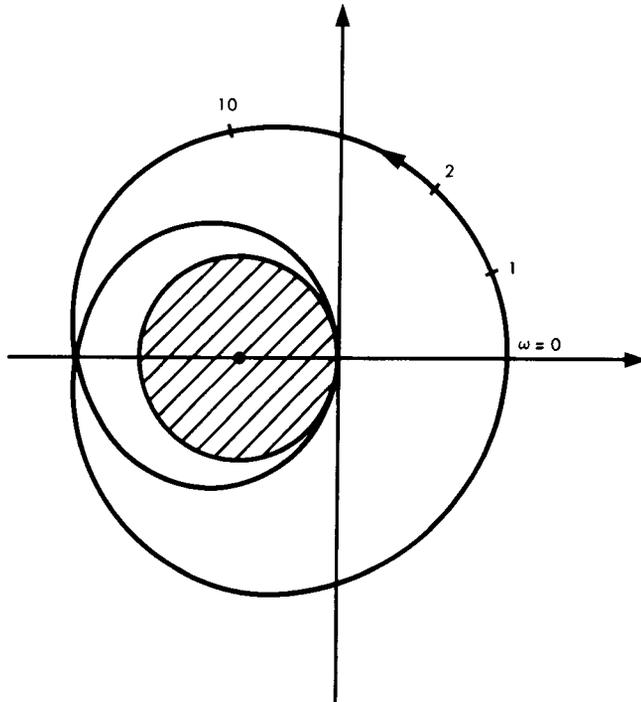
Figure 5.4-3 Nyquist plot for which open-loop transfer function has two poles in Re $[s] > 0$, and closed-loop is stable.

inspection: a few experiments will quickly show that if the preceding remarks were not true, the Nyquist diagram would have to enter the circle from which it has been excluded.

It is known that the closed-loop system with gain multiplied by a constant factor β will continue to be asymptotically stable if the Nyquist diagram of $-\beta k'(j\omega I - F)^{-1}g$ encircles $-1 + j0$ in a counterclockwise direction a number of times equal to the number of poles of $-\beta k'(sI - F)^{-1}g$ lying in Re $[s] \geq 0$. Equivalently, asymptotic stability will follow if the Nyquist diagram of $-k'(j\omega I - F)^{-1}g$ encircles $-(1/\beta) + j0$ this number of times. But our previous observation precisely guarantees this for a range of β. The points $-(1/\beta) + j0$ for all real $\beta > \frac{1}{2}$ lie inside the critical circle, and thus are encircled counterclockwise the same number of times as the point $-1 + j0$. As we have argued, this number is the same as the number of poles of $-k'(sI - F)^{-1}g$ in Re $[s] \geq 0$. Consequently, with asymptotic stability following for all real $\beta > \frac{1}{2}$, we have established the infinite gain margin property, and a downside margin of $\frac{1}{2}$.

Let us now turn to consideration of the phase margin property. First, we recall the definition of phase margin. It is the amount of negative phase shift that must be introduced (without gain increase) to make that part of the Nyquist plot corresponding to $\omega \geq 0$ pass through the $-1 + j0$ point. For example, consider the three plots of Fig. 5.4-2; points A at unit distance from the origin on the $\omega \geq 0$ part of the plot have been marked. The negative phase shift that will need to be introduced in the first case is about 80 deg, and that in the second case about 280 deg. Thus, 80 deg is approximately the phase margin in the first case, and 280 deg in the second.

We shall now show that the phase margin of an optimal regulator is always at least 60 deg. The phase margin is determined from that point or those points on the $\omega \geq 0$ part of the Nyquist plot which are at unit distance from the origin. Since the Nyquist plot of an optimal regulator must avoid the circle with center $-1 + j0$ and unity radius, the points at unit distance from the origin and lying on the Nyquist plot of an optimal regulator are restricted to lying on the shaded part of the circle of unit radius and center the origin, shown in Fig. 5.4-4. The smallest angle through which
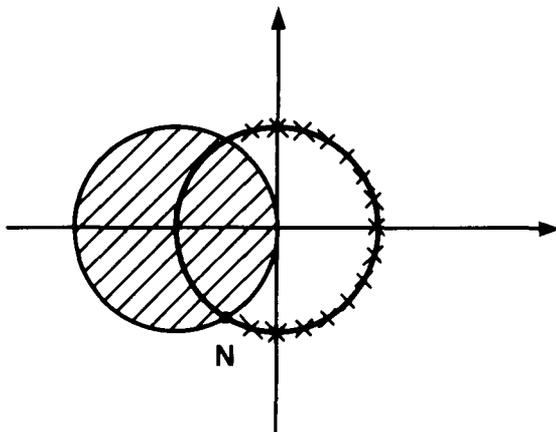
Figure 5.4-4   Shaded points denote permissible points on Nyquist plot of optimal regulator at unit distance from origin.

one of the allowable points could move in a clockwise direction to reach $-1 + j0$ is 60 deg, corresponding to the point N of Fig. 5.4-4. Any other point in the allowed set of points (those outside the circle of center $-1 + j0$, and unity radius, but at unit distance from the origin) must move through more than 60 deg to reach $-1 + j0$. The angle through which a point such as N must move to reach $-1 + j0$ is precisely the phase margin. Consequently, the lower bound of 60 deg is established.

Note that there is no assertion here that stability is retained if *both* a phase shift of no more than 60 deg is introduced *and* the gain is changed by a factor in the interval $(\frac{1}{2}, \infty)$.

We now turn to a discussion of the tolerance of time delay in the closed loop. Accordingly, we shall consider the scheme of Fig. 5.4-5, where T is a certain time delay. (The delay block could equally well be anywhere else in the loop, from the point of view of the following discussion.) We shall be concerned with the stability of the closed-loop system.

The effect of the time delay is to insert a frequency-dependent negative phase shift into the open-loop transfer function. Thus, instead of $-k'(j\omega I - F)^{-1}g$ being the open-loop transfer function, it will be $-k'(j\omega I - F)^{-1}g e^{-j\omega T}$. This has the same magnitude as $-k'(j\omega I - F)^{-1}g$, but a negative phase shift of ωT radians.

It is straightforward to derive allowable values of the time delay that do not cause instability. Suppose the transfer function $-k'(j\omega I - F)^{-1}g$ has unity gain at the frequencies $\omega_1, \omega_2, \ldots, \omega_r$ with $0 < \omega_1 < \omega_2 < \ldots < \omega_r$, and let the amount of negative phase shift that would bring each of these unity gain points to the $-1 + j0$ point be $\varphi_1, \varphi_2, \ldots, \varphi_r$, respectively. Of course, $\varphi_i \geq \pi/3$ for all i. Then, if a time delay T is inserted, so long as $\omega_i T < \varphi_i$ or $T < \varphi_i/\omega_i$ for all i, stability will prevail. In particular, if $T < \pi/(3\omega_r)$, stability is assured.

The introduction of time delay will destroy the infinite gain margin property. To see this, observe that as ω approaches infinity, the phase shift introduced—viz., ωT radians—becomes infinitely great for any nonzero T. In particular, one can be assured that the Nyquist plot of $-k'(j\omega I - F)^{-1}g$ will be rotated for suitably large ω such that the rotated plot crosses the real axis just to the left of the origin. (In fact, the plot will cross the axis infinitely often.) If for a given T, the leftmost point of the real axis that is crossed is $-(1/\beta) + j0$, then the gain margin of the closed loop with time delay is $20 \log_{10}\beta$ dB.

Of course, the introduction of a multiplying factor, phase shift, or time delay
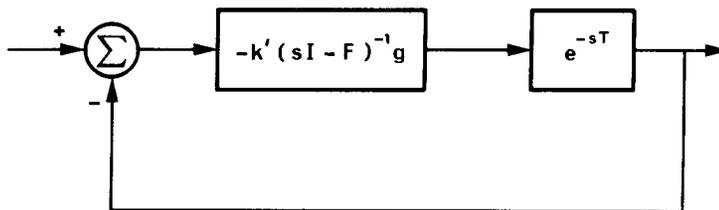
Figure 5.4-5 Single-input optimal regulator with time delay T inserted.

in the loop will destroy the optimality of the original system. But the important point to note is that the optimality of the original system allows the introduction of these various modifications while maintaining the required closed-loop stability. Optimality may not, in any specific situation, be of direct engineering significance, but stability is; therefore, optimality becomes of engineering significance indirectly.

The key to the above conclusions is the inequality (5.4-1). Notice that this gives an upper bound on the return difference. In the notation of the previous sections, it states that for all ω,

$$|S| \leq 1 \qquad (5.4\text{-}2)$$

Other implications of this inequality are explored in the last section. The complementary sensitivity function is

$$
\begin{aligned}
T &= 1 - S \\
&= 1 - [1 - k'(j\omega I - F)^{-1}g]^{-1} \qquad (5.4\text{-}3) \\
&= -k'(j\omega I - F - gk')^{-1}g
\end{aligned}
$$

(The last equality follows from an Appendix B result, which can be established by direct algebraic manipulation.)

Observe that

$$\lim_{\omega \to \infty} j\omega T = -k'g = r^{-1}g'\bar{P}g \qquad (5.4\text{-}4)$$

Except in the abnormal case when $k = 0$, and so $\bar{P}g = 0$, there holds $g'\bar{P}g > 0$. So $|T|$ rolls off at a rate of ω^{-1}. This is *not* an attractive property if multiplicative unstructured uncertainty in the plant is contemplated, as it might not be fast enough. (Recall that $|T|$ should be small where the uncertainty is large.) In a sense, this is the penalty being paid for the attractive features such as good gain margin and phase margin described earlier.

To reemphasise the fact that the attractive properties described hitherto will not resolve all robustness or sensitivity issues satisfactorily, let us note another potential danger. Suppose a design of k is obtained from a relatively low penalty in control energy with the property that the closed-loop bandwidth, dictated by the eigenvalues of $F + gk'$, significantly exceeds the open-loop bandwidth, dictated by the eigenvalues of F. (Classical control would suggest that this is unwise, requiring overly large control signals.) Reference [9] shows via one example that there can be reduced robustness to variations in individual entries of the g vector, and with very large k, there may be almost no robustness. The idea is developed in Problem 5.4-3.

The infinite gain margin property is consistent with the just proved result on the roll-off rate (relative degree of $k'(sI - F)^{-1}g$) and the minimum phase property established earlier of $k'(sI - F)^{-1}g$. The minimum phase property indicates that as the loop gain goes to infinity, certain closed-loop poles tend to the zeros of $k'(sI - F)^{-1}g$ and the remainder tend to infinity, with the number of such poles equal to the relative degree of $k'(sI - F)^{-1}g$.

Multi-input systems.   Multi-input systems are, as one might expect, more awkward to consider than single-input systems. However, our starting point is still the return difference equality, from which we have

$$[I - G'(-j\omega I - F')^{-1}K]R[I - K'(j\omega I - F)^{-1}G] \geq R \qquad (5.4-5)$$

So if $R = \rho I$, we can state that

$$\underline{\sigma}[I - K'(j\omega I - F)^{-1}G] \geq 1 \qquad (5.4-6)$$

for all ω and further that

$$\overline{\sigma}[S] \leq 1 \qquad (5.4-7)$$

since $S^{-1} = I - K'(j\omega I - F)^{-1}G$.

The inequality (5.4-7) also gives us an inequality for $\overline{\sigma}(T)$. Since $T = I - S$, $\overline{\sigma}(T) \leq 1 + \overline{\sigma}(S) \leq 2$ or

$$2 \geq \overline{\sigma}(T) \qquad (5.4-8)$$

It is not hard to find particular examples where the bound in (5.4-8) can be attained. So (5.4-8) is the best general result. In the light of (5.3-14), this means that multiplicative unstructured uncertainty which is guaranteed not to disturb stability must satisfy $l^{-1}(j\omega) > 2$ or $l(j\omega) < \frac{1}{2}$. This rules out, for example, an unmodeled high-frequency mode. This disappointing fact is consistent with the slow roll-off rate, viz. ω^{-1}, of $\overline{\sigma}(T)$ as $\omega \to \infty$. But note that for some particular examples, there may be greater tolerance of multiplicative unstructured uncertainty.

For more general R, we can claim only that

$$\underline{\sigma}\{R^{1/2}[I - K'(j\omega I - F)^{-1}G]R^{-1/2}\} \geq 1 \qquad (5.4-9)$$

and

$$\overline{\sigma}\{R^{1/2}SR^{-1/2}\} \leq 1 \qquad (5.4-10)$$

Are (5.4-9) and (5.4-10) even desirable? We shall study this question, investigating in the process the gain and phase margin results that flow from (5.4-7) in case $R = \rho I$ for some ρ.

Our first goal will be to show that with R any positive definite diagonal matrix we do again get attractive gain and phase margin properties, as outlined in [10]. To establish this claim, we need two preliminary results, one a result of linear algebra and the second a robust stability result.

Lemma A: Let V, W be square complex matrices of the same dimensions such that

$$(I + V^*)(I + V) \geq I \qquad (5.4-11)$$

$$W^* + W > I \qquad (5.4-12)$$

Then $I + VW$ is nonsingular.

Proof.   Suppose $VWu = -u$ for some $u \neq 0$ to secure a contradiction. Now $I + V^* + V + V^*V \geq I$ implies $W^*V^*W + W^*VW + W^*V^*VW^* \geq 0$. Premultiply by u^* and postmultiply by u. There results $-u^*Wu - u^*W^*u + u^*u \geq 0$, that is, $u^*[W^* + W - I]u \leq 0$. This is a contradiction.

Lemma B: Consider the closed-loop system defined in Fig. 5.4-6. Suppose there are no unstable pole-zero cancellations in the product VW, and that for all ω, $V(j\omega)$ and $W(j\omega)$ satisfy (5.4-11) and (5.4-12). Suppose also that with W replaced by I the closed loop is stable. Then the closed loop of Fig. 5.4-6. is stable.

Proof.   Replace W by $\overline{W}_\epsilon = (1 - \epsilon)I + \epsilon W$, where ϵ can vary in $[0, 1]$. Observe that $\epsilon = 0$ corresponds to a known stable situation, and $\epsilon = 1$ to the situation of interest. Further,

$$\overline{W}_\epsilon^* + \overline{W}_\epsilon = 2(1 - \epsilon)I + \epsilon(W^* + W)$$
$$> 2(1 - \epsilon)I + \epsilon I$$
$$= (2 - \epsilon)I$$

Hence

$$\overline{W}_\epsilon^* + \overline{W}_\epsilon > I$$

Also, since $(I + V^*)(I + V) \geq I$, by assumption, then from Lemma A, $I + V(j\omega)\overline{W}_\epsilon(j\omega)$ is nonsingular for all $\omega, \epsilon \in [0, 1]$.

Now with W replaced by $\overline{W}_\epsilon$, the closed-loop transfer function matrix is $V\overline{W}_\epsilon(I + V\overline{W}_\epsilon)^{-1}$. With some work, one can argue that unstable pole-zero cancellations pose no problem. Then as ϵ moves from 0 to 1, an instability can arise only when a closed-loop pole moves from the open left half-plane to the right half-plane. In so doing, it must cross the $j\omega$-axis. That is, $I + V(j\omega)\overline{W}_\epsilon(j\omega)$ becomes singular for some $\epsilon \in [0, 1]$ and some ω, a contradiction. Thus there is closed-loop stability with W present.

Lemma B provides the recipe for obtaining phase and gain margin type results. Let us identify

$$V(s) = -R^{1/2}K'(j\omega I - F)^{-1}GR^{-1/2}$$

The return difference equality states

$$(I + V^*)(I + V) \geq I$$
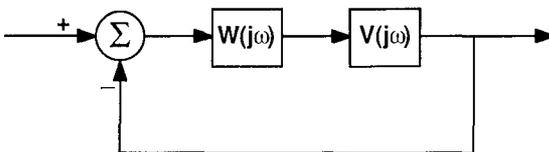
Figure 5.4-6   Closed-loop systems used for robustness result.

which is (5.4-11). Suppose that W satisfies (5.4-12) and has the structure $W(j\omega) = R^{1/2}L(j\omega)R^{-1/2}$. Then reference to Fig. 5.4-7 indicates that the setup of Fig. 5.4-7(b) will be stable when

$$(R^{-1/2})L*R^{1/2} + R^{1/2}LR^{-1/2} > I$$

or equivalently

$$L*R + RL > R \tag{5.4-13}$$

Now suppose R is diagonal and L is diag $(l_1, \ldots, l_m)$. Then (5.4-13) is satisfied if and only if

$$l_i^* + l_i > 1 \tag{5.4-14}$$

In particular if l_i is real, in the interval $(\frac{1}{2}, \infty)$, this is satisfied. This is the gain margin analogy. Also, if $l_i = e^{-j\varphi}$ with $|\varphi| < \pi/3$, (5.4-14) is satisfied. This is the phase margin analogy. Notice that simultaneous variations in different l_i, l_j are permitted.

In summary, if R is diagonal we can tolerate independent scalar gain variations between $\frac{1}{2}$ and ∞ and phase variations less than 60 deg in each scalar input, without disturbing stability. Of course, in any one scalar input, there cannot be simultaneously a gain variation of $\frac{1}{2}$ and phase shift of 60 deg.

This result may suggest that any diagonal R is acceptable. This is a misleading statement. If one diagonal entry of R is much lower than the others, the corresponding entry of the control vector defines a signal that may take very large, even unacceptably large, values. Further, there is a limit on the cross coupling that can be tolerated. Suppose that

$$R = \begin{bmatrix} R_1 & 0 \\ 0 & R_2 \end{bmatrix} \tag{5.4-15}$$

and

$$L = \begin{bmatrix} I & X \\ 0 & I \end{bmatrix} \tag{5.4-16}$$

(a)

(b)
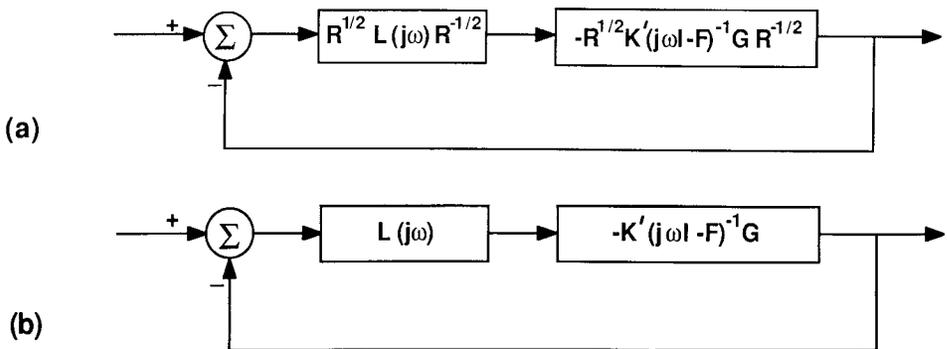
Figure 5.4-7   Two structures with the same stability properties.

Thus X represents a unidirectional cross coupling. Then (5.4-13) is satisfied when

$$\overline{\sigma}^2(X) < \frac{\lambda_{\min}(R_2)}{\lambda_{\max}(R_1)} \qquad (5.4\text{-}17)$$

Problem 5.4-6 seeks verification of this. A consequence of the condition (5.4-17) is that if there is the possibility of coupling in only one direction, but this direction is unknown, that is, if there is a possibility that

$$L = \begin{bmatrix} I & X \\ 0 & I \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} I & 0 \\ X & I \end{bmatrix} \qquad (5.4\text{-}18)$$

then the bound is

$$\overline{\sigma}^2(X) < \min\left[\frac{\lambda_{\min}(R_1)}{\lambda_{\max}(R_2)}, \frac{\lambda_{\min}(R_2)}{\lambda_{\max}(R_1)} \right] \qquad (5.4\text{-}19)$$

The greater the discrepancy between the eigenvalues of R_1 and R_2, the smaller this will be.

If R is diagonal and the cross coupling can be between any two input lines, or collections of lines, then

$$\overline{\sigma}^2(X) < \frac{\lambda_{\min}(R)}{\lambda_{\max}(R)} \qquad (5.4\text{-}20)$$

Of course, (5.4-19) and (5.4-20) are only sufficient conditions. But they serve nevertheless to highlight the potential difficulties arising when R departs from the form ρI.

Having seen what happens when R is diagonal, let us note the effect of nondiagonal R. Reference [10] includes an example that shows with nondiagonal R, the gain margins can become arbitrarily small; see Problem 5.4-5. Of course, it may also be the case for some examples that with a nondiagonal R, there is nevertheless a considerable gain margin on each line.

Main points of the section.   The return difference equality ensures for single-input plants that $|S| \leq 1$ and for multiple-input plants that $\overline{\sigma}[R^{1/2}SR^{-1/2}] \leq 1$ and for all frequencies. This condition translates, provided that R is diagonal in the multivariable case, to gain margins in each loop of $(\tfrac{1}{2}, \infty)$ and phase margins of 60 deg or more. However, the loop gain and complementary sensitivity function roll off at a rate of only ω^{-1}, and there is poor tolerance of multiplicative unstructured uncertainty. For multiple input systems, having a diagonal R with entries of very different sizes gives poor robustness to input cross coupling. In the nondiagonal R case there is no guarantee in a universal way of robustness to gain or phase variations on any single input.

Problem 5.4-1.   Show for a single-input plant that if $e^{2\alpha t}$ is introduced in the usual performance index to secure degree of stability α, this degree of stability is preserved if the loop gain is varied by a factor in the range $(\tfrac{1}{2}, \infty)$. [*Hint:* Study the

Nyquist plot of $k'(j\omega I - \alpha I - F)^{-1}g$ and prove a return difference equality involving this transfer function.]

Problem 5.4-2. Suppose that $\underline{\sigma}(I - X) \geq 1$. Show that $\underline{\sigma}(I - X^{-1}) \geq \frac{1}{2}$. [*Hint:* Observe that $(I - X^{-1})^{-1} = I - (I - X)^{-1}$ and use the triangle inequality.]

Problem 5.4-3. Suppose that G is $n \times m$ with rank m. For the multivariable case, show that for large ω, $T(j\omega)$, the complementary sensitivity function, is approximated by $(j\omega)^{-1}N$, where N is a product of two positive definite $m \times m$ matrices.

Problem 5.4-4. Consider the system

$$\dot{x} = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix} x + \begin{bmatrix} 1 + \epsilon \\ 1 \end{bmatrix} u$$

and performance index

$$V = \int_0^\infty [ru^2(t) + (x_1 - x_2)^2]\, dt$$

(i) Suppose $\epsilon = 0$. Show, using the return difference equality, that

$$1 - k'(sI - F)^{-1}g = \frac{s^2 + \sqrt{5 + 2\gamma}\, s + \gamma}{s^2 + 3s + 2}$$

where $\gamma = \sqrt{4 + r^{-1}}$.

Deduce the value of k'.

(ii) With the k from (i), consider the stability of

$$\begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix} + \begin{bmatrix} 1 + \epsilon \\ 1 \end{bmatrix} k'$$

Show that, given any $\epsilon > 0$, there exists $r(\epsilon)$ such that with $r \leq r(\epsilon)$, instability results.

Problem 5.4-5. Consider a system with

$$\dot{x} = x + \begin{bmatrix} 1 & \beta \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$V = \int_0^\infty (u'Ru + x'x)\, dt$$

with

$$R = \begin{bmatrix} 1 & 0 \\ \beta & 1 \end{bmatrix} [N^{-2} + 2N^{-1}]^{-1} \begin{bmatrix} 1 & \beta \\ 0 & 1 \end{bmatrix}$$

and $N > 0$ is arbitrary but not diagonal. Show that $\bar{P} = N$ is the solution of the steady state Riccati equation. Consider an input perturbation by the constant matrix

$$L = \begin{bmatrix} 1 & 0 \\ 0 & 1 + \epsilon \end{bmatrix}$$

Write down the closed-loop matrix $F + GLK'$ parametrized in terms of the entries of $\bar{P}^{-1}$, β, and ϵ. Show that for any $\epsilon \neq 0$, there exists a β which makes $\text{tr}(F + GLK') > 0$, implying closed-loop instability.

Problem 5.4-6.   Suppose that

$$R = \begin{bmatrix} R_1 & 0 \\ 0 & R_2 \end{bmatrix} \quad \text{and} \quad L = \begin{bmatrix} I & X \\ 0 & I \end{bmatrix}$$

Show that

$$\bar{\sigma}^2(X) < \lambda_{\min}(R_2)/\lambda_{\max}(R_1)$$

implies $L^*R + RL - R > 0$. [*Hint:* Evaluate $L^*R + RL - R$. Use the fact that if $A > 0$, $C > 0$, then

$$\begin{bmatrix} A & B \\ B^* & C \end{bmatrix} > 0$$

if and only if $C - B^*A^{-1}B > 0$.]

Problem 5.4-7.   (Requires computer solution). For the cases studied in Problem 3.3-3 examine Nyquist and Bode plots of the various designs, noting gain and phase margins.

5.5 INSERTION OF NONLINEARITIES

In the previous section, we considered the introduction of various linear system perturbations at the plant input, corresponding to gain changes, introduction of phase shift, and so on. In this section, we consider the arrangement of Fig. 5.5-1. The input nonlinearity depicted might well capture the inexact behavior of an actuator.
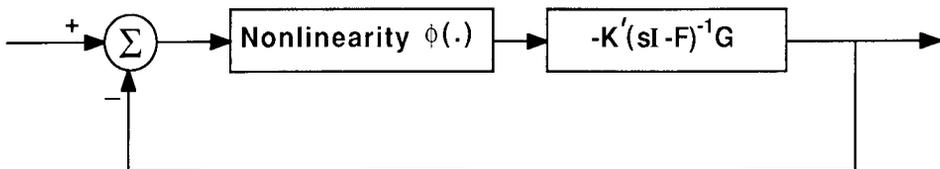
Figure 5.5-1   Introduction of an input nonlinearity.

For simplicity, we shall first consider the single-input case. We shall impose a restriction on the nonlinearity.

Nonlinearity description. The nonlinearity $\varphi(\sigma)$ is memoryless and confined within a sector that itself is confined strictly within a first-third quadrant sector. The outer sector slope bounds are $\frac{1}{2}$ and ∞; see Fig. 5.5-2. Analytically

$$\tfrac{1}{2} + \epsilon_1 \leq \frac{\varphi(\sigma)}{\sigma} \leq \epsilon_2^{-1} \tag{5.5-1}$$

for all $\sigma \neq 0$, and small positive ϵ_1, ϵ_2.

We showed in the last section that for all linear gains in the sector, closed-loop stability is preserved. This result is now generalized to achieve the following remarkable robustness result.

Robustness property. For the optimal state feedback gain K' arising from an optimal linear quadratic design (with no cross-product terms), the optimal closed-loop system maintains asymptotic stability when arbitrary nonlinearities $\varphi(\cdot)$ satisfying (5.5-1) are inserted at the plant input. That is, the closed-loop system of Fig. 5.5-1 is stable under (5.5-1). Moreover, the nonlinearities may be arbitrarily time-varying.
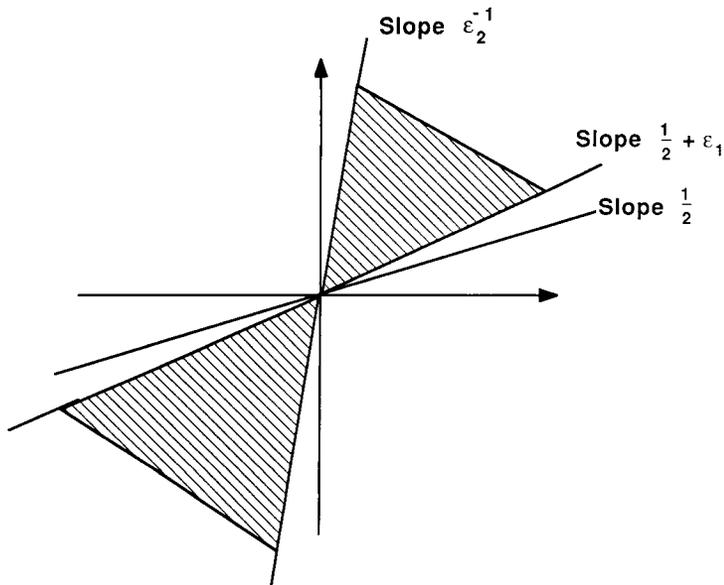
Figure 5.5-2 Sector bounds.

***Proof.*†**   For those readers familiar with the circle criterion [2], the result follows from the fact that the Nyquist plot avoids the critical unit disk centered at [−1, 0]—a direct consequence of the return difference equality.

For those who know that a positive real system back to back with a strictly passive system is asymptotically stable [2], it is instructive to see that the feedback arrangement of Fig. 5.5-1 can be reorganized as the positive real system $Z(s)$ of (5.2-13) with a feedback nonlinearity $\tilde{\varphi}(\tau) = \varphi(\sigma) - \frac{1}{2}\sigma$ scaled to be in the range $(0, \infty)$, that is, $[\epsilon_1, \epsilon_2^{-1}]$. This rearrangement is achieved by adding and subtracting a feedback gain of $\frac{1}{2}$ into the loop. Now the rescaled $\tilde{\varphi}(\cdot)$ defines a strictly passive system since, with input σ and output $\tilde{\varphi}(\sigma)$

$$\int_0^t \sigma(\tau)\varphi[\sigma(\tau)]\, d\tau \geq \epsilon_1 \int_0^t \sigma^2(\tau)\, d\tau \quad \text{for all } t$$

These observations constitute alternative proofs of the robustness result, which extend to the multivariable case and are valid whether $\varphi(\cdot)$ is time-varying or not.

For those more familiar with Lyapunov theory, we shall establish the result for the case when the regulator is designed with $[F, g]$ controllable and $[F, D]$ observable, where $DD' = Q$. The result is extendable to the stabilizable/detectable case, but the details are intricate. We shall assume that $\varphi(\cdot)$ is time-invariant. Under the stated conditions, $V(x) = x'\bar{P}x$ (with $\bar{P}$ the steady state Riccati equation solution) is positive definite. We shall evaluate $\dot{V}(x)$ along trajectories of the closed-loop system, which is evidently

$$\dot{x} = Fx + g\varphi(k'x) \tag{5.5-2}$$

We have

$$\dot{V} = x'\bar{P}[Fx + g\varphi(k'x)] + [x'F' + \varphi(k'x)g']\bar{P}x$$

Use the fact that $\bar{P}F + F'\bar{P} = krk' - Q$, and $\bar{P}g = -kr$. Then

$$\dot{V} = -x'Qx + r(k'x)^2 - 2r\varphi(k'x)(k'x) \tag{5.5-3}$$

The restriction on φ ensures that $\varphi(k'x)(k'x) \geq (\frac{1}{2} + \epsilon_1)(k'x)^2$. So

$$\dot{V} \leq -x'Qx - \epsilon_1 r(k'x)^2 \tag{5.5-4}$$

Since $\dot{V} \leq 0$, we obviously have stability. To show asymptotic stability, we have to show that $\dot{V} \equiv 0$ implies $x \equiv 0$ (see Appendix D). Now $\dot{V} \equiv 0$ implies $k'x \equiv 0$, $D'x \equiv 0$. Also $k'x \equiv 0$ implies that $\dot{x} \equiv Fx + g\varphi(k'x) \equiv Fx$. With $\dot{x} \equiv Fx$ and $D'x \equiv 0$, it follows that $x \equiv 0$. Hence asymptotic stability is established.

The multivariable case is almost as straightforward. Suppose that the plant input is $\Phi(K'x)$ instead of $K'x$. Here, Φ is a vector function of $K'x$. We shall restrict it below. Now with $V = x'\bar{P}x$, from the closed-loop system

$$\dot{x} = Fx + G\Phi(K'x) \tag{5.5-5}$$

___
† This proof may be omitted without loss of continuity.

we derive

$$\dot{V} = -x'Qx + x'KRK'x - x'KR\Phi(K'x) - \Phi'(K'x)RK'x \qquad (5.5\text{-}6)$$

Obviously, we need some condition that forces

$$2x'KR\Phi(K'x) \geq x'KRK'x \qquad (5.5\text{-}7)$$

One simple way to ensure this (as checked below) is to require that R is diagonal, and that $\Phi(\cdot)$ is a diagonal nonlinearity satisfying (5.5-1); that is,

$$\Phi(K'x) = \begin{bmatrix} \varphi_1[(K'x)_1] \\ \varphi_2[(K'x)_2] \\ . \\ . \\ . \end{bmatrix} \qquad (5.5\text{-}8)$$

with

$$\tfrac{1}{2} + \epsilon_1 \leq \varphi_i(u)/u \leq \epsilon_2^{-1} \quad \text{for all } u \qquad (5.5\text{-}9)$$

Note the analogy with the last section, where with R diagonal we concluded that independent linear gain variations between $\tfrac{1}{2}$ and ∞ on each scalar input do not destroy stability.

Let us check (5.5-7).

$$2x'KR\Phi(K'x) = \sum_i 2r_i(K'x)_i\varphi_i[(K'x)_i]$$

$$\geq \sum_i 2r_i(\tfrac{1}{2} + \epsilon_1)(K'x)_i^2$$

$$= (1 + 2\epsilon_1)x'KRK'x$$

The rest of the proof of stability, of course, proceeds as before.

For the time-varying $\Phi(\cdot)$ case, because $\dot{V} \leq 0$, the proof of asymptotic stability as opposed just to stability is awkward; one cannot appeal in the time-varying case to the fact that $\dot{V} \equiv 0$ implies $x \equiv 0$. Problem 5.5-2 deals with time-varying $\Phi(\cdot)$.

Main points of the section. Nonlinear, in fact time-varying, input gains that are sector restricted to be strictly inside a sector bounded by slopes $\tfrac{1}{2}$ and ∞ do not destroy the stability of a single-input optimal system. A multivariable result is also available which has an attractive interpretation in case R is diagonal.

Problem 5.5-1. Assume a regulator is designed to have degree of stability α, by the inclusion of the exponential weighting $e^{2\alpha t}$ in the performance index integrand. Show that nonlinearities of the type discussed in this section allow retention of the degree of stability α. [*Hint:* Show first that $\dot{V}/V \leq -2\alpha$.]

Problem 5.5-2. Extend the Lyapunov-based proof to the case of time-varying $\varphi(\cdot)$ for single-input systems. Do this as follows:

(i) From $V > 0$ and $\dot{V} \le 0$, show that $k'x$ and $\varphi(k'x)$ are square integrable.

(ii) Show that if $\dot{x} = Ax + bu$ with Re $\lambda_i(A)$ stable and u is square integrable then $\|x(t)\| \to 0$.

(iii) Use the fact that $\dot{x} = (F + gk')x + gv$ where v is square integrable.

5.6 THE INVERSE OPTIMAL CONTROL PROBLEM†

The inverse optimal control is easily stated. Given a triple $\{F, G, K\}$, does K have the property that it is the optimal control law associated with F, G, some non-negative symmetric Q, and some positive definite R?

In this section, we shall give a reasonably complete answer to this question. Relevant references include [11, 12, 13].

What assumption is it reasonable to make? Certainly, stabilizability of $[F, G]$ and stability of $F + GK'$. Also, we know that a return difference equality necessarily holds, and thus it is reasonable in the inverse problem to postulate its satisfaction. For scalar systems, this implies that for all ω

$$|1 - k'(j\omega I - F)^{-1}g| \ge 1 \tag{5.6-1}$$

and for multivariable systems, it implies that for some positive definite R,

$$\underline{\sigma}\{R^{1/2}[I - K'(j\omega I - F)^{-1}G]R^{-1/2}\} \ge 1 \tag{5.6-2}$$

Note that (5.6-2) is equivalent to

$$(R^{-1/2})'[I - G'(-j\omega I - F')^{-1}K]R[I - K'(j\omega I - F)^{-1}G]R^{-1/2} \ge I$$

or

$$[I - G'(-j\omega I - F')^{-1}K]R[I - K'(j\omega I - F)^{-1}G] \ge R \tag{5.6-3}$$

It is much easier to postulate that R is part of the data than to regard it, as well as Q, as unknown. In the scalar case, this is a costless assumption. Now with the assumptions that (1) $\{F, G, K, R\}$ are known, (2) $[F, G]$ is stabilizable, (3) $F + GK'$ is stable, (4) the return difference inequality (5.6-3) holds, the crucial question becomes: Does there exist $Q = DD' \ge 0$ with $[F, D]$ detectable such that the associated optimal control law is K? Equivalently, does there exist $Q = DD'$ with $[F, D]$ detectable such that

$$[1 - g'(-j\omega I - F')^{-1}k][1 - k'(j\omega I - F)^{-1}g]$$
$$= 1 + g'(-j\omega I - F')^{-1}Q(j\omega I - F)^{-1}g \tag{5.6-4}$$

or

$$[I - G'(-j\omega I - F')^{-1}K]R[I - K'(j\omega I - F)^{-1}G]$$
$$= R + G'(-j\omega I - F')^{-1}Q(j\omega I - F)^{-1}G \tag{5.6-5}$$

---

† This section may be omitted without loss of continuity.

and such that the K we started with is the solution of the optimal control problem defined by F, G, Q, and R.

The first question is relatively easy to answer. For the second question, it turns out that the best we can achieve, given stabilizability of $[F, G]$ rather than controllability, is to show that $\overline{K}'(j\omega I - F')^{-1}G = K'(j\omega I - F')^{-1}G$, where $\overline{K}'$ is the optimal gain. Of course, this is as much as one could hope for, and is a satisfactory resolution of the problem. See also Problem 5.6-1.

We shall describe first the construction of Q for the single-input case. Let $|sI - F| = p_0(s)$ and $|sI - F - GK'| = p_c(s)$, which is a stable polynomial. Assume also that $[F, G]$ is controllable, not just stabilizable. Then (5.6-1) implies that, via (5.2-20),

$$\frac{p_c(-j\omega)p_c(j\omega)}{p_0(-j\omega)p_0(j\omega)} \geq 1$$

whence

$$p_c(-j\omega)p_c(j\omega) - p_0(-j\omega)p_0(j\omega) \geq 0 \qquad (5.6\text{-}6)$$

The quantity on the left is the evaluation on the $j\omega$-axis of a polynomial, call it $e(s)$, for which $e(s) = e(-s)$. Such a polynomial necessarily has zeros that are symmetrically located with respect to the $j\omega$-axis; that is, one can write $e(s) = \pm m(s)m(-s)$ for some polynomial $m(s)$ with real coefficients and with all roots in Re $[s] \leq 0$. A polynomial factorization procedure gives us $m(s)$. Taking into account that $e(j\omega) \geq 0$ for all ω, we see that the plus sign applies. Then

$$e(j\omega) = p_c(-j\omega)p_c(j\omega) - p_0(-j\omega)p_0(j\omega) = m(-j\omega)m(j\omega) \qquad (5.6\text{-}7)$$

from which

$$\frac{p_c(-j\omega)p_c(j\omega)}{p_0(-j\omega)p_0(j\omega)} = 1 + \frac{m(-j\omega)m(j\omega)}{p_0(-j\omega)p_0(j\omega)} \qquad (5.6\text{-}8)$$

Now p_c and p_0 are both monic, of degree n if $n = \dim F$. So $e(s)$ has degree at most $2n - 2$, and $m(s)$ has degree at most $(n - 1)$. Hence $m(s)/p_0(s)$ is strictly proper. Accordingly, there exists an n-vector d such that

$$\frac{m(s)}{p_0(s)} = d'(sI - F)^{-1}g \qquad (5.6\text{-}9)$$

And now with $Q = dd'$, we recover (5.6-4) from (5.6-8). Is $[F, d]$ detectable? Lack of detectability would mean that m and p_0 have one or more common zeros in Re $[s] \geq 0$. Since $m(s)$ has all zeros in Re $[s] \leq 0$, we see that lack of detectability implies for some ω_0 that $p_0(j\omega_0) = m(j\omega_0) = 0$ and then from (5.6-8) $p_c(-j\omega_0)p_c(j\omega_0) = 0$. Since $p_c(-j\omega_0) = p_c^*(j\omega_0)$, this means that $p_c(j\omega_0) = 0$, which contradicts the assumption that $F + GK'$ has all negative real part eigenvalues, or equivalently, $p_c(s)$ has all zeros in Re $[s] < 0$.

In the multivariable case, a similar argument can be given, if we agree to represent transfer function matrices using polynomial matrix fraction descriptions.

Associated with any controllable pair $[F, G]$ there exists a polynomial matrix $A(s)$ with det $A(s) = $ det $(sI - F)$ and a bijection $H \leftrightarrow B_H(s)$ (with $B_H(s)$ polynomial) such that

$$H'(sI - F)^{-1}G = B_H(s)A^{-1}(s) \qquad (5.6\text{-}10)$$

We represent $I - K'(sI - F)^{-1}G$ as $[A(s) - B_K(s)]A^{-1}(s)$, and infer from (5.6-3) a matrix polynomial inequality analogous to (5.6-6). We can then appeal to a factorization theorem [14] to conclude a matrix version of (5.6-7). Analogues to (5.6-8) and (5.6-9) follow straightforwardly, and finally the detectability can be checked.

In the above discussion, we assumed that $[F, G]$ was controllable, not stabilizable. The argument can be extended to cover this last case. Problem 5.6-1 sets up the argument for the single-input case.

To conclude the argument, we need to show that if $[F, G]$ is stabilizable and Re $\lambda_i(F + GK') < 0$, and if (5.6-5) is satisfied, then $K'(j\omega I - F)^{-1}G = \bar{K}'(j\omega I - F)^{-1}G$, where $\bar{K}'$ is the optimal control gain associated with F, G, Q, and R. Now $\bar{K}$ appears in a return difference equality by virtue of its optimality, and so

$$[I - G'(-j\omega I - F')^{-1}K]R[I - K'(j\omega I - F)^{-1}G]$$
$$= [I - G'(-j\omega I - F')^{-1}\bar{K}]R[I - \bar{K}'(j\omega I - F)^{-1}G]$$

whence

$$R[I - K'(j\omega I - F)^{-1}G][I - \bar{K}'(j\omega I - F)^{-1}G]^{-1}$$
$$= [I - G'(-j\omega I - F')^{-1}K]^{-1}[I - G'(-j\omega I - F')^{-1}\bar{K}]R$$

or, after some manipulation,

$$R[I + (\bar{K}' - K')(j\omega I - F - G\bar{K}')^{-1}G] = [I - G'(-j\omega I - F' - KG')^{-1}(\bar{K} - K)]R$$

The left side is a transfer function matrix with all poles in Re $(s) < 0$, the right side with all poles in Re $[s] > 0$. Hence each side is a constant, obtainable by setting $\omega = \infty$:

$$R + (\bar{K}' - K')(j\omega I - F - G\bar{K}')^{-1}G = R \qquad (5.6\text{-}11)$$

If $[F, G]$ is completely controllable, (5.6-11) implies that $K = \bar{K}$. Otherwise, it implies merely that $K'(sI - F)^{-1}G = \bar{K}'(sI - F)^{-1}G$.

Let us summarize what we have achieved:

> *Inverse optimal control problem.*   Suppose a quadruple $\{F, G, K, R\}$ satisfies (1) $[F, G]$ is stabilizable; (2) Re $\lambda_i(F + GK') < 0$; (3) the return difference inequality (5.6-3) holds. Then there exists $Q = DD'$ with $[F, D]$ detectable such that the optimal control gain $\bar{K}'$ associated with the optimal control problem $\{F, G, Q, R\}$ satisfies $\bar{K}'(j\omega I - F)^{-1}G = K'(j\omega I - F)^{-1}G$.

In the multivariable case, one can examine a more complicated problem. Given F, G, K, one can ask if there exists a positive definite R such that (5.6-3)

holds. There does not seem a tidy answer to this question. Problem 5.6-2 states a necessary condition.

Main points of the section.   Under assumptions of stabilizability of $[F, G]$ and stability of $F + GK'$, satisfaction of a return difference inequality is not just necessary but in practical terms sufficient for K to be optimal.

Problem 5.6-1.   Suppose that (5.6-1) holds with $[F, g]$ stabilizable and $F + gk'$ stable. Suppose further that

$$F = \begin{bmatrix} F_{11} & F_{12} \\ 0 & F_{22} \end{bmatrix} \qquad g = \begin{bmatrix} g_1 \\ 0 \end{bmatrix} \qquad k = \begin{bmatrix} k_1 \\ k_2 \end{bmatrix}$$

with Re $\lambda_i(F_{22}) < 0$. Observe that $k'(sI - F)^{-1}g = \bar{k}'(sI - F)^{-1}g$ where $\bar{k}' = [k_1'\ 0]$, and $F + g\bar{k}'$ is stable. Verify that there exists d with $[F, d]$ detectable such that $\bar{k}$ is the optimal control law associated with $F, g, Q = dd'$ and $r = 1$.

Problem 5.6-2.   Show that a necessary condition for (5.6-3) to hold for some $R = R' > 0$ is that

$$|\det [I - K'(j\omega I - F)^{-1}G]| \geq 1$$

for all ω.

Problem 5.6-3.   Consider two optimal control problems parametrized by F, G, Q_1, R, and F, G, Q_2, R with all the usual restrictions satisfied. Suppose further that $[F, G]$ is completely controllable, and

$$G'(-sI - F')^{-1}Q_1(sI - F)^{-1}G = G'(-sI - F')^{-1}Q_2(sI - F)^{-1}G$$

even though $Q_1 \neq Q_2$. Show that the optimal control laws for the two problems are the same. Is it possible for the optimal performance indices to be the same for all x_0?

5.7 RETURN DIFFERENCE EQUALITY FOR DISCRETE-TIME REGULATORS

In this section, we shall present the return difference equality for the discrete-time regulator, and explain some variations that must be made to the continuous-time robustness results (e.g., gain margin) which follow. Recall that for a problem parametrized by F, G, Q, R, the optimal performance index is $x'(t_0)\bar{P}x(t_0)$ with $\bar{P} = \bar{S} - Q$, and $\bar{S}$ satisfying

$$\bar{S} = F'\{\bar{S} - \bar{S}G[G'\bar{S}G + R]^{-1}G'\bar{S}\}F + Q \qquad (5.7\text{-}1)$$

The optimal control is $u^* = K'x$ when

$$K' = -[G'\bar{S}G + R]^{-1}G'\bar{S}F \qquad (5.7\text{-}2)$$

Rewrite (5.7-1) as

$$\bar{S} - F'\bar{S}F + F'\bar{S}G[G'\bar{S}G + R]^{-1}G'\bar{S}F = Q$$

or

$$(z^{-1}I - F)'\bar{S}(zI - F) + (z^{-1}I - F')\bar{S}F + F'\bar{S}(zI - F)$$
$$+ F'\bar{S}G[G'\bar{S}G + R]^{-1}G'\bar{S}F = Q$$

Now premultiply by $G'(z^{-1}I - F')^{-1}$ and postmultiply by $(zI - F)^{-1}G$. There results

$$G'\bar{S}G + G'\bar{S}F(zI - F)^{-1}G + G'(z^{-1}I - F')^{-1}F'\bar{S}G$$
$$+ G'(z^{-1}I - F')^{-1}F'\bar{S}G[G'\bar{S}G + R]^{-1}G'\bar{S}F(zI - F)^{-1}G$$
$$= G'(z^{-1}I - F')^{-1}Q(zI - F)^{-1}G$$

Now use (5.7-2):

$$G'\bar{S}G - [G'\bar{S}G + R]K'(zI - F)^{-1}G - G'(z^{-1}I - F')^{-1}K[G'\bar{S}G + R]$$
$$+ G'(z^{-1}I - F')^{-1}K[G'\bar{S}G + R]K'(zI - F)^{-1}G$$
$$= G'(z^{-1}I - F')^{-1}Q(zI - F)^{-1}G$$

or

$$[I - G'(z^{-1}I - F')^{-1}K][G'\bar{S}G + R][I - K'(zI - F)^{-1}G]$$
$$= R + G'(z^{-1}I - F')^{-1}Q(zI - F)^{-1}G \qquad (5.7\text{-}3)$$

This is the sought return difference equality. There are two differences with the continuous-time result: $-s$ is replaced by z^{-1}, which is not a surprise, and R on the left side is replaced by $G'\bar{S}G + R$, but not on the right side. This perhaps is a surprise.

From (5.7-3), it is easy to follow the continuous-time argument to establish an equation involving the closed-loop characteristic polynomial. Suppose $p_c(z) = \det (zI - F - GK')$ and $p_0(z) = \det (zI - F)$. Then

$$p_c(z^{-1})p_c(z) = \frac{p_0(z^{-1})p_0(z)\det [R + G'(z^{-1}I - F')^{-1}Q(zI - F)^{-1}G]}{\det [G'\bar{S}G + R]} \qquad (5.7\text{-}4)$$

Without knowing $\bar{S}$, we can find $p_c(z)$ in the following way. Evaluate the numerator on the right side of (5.7-4). It will be a polynomial in z and z^{-1} with the property that if z_0 is a root, so is z_0^{-1}. Find all roots z_i with $|z_i| < 1$. Then $p_c(z) = \Pi(z - z_i)$.

In the continuous-time case, the return difference equality is used to infer robustness results, especially those concerned with phase and gain margin and tolerance of sector nonlinearities at plant inputs. The same can be done in discrete time, but some important differences occur. For example, let us consider the gain margin property. In a continuous-time problem, when the optimal gain K is replaced by βK and $\beta \to \infty$, all closed-loop eigenvalues remain stable, and one or

more tend to infinity. In discrete time, similarly if K is replaced by βK and $\beta \to \infty$, one or more of the zeros of $I - K'(zI - F)^{-1}G$, which are the closed-loop eigenvalues, must tend to infinity. But in discrete time, stability is equivalent to having all closed-loop eigenvalues inside the unit circle. This means that an infinite gain margin result cannot be expected.

The lack of parallel is a consequence of an immediately observable difference in the continuous-time and discrete-time return difference equalities. Consider a single-input problem, with $R = 1$. Then we have for continuous and discrete time

$$|1 - k'(j\omega I - F)^{-1}g| \geq 1 \tag{5.7-5}$$

and, using (5.7-3),

$$|1 - k'(e^{j\omega}I - F)^{-1}g| \geq \frac{1}{(g'\bar{S}g + 1)^{1/2}} \tag{5.7-6}$$

Because $g'\bar{S}g > 0$, this means that the Nyquist plot of $k'(e^{j\omega}I - F)^{-1}g$ avoids a smaller circle centered at the -1 point than does the Nyquist plot of $k'(j\omega I - F)^{-1}g$. In particular, the Nyquist plot of $k'(e^{j\omega}I - F)^{-1}g$ must avoid the interval $(-1 - \gamma, -1 + \gamma)$ where

$$\gamma = (g'\bar{S}g + 1)^{-1/2} < 1 \tag{5.7-7}$$

and the gain margin has a lower limit of $(1 + \gamma)^{-1}$ and an upper limit of $(1 - \gamma)^{-1}$. (See Fig. 5.7-1). The point A on the circle boundary is at unit distance from the
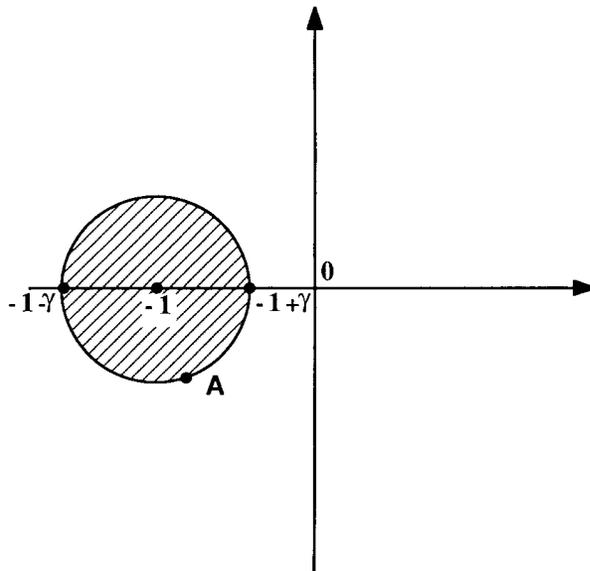
Figure 5.7-1   Circle of center $-1 + j0$ and radius $\gamma < 1$ avoided by Nyquist plot of $k'(\exp(j\omega)I - F)^{-1}g$.

origin, and the acute angle between AO and the real axis defines the phase margin. This is evidently $2 \sin^{-1} (\gamma/2)$, and is smaller than 60 deg.

Most of these ideas can be found in [15], which points out that there exist first-order systems for which the guaranteed margins are arbitrarily small. In [16], a multivariable version of these results is discussed. More complicated formulas, involving Q, R, F, and so forth are used to define the equivalent of γ above.

Discrete-time plant models very frequently arise through discretization of a continuous-time model. The discretization interval can be chosen by the designer. It turns out that as the discretization interval approaches zero and the adjustments to F, G, Q, and R are made to maintain the appropriate relationship with the continuous-time model, the quantity $\gamma \to 1$; that is, the continuous-time results are recovered. Of course, this is not altogether surprising.

Main points of the section.   A discrete-time return difference equality can be found. From the equality, robustness results, including phase and gain margin bounds, can be derived. These are not as attractive as in continuous-time, but this is to be expected, since no discrete-time system could have an infinite gain margin.

Problem 5.7-1.   Suppose that F is nonsingular. In the discrete-time regulator, it is possible to construct a discrete-time Hamiltonian matrix

$$M = \begin{bmatrix} F + GR^{-1}G'(F^{-1})'Q & -GR^{-1}G'(F^{-1})' \\ -(F^{-1})'Q & (F^{-1})' \end{bmatrix}$$

Show that the characteristic polynomial of M is, to within multiplication by a scaling constant, and powers of z,

$$\det(zI - F)\det(z^{-1}I - F')\det[R + G'(z^{-1}I - F')^{-1}Q(zI - F)^{-1}G]$$

Hint: An intermediate result is

$$\det(zI - M) = \det \begin{bmatrix} I & 0 \\ 0 & (F^{-1})' \end{bmatrix} \det \begin{bmatrix} zI - F & GR^{-1}G' \\ Q & F' - z^{-1}I \end{bmatrix} \det \begin{bmatrix} I & 0 \\ 0 & zI \end{bmatrix}$$

REFERENCES

[1] B. D. O. Anderson and S. Vongpanitlerd, *Network Analysis and Synthesis*. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1973.

[2] C. A. Desoer and M. Vidyasagar, *Feedback Systems: Input-Output Properties*. New York: Academic Press, 1975.

[3] J. C. Doyle and G. Stein, "Multivariable Feedback Design: Concepts for a Classical/ Modern Synthesis," *IEEE Trans. Auto. Control*, Vol. AC-26, No. 1 (February 1981), pp. 4–16.

[4] J. B. Cruz and W. R. Perkins, "A New Approach to the Sensitivity Problem in Multi-

variable Feedback System Design," *IEEE Trans. Auto. Control*, Vol. AC-9, No. 4 (July 1964), pp. 216–223.

[5] W. R. Perkins and J. B. Cruz, Jr., "Feedback Properties of Linear Regulators," *IEEE Trans. Auto. Control*, AC-16, No. 6 (special issue on LQG problem) (Dec. 1971) pp. 659–664.

[6] J. B. Cruz, Jr., J. S. Freudenberg, and D. P. Looze, "A Relationship Between Sensitivity and Stability of Multivariable Feedback Systems," *IEEE Trans. Auto. Control*, Vol. AC-26, No. 1 (special issue on linear multivariable control systems) (Feb. 1981), pp. 66–74.

[7] J. S. Freudenberg and D. P. Looze, "Right Half Plane Zeros and Design Tradeoffs in Feedback Systems," *IEEE Trans. Auto. Control*, Vol. AC-30, No. 6 (June 1985), pp. 555–565.

[8] S. D. O'Young and B. A. Francis, "Sensitivity Tradeoffs for Multivariable Plants," *IEEE Trans. Auto. Control*, Vol. AC-30, No. 7 (July 1985), pp. 625–632.

[9] M. J. Grimble and T. J. Owens, "On Improving the Robustness of LQ Regulators," *IEEE Trans. Auto. Control*, Vol. AC-31, No. 1 (January 1986), pp. 54–55.

[10] N. A. Lehtomaki, N. R. Sandell, Jr., and M. Athans, "Robustness Results in Linear-Quadratic Gaussian Based Multivariable Control" *IEEE Trans. Auto. Control,* Vol. AC-26, No. 1 (special issue on linear multivariable control) (Feb. 1981), pp. 75–92.

[11] R. E. Kalman, "When Is a Linear Control System Optimal?" *Trans. ASME Ser. D: J. Basic Eng.*, Vol. 86 (March 1964), pp. 1–10.

[12] B. D. O. Anderson, "The Inverse Problem of Optimal Control," *Rep. No. SEL-66-038 (TR No. 6560-3)*, May 1966, Stanford Electronics Laboratories, Stanford, California.

[13] T. Fujii and M. Narazaki, "Complete Optimality Conditions in the Inverse Problem of Optimal Control," *SIAM J. Control and Optimization*, Vol. 22, No. 2 (March 1984), pp. 327–341.

[14] D. C. Youla, "On the Factorization of Rational Matrices," *IRE Trans. Information Theory*, Vol. IT-7, No. 3 (July 1961), pp. 172–189.

[15] J. L. Willems and H. Van de Voorde, "The Return Difference for Discrete-Time Optimal Feedback Systems," *Automatica*, Vol. 14 (1978), pp. 511–513.

[16] U. Shaked, "Guaranteed Stability Margins for the Discrete-time Linear Quadratic Optimal Regulator," *IEEE Trans. Auto. Control*, Vol. AC-31, No. 2 (February 1986), pp. 162–165.

6

Asymptotic Properties and Quadratic Weight Selection

6.1 SINGLE INPUT SYSTEMS

In this chapter, we shall address many of the factors associated with quadratic weight selection, some of which were foreshadowed by consideration of the scalar example of Chapter 3. A great many of the ideas are most readily approached by considering the simpler case of single-input systems, so we begin with this case. It is very revealing to study situations in which the ratio of state to control weighting is very large or small, so much of our attention will be given to these situations.

Low state weighting. Consider a collection of optimization problems parametrized by F, g, $r = 1$, and ρQ, where ρ is variable. (We shall be especially interested in $\rho \to 0$.) Of course, we assume the usual stabilizability and detectability conditions. Also, there is no loss of generality in taking $r = 1$. (Why not?)

If the open-loop plant is stable, the control $u \equiv 0$ will incur a cost

$$\rho x_0' \int_0^\infty e^{F't} Q e^{Ft} \, dt \, x_0$$

which tends to zero as $\rho \to 0$. So we might expect that the control gain $k \to 0$ as $\rho \to 0$. On the other hand, if the plant is unstable, we could not expect $k \to 0$, for

then the closed-loop would be unstable for suitably small $\|k\|$. To understand what happens, consider the return difference equation

$$[1 - g'(-sI - F')^{-1}k_\rho][1 - k_\rho'(sI - F)^{-1}g]$$
$$= 1 + \rho g'(-sI - F')^{-1}Q(sI - F)^{-1}g \qquad (6.1\text{-}1)$$

With $p_0(s)$ and $p_{c\rho}(s)$ denoting the open-loop and closed-loop characteristic polynomials, we know that for some polynomial $q(s)$, even in s and nonnegative for $s = j\omega$,

$$p_{c\rho}(-s)p_{c\rho}(s) = p_0(-s)p_0(s) + \rho q(s) \qquad (6.1\text{-}2)$$

This is the relevant version of (5.2-19). Clearly, as $\rho \to 0$,

$$p_{c\rho}(-s)p_{c\rho}(s) \to p_0(-s)p_0(s) \qquad (6.1\text{-}3)$$

Suppose no zero of $p_0(s)$ lies on the $j\omega$-axis. Then as $\rho \to 0$, the zeros of $p_{c\rho}(s)$, which must be stable, approach the stable zeros of $p_0(s)$ and the reflection through the $j\omega$-axis of the unstable zeros of $p_0(s)$. If $p_0(s)$ has a $j\omega$-axis zero, then $p_{c\rho}(s)$ has a zero that approaches the zero of $p_0(s)$ on the $j\omega$-axis, from the left half plane. Now

$$1 - k_\rho'(sI - F)^{-1}g = \frac{p_{c\rho}(s)}{p_0(s)} \qquad (6.1\text{-}4)$$

If $p_{c\rho}(s) \to p_0(s)$, then $k_\rho'(sI - F)^{-1}g \to 0$; that is, $k_\rho \to 0$. Otherwise, k_ρ approaches a nonzero quantity, call it k_0. Clearly, k_0 is independent of the particular Q used in defining k_ρ for $\rho \neq 0$, since $p_0(s)$ is independent of Q.

High state weighting.

Now our interest is in letting $\rho \to \infty$. It is a little easier to see what happens if we make the restriction that $Q = dd'$ for some vector d. Then if we identify

$$d'(sI - F)^{-1}g = \frac{m(s)}{p_0(s)} \qquad (6.1\text{-}5)$$

with degree $m(s) <$ degree $p_0(s)$, (6.1-2) is replaced by

$$p_{c\rho}(-s)p_{c\rho}(s) = p_0(-s)p_0(s) + \rho m(-s)m(s) \qquad (6.1\text{-}6)$$

In case we do not have $Q = dd'$, the polynomial $q(s)$, being even in s and nonnegative for $s = j\omega$, necessarily has a factorization $q(s) = m(-s)m(s)$. The choice $Q = dd'$ simply makes the origin of $m(j\omega)$ more transparent.

Suppose that degree $m(s) = l$ and degree $p_0(s) = n$. It is immediately clear that any zero of the left side which remains finite as $\rho \to \infty$ must approach a zero of $m(-s)m(s)$. Hence any zero of $p_{c\rho}(s)$ which remains finite as $\rho \to \infty$ must approach a stable zero of $m(-s)m(s)$. Since there are only $2l$ zeros of $m(-s)m(s)$, we must expect the remaining $2(n - l)$ zeros of $p_{c\rho}(-s)p_{c\rho}(s)$ to become infinite. Suppose that $m(-s)m(s) = (-1)^l \alpha s^{2l} +$ lower-order terms; note that $\alpha > 0$, and that $p_0(-s)p_0(s) = (-1)^n s^{2n} +$ lower-order terms. It follows that the zeros of

$p_{cp}(-s)p_{cp}(s)$ which tend to infinity must also tend to the nonzero roots of the equation

$$(-1)^n s^{2n} + (-1)^l \rho\alpha s^{2l} = 0$$

or of

$$s^{2(n-l)} = (-1)^{l-n+1}\rho\alpha \qquad (6.1\text{-}7)$$

In summary then, as $\rho \to \infty$, $2l$ zeros of $p_0(-s)p_0(s) + \rho m(-s)m(s)$ approach the zeros of $m(-s)m(s)$ and $2(n-l)$ zeros approach the roots of (6.1-7). The zeros of $p_{cp}(s)$ tend to the l zeros of $m(-s)m(s)$ which have negative real parts, and the left half plane roots of (6.1-7). The latter zeros lies in a pattern on a circle of radius $(\rho\alpha)^{1/(2n-2l)}$, which network theory terms a Butterworth configuration [1]. The phase angles are given in the following table.

$n - l = 1$	$+180°$
$n - l = 2$	$\pm135°$
$n - l = 3$	$\pm120°, +180°$
$n = l = 4$	$\pm112.5°, \pm157.5°$
etc.	

Of course, if d is chosen so that $d'(sI - F)^{-1}g$ is minimum phase, then the zeros of $m(-s)m(s)$ that are relevant are the zeros of $m(s)$.

In Figure 6.1-1 we illustrate the effect of changing ρ from a very small value through to a very large value. The example has

$$d'(sI - F)^{-1}g = \frac{s + a}{s^2 + 2\zeta s + 1}$$

with the four cases

$$
\begin{array}{ll}
a = \ \ 0.5 & \zeta = \ \ 0.1 \\
a = \ \ 0.5 & \zeta = -0.1 \\
a = -0.5 & \zeta = \ \ 0.1 \\
a = -0.5 & \zeta = -0.1
\end{array}
$$

Figure 6.1-1 displays the root locus for the closed-loop regulator poles as ρ varies from 0 to 1000. The root locus is the same for all four cases. (Why?) Notice that for small ρ, the poles in all cases approximate those of $s^2 + 0.2s + 1$, and for large ρ, one approximates -0.5.

Pole positioning and loop gain setting via high state weighting.   High state weighting evidently provides a technique for positioning closed-loop poles. First, one can control up to $(n - 1)$ poles of the closed-loop system, by choosing $Q = dd'$ with the zeros of $d'(sI - F)^{-1}g$ coinciding with these $(n - 1)$ poles. Then one can choose ρ to move the remaining pole towards infinity. Of course, the smaller ρ is, the less accurate will be the pole positioning.
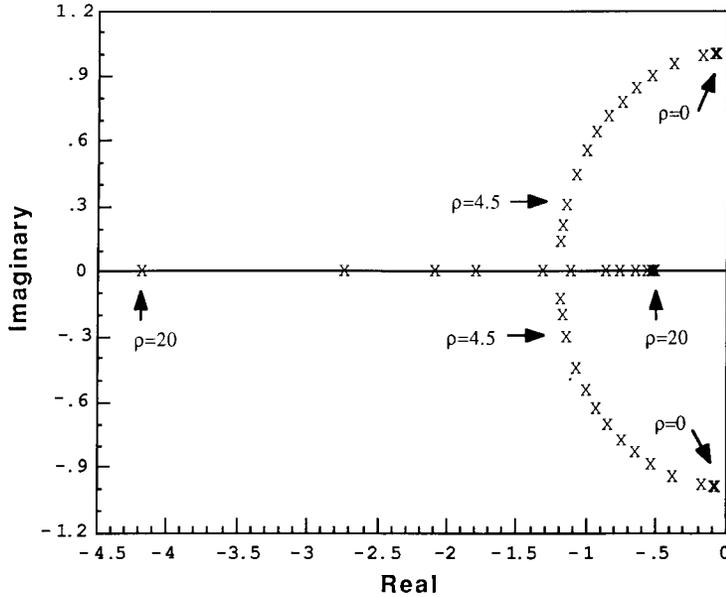
Figure 6.1-1   Root locus as ρ is varied.

In the high state weighting case, we can obtain an approximate expression for k_ρ when $d'(sI - F)^{-1}g$ has left plane zeros and $[F, g]$ is controllable, rather than just stabilizable. [The assumption that $[F, d]$ is detectable remains, so that if there are pole-zero cancellations in forming $d'(sI - F)^{-1}g$, they are in Re $[s] < 0$]. The approximate expression is

$$k_\rho = \pm\sqrt{\rho}\,d \qquad (6.1\text{-}8)$$

(with $\pm \rho^{-1/2}\, k_\rho \to d$ as $\rho \to \infty$) In case $d'g \neq 0$, the plus sign applies if $d'g < 0$, and the minus sign if $d'g > 0$. Otherwise, the sign is determined by the requirement that the zeros of $1 - k'_\rho(sI - F)^{-1}g$ must be stable.

This result comes about as follows. From the return difference equality, we have

$$|1 - k'_\rho(j\omega I - F)^{-1}g| = [1 + \rho|d'(j\omega I - F)^{-1}g|^2]^{1/2}$$

It follows as $\rho \to \infty$ that for $\omega \neq \infty$ (so that $|d'(j\omega I - F)^{-1}g| \neq 0$),

$$|1 - k'_\rho(j\omega I - F)^{-1}g| \simeq \rho^{1/2}\,|d'(j\omega I - F)^{-1}g|$$

Since this can only hold if $|k'_\rho(j\omega I - F)^{-1}g|$ grows as fast as $\rho^{1/2}$, the following further approximation is valid:

$$|k'_\rho(j\omega I - F)^{-1}g| \simeq \rho^{1/2}\,|d'(j\omega I - F)^{-1}g| \qquad (6.1\text{-}9)$$

This equality forces the zeros of $k'_\rho(sI - F)^{-1}g$ and $d'(sI - F)^{-1}g$ to be related (either

equal or reflections through the imaginary axis of one another). We now must show that all zeros of one coincide with all zeros of the other as $\rho \to \infty$.

Recall from the last chapter that all zeros of $k_\rho'(sI - F)^{-1}g$ must necessarily be in Re $[s] \leq 0$; that is, $k_\rho'(sI - F)^{-1}g$ is minimum phase. Now $d'(sI - F)^{-1}g$ is also minimum phase. In view of (6.1-9) and the fact that both $k_\rho'(j\omega I - F)^{-1}g$ and $d'(sI - F)^{-1}g$ are minimum phase, it follows that both transfer functions agree, to within a multiple $\pm\sqrt{\rho}$. Then (6.1-8) follows, when we appeal to controllability.

Cross-over frequency and closed-loop bandwidth
setting.   With ρ large, it is also possible to obtain an approximate value for the cross-over frequency of the loop gain $k_\rho'(j\omega I - F)^{-1}g$. At high frequencies, there holds

$$|k_\rho'(j\omega I - F)^{-1}g| \simeq \frac{k_\rho'g}{\omega}$$

and so the loop gain is approximately 1 where

$$\omega_1 = k_\rho'g = \sqrt{\rho}\,|d'g| \tag{6.1-10}$$

Equations (6.1-8) and (6.1-10) provide information about shaping of the loop gain $k'(sI - F)^{-1}g$ in an optimal design. Choose d so that $d'(sI - F)^{-1}g$ has the desirable shape, apart from a multiplicative gain factor. Then set ρ by invoking a specification regarding the cross-over frequency, see (6.1-10).

By way of example, consider the control of a simple harmonic oscillator

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

Suppose we wish to have a single dominant pole at $s = -1 + j0$. Suppose also that we seek (and can tolerate) a cross-over frequency of about 10 rad/sec. (Thus unstructured uncertainty ought all be confined to $|\omega| > 10$.)

We take

$$d'(sI - F)^{-1}g = \frac{s + 1}{s^2 + 1}$$

which implies that $d' = [1 \quad 1]$. Then $d'g = 1$, and (6.1-10) with $\omega_1 = 10$ leads to $\rho = 100$. For convenience, we seek to minimize the index

$$V = \int_0^\infty [u^2 + 99(x_1 + x_2)^2]\, dt$$

The return difference equality (in its polynomial form) yields here

$$\begin{aligned}
p_c(-s)p_c(s) &= (s^2 + 1)(s^2 + 1) + 99(-s + 1)(s + 1) \\
&= s^4 - 97s^2 + 100 \\
&= s^4 + 20s^2 + 100 - 117s^2 \\
&= (s^2 + \sqrt{117}s + 10)(s^2 - \sqrt{117}s + 10)
\end{aligned}$$

Hence

$$1 - k'(sI - F)^{-1}g = \frac{s^2 + \sqrt{117}s + 10}{s^2 + 1}$$

from which we derive

$$k' = [-9 \quad -\sqrt{117}]$$

The actual closed-loop poles are -1.02 and -9.80. The loop gain $|k'(j\omega I - F)^{-1}g|$ is plotted in Fig. 6.1-2. The cross-over frequency is close to 10 rad/s.

It is also possible to exercise some control over the closed-loop system bandwidth by choice of ρ. The closed-loop transfer function is defined as

$$w_c(s) = k_\rho'(sI - F - gk_\rho')^{-1}g \qquad (6.1\text{-}11)$$

and its bandwidth may be defined (somewhat arbitrarily) by that ω_0 for which

$$|w_c(j\omega_0)|^2 = \tfrac{1}{2}|w_c(0)|^2 \qquad (6.1\text{-}12)$$

We assert that ω_0 is approximately determined as that frequency for which

$$|d'(j\omega_0 I - F)^{-1}g|^2 = \frac{1}{\rho + (2/|d'F^{-1}g|^2)} \quad (F \text{ nonsingular})$$

$$= \rho^{-1} \qquad\qquad\qquad (F \text{ singular}) \qquad (6.1\text{-}13)$$
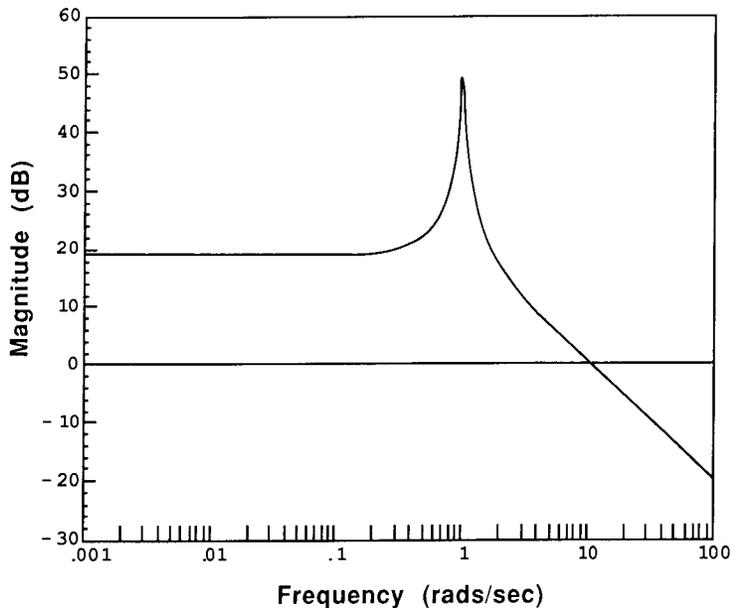
Figure 6.1-2 Bode plot of $k'(sI - F)^{-1}g$.

Proof of this result is left to the problems. In the example above, choice of $\omega_0 = 10$ rad/s leads to $\rho \approx 100$. The closed-loop frequency response is depicted in Fig. 6.1-3 and exhibits a bandwidth (3 dB cut-off frequency) of approximately 12 rad/s.

The cross-over frequency and the closed-loop bandwidth are not independent. For a first-order system $\alpha(s + \beta)^{-1}$ with associated closed-loop transfer function $\alpha(s + \beta + \alpha)^{-1}$, the open-loop cross-over frequency is $\omega_1 = \sqrt{\alpha^2 - \beta^2}$ while the closed-loop 3-dB frequency is $\alpha + \beta$. These will be approximately the same when α is large. Similarly, as a rough approximation (to be checked after a design), one could assume that with ρ large, there holds $\omega_1 = \omega_0$.

Design insights.   Some remarks pertinent to design problems are now made.

1. If a d can be selected to achieve (modulo scaling) a desirable open-loop transfer function $d'(sI - F)^{-1}g$ with left half-plane zeros, then the performance index parameter Q can be taken as $\rho dd'$, with ρ selected as in a subsequent remark.

2. A simpler approach than in the above remark is to merely select some d_1 such that $d_1'(sI - F)^{-1}g$ has a desirable magnitude response. The zeros of $d_1'(sI - F)^{-1}g$ and its phase response can be ignored, since we know that quadratic designs with $Q = \rho d_1 d_1'$ give attractive phase margins, irrespective of
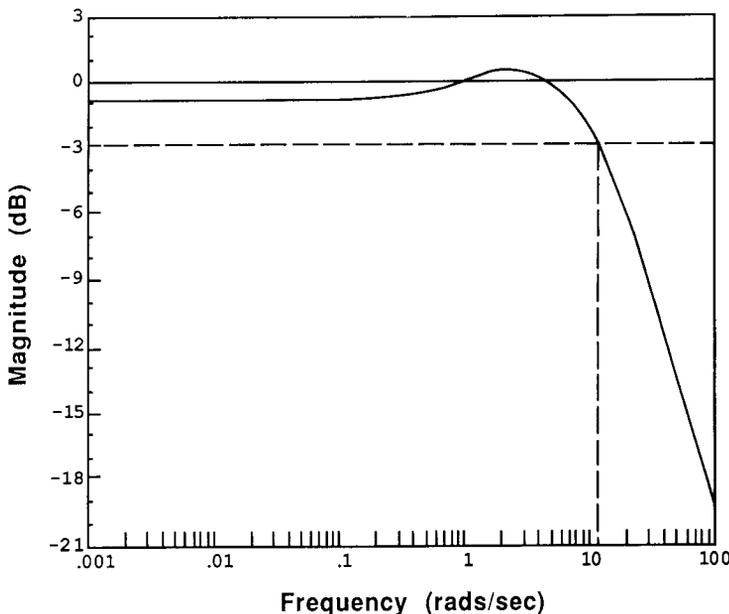
Figure 6.1-3 Closed-loop frequency response.

the ρ selection, and for large ρ will yield a loop gain (apart from a scaling factor) with the desired magnitude response. But, of course, we do not have $k_\rho \simeq \pm\sqrt{\rho}d_1$ unless $d_1'(sI - F)^{-1}g$ is minimum phase.

3. What, then, is a desirable gain characteristic shape for $d_1'(sI - F)^{-1}g$? The answer is that usually something like an integrator response or a first- or second-order bandpass filter characteristic is ideal, since such systems yield desirable closed-loop characteristics in classical control design. Of course, such characteristics for plants higher than second order may not be possible to achieve, even approximately, with any d selection. There may be notches in the high gain region indicating zeros in unfortunate locations. In such cases it may then be possible to use the frequency shaping ideas of Chapter 9. [That is, augment the plant with filters to achieve "nice" gain responses $d_a'(sI - F_a)^{-1}g_a$ for the augmented plant. Here the subscript a denotes augmented matrices.]

4. After d has been set to achieve a "nice" open-loop gain characteristic shape, how then is ρ selected? An initial selection can be made to aim for an appropriate bandwidth for the closed-loop system. This is often roughly the cross-over bandwidth of the open-loop system, denoted here by ω_1. With such a selection, a trial design can be carried out with a view to refining the selection in subsequent trials. If the bandwidth achieved in a trial is too low, then increase ρ, keeping in mind that ω_1 is proportional to $\sqrt{\rho}$.

Low control weighting.   We conclude this section by making an observation that tells us very little more about Q and R selection, but does highlight one further property associated with high state weighting. In computing an optimal control law, an alternative to increasing the state weighting is to decrease the control weighting. (The optimal performance index is naturally affected.) Let us then suppose the index is $\int_0^\infty (\mu u^2 + x'Qx)\, dt$, where we think of μ going to zero. Clearly, for fixed μ the optimal control for $\int_0^\infty (\mu u^2 + x'Qx)\, dt$ is the same as that for $\int_0^\infty (u^2 + \mu^{-1}x'Qx)\, dt)$, so that all we have said about the optimal control closed-loop poles, loop gain and so on with high state weighting applies with $\mu \to 0$, with only unessential notational adjustment caused by replacement of ρ with μ^{-1}. The value of the optimal index is, however, a different matter. Clearly, we are penalizing the control less and less and it is almost immediate that the optimal cost $x_0'\bar{P}_\mu x_0$ decreases with μ. The question then arises: in the limit as $\mu \to 0$, will the optimal cost also be zero, or is there some irreducible minimum below which we cannot go? If there are uncontrollable states, they will necessarily contribute to the irreducible minimum of the performance, irrespective of the control signals. Let us therefore separate out this issue by eliminating it from consideration.

Suppose that with the usual assumptions and with $[F, g]$ controllable,

$$g'(-sI - F')^{-1}Q(sI - F)^{-1}g = \frac{q(s)}{p_0(-s)\,p_0(s)} = \frac{m(-s)\,m(s)}{p_0(-s)\,p_0(s)}$$

where $m(\cdot)$ is obtained by factoring $q(\cdot)$, and assigning all left half-plane roots to $m(\cdot)$. Suppose that d is then defined from $m(s)$ by

$$d'(sI - F)^{-1}g = \frac{m(s)}{p_0(s)}$$

We know that with Q as given, or replaced by dd', the closed-loop poles are the zeros of the stable polynomial $p_{c\mu}(s)$ given by

$$p_{c\mu}(-s)p_{c\mu}(s) = p_0(-s)p_0(s) + \mu^{-1}m(-s)m(s)$$

and so the optimal k_μ is the same with weighting Q or dd'. The corresponding Riccati equation solutions satisfy

$$\bar{P}_{1\mu}F + F'\bar{P}_{1\mu} - \bar{P}_{1\mu}g\mu^{-1}g'\bar{P}_{1\mu} + Q = 0$$

$$\bar{P}_{2\mu}F + F'\bar{P}_{2\mu} - \bar{P}_{2\mu}g\mu^{-1}g'\bar{P}_{2\mu} + dd' = 0$$

with $\bar{P}_{1\mu}g\mu^{-1} = \bar{P}_{2\mu}g\mu^{-1} = -k_\mu$. We also know from the high state weighting result—see (6.1-8)—that $\mu^{1/2}k_\mu \to d$ as $\mu \to 0$. Consequently, as $\mu \to 0$, these two equations give

$$\bar{P}_{2\mu}F + F'\bar{P}_{2\mu} - dd' + Q \to 0$$

$$\bar{P}_{1\mu}F + F'\bar{P}_{1\mu} \to 0$$

From the first of these limits, it is clear that if $Q \neq dd'$, then $\bar{P}_{2\mu} \nrightarrow 0$. (Argument by contradiction is trivial.) The second implies $\bar{P}_{1\mu} \to 0$, by the following argument. Observe that $\bar{P}_{1\mu}g = -\mu k_\mu \to -\mu^{1/2}d \to 0$ as $\mu \to 0$. Multiplying the second limit above by g yields $\bar{P}_{1\mu}Fg + F'\bar{P}_{1\mu}g \to 0$ and so $\bar{P}_{1\mu}Fg \to 0$. Similarly, $\bar{P}_{1\mu}F^2g$, $\bar{P}_{1\mu}F^3g, \ldots \to 0$ whence $P_{1\mu} \to 0$. Hence, we conclude that if $Q = dd'$ with $d'(sI - F)^{-1}g$ possessing zeros in the left half-plane, the optimal index goes to zero as $\mu \to 0$. Otherwise, it tends to a nonzero quantity.

If we set $y = d'x$, the index is $\int_0^\infty (\mu u^2 + y'y)\, dt$. If we think of the transfer function $d'(sI - F)^{-1}g$ as that of a controllable plant with input u and output y, we have argued that the cost of controlling this plant with vanishingly small control penalty will be zero if the plant is minimum phase and nonzero otherwise.

Main points of the section.   As state weighting goes to zero, the closed-loop poles approach the stable open-loop poles or the reflections through the $j\omega$-axis of the unstable poles. If $Q = dd'$ with $d'(sI - F)^{-1}g$ possessing stable zeros, as the state weighting goes to infinity, the closed-loop poles that remain finite approach the zeros of $d'(sI - F)^{-1}g$ and the remainder lie in a Butterworth pattern on a circle with radius that grows as $\rho^{1/(2n - 2l)}$ where l, n are the numerator and denominator degree of $d'(sI - F)^{-1}g$. With ρ large, approximate expressions are available for the optimal gain, the open-loop response and in particular the loop gain cross-over frequency, and the closed-loop bandwidth. If high state weighting is replaced by low control weighting, the optimal cost tends to zero as the weighting

tends to zero if and only if the transfer function from u to $y = d'x$ is minimum phase, where $Q = dd'$, and $[F, g]$ is controllable.

The just noted asymptotic properties lead to a design approach for weighting coefficient selection to achieve robustness.

Problem 6.1-1. Show that for any second-order single-input system $\dot{x} = Fx + gu$ with performance index defined by $r = 1$ and $Q \geq 0$ (under stabilizability and detectability), the closed-loop poles can be determined analytically. Suppose that det $(sI - F) = s^2 + as + b$ and $g'(-sI - F')^{-1}Q(sI - F)^{-1}g = (-c^2s^2 + d^2)(s^2 + as + b)^{-1}(s^2 - as + b)^{-1}$. Find the closed-loop characteristic polynomial.

Problem 6.1-2. For the standard problem with $Q = dd'$, with $d'(sI - F)^{-1}g$ possessing stable zeros, define the closed-loop transfer function as

$$w_c(s) = k'_\rho(sI - F - gk'_\rho)^{-1}g$$

and its cut-off frequency by that ω_0 for which

$$|w_c(j\omega_0)|^2 = \tfrac{1}{2}|w_c(0)|^2$$

Show that approximately (ρ large)

$$|d'(j\omega_0 I - F)^{-1}g|^2 = \frac{1}{\rho + (2/|d'F^{-1}g|^2)} \quad (F \text{ nonsingular})$$

$$= \rho^{-1} \quad (F \text{ singular})$$

Problem 6.1-3. Consider a system

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} x + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u$$

with performance index $V_1 = \int_0^\infty (u^2 + x_1^2 + x_2^2)\, dt$

(i) Find the optimal control law
(ii) Change the index to $V_2 = \int_0^\infty [u^2 + 900(x_1^2 + x_2^2)]\, dt$
   Determine approximately the optimal control law and cross-over frequency. Compare the results with the actual values.

6.2 MULTIVARIABLE SYSTEMS

For a single-input system $\dot{x} = Fx + gu$, once the closed-loop eigenvalues of $F + gk'$ have been specified, there is no further freedom that can be taken up in the choice of k. However, for a multi-input system

$$\dot{x} = Fx + Gu \tag{6.2-1}$$

knowledge of the eigenvalues of $F + GK'$ does not of itself determine K uniquely.

It is this additional freedom which, in a rough sense, lies behind the greater complexity of the asymptotic properties of multivariable systems.

As in the previous section, we shall focus our attention first on low state weighting, and then low control weighting. Throughout this section, we shall assume for convenience that $[F, G]$ is controllable, and not just stabilizable.

Low state weighting.   For the problem parametrized by F, G, ρQ, R with Re $\lambda_i(F) < 0$, the optimal performance index, call it $x_0' \bar{P}_\rho x_0$, is overbounded by the value obtained with $u \equiv 0$, viz.

$$\rho x_0' \hat{P} x_0 = \rho x_0' \int_0^\infty e^{F't} Q e^{Ft} \, dt \, x_0$$

(Here, $\hat{P}$ is the solution of $\hat{P}F + F'\hat{P} = -Q$; see Appendix A). It follows that as $\rho \to 0$, $\bar{P}_\rho \to 0$. Hence $K_\rho \to 0$. Thus, as for the single-input case, the gain goes to zero. Further, the return difference equality, viz.

$$[I - G'(-sI - F')^{-1}K_\rho]R[I - K_\rho'(sI - F)^{-1}G]$$

$$= R + \rho G'(-sI - F')^{-1}Q(sI - F)^{-1}G \tag{6.2-2}$$

yields as $\rho \to 0$ for $s = j\omega$

$$|\det [I - K_\rho'(j\omega I - F)^{-1}G]|^2 \det R$$

$$= \det [R + \rho G'(-j\omega I - F')^{-1}Q(j\omega I - F)^{-1}G] \to \det R$$

Now

$$\det [I - K_\rho'(j\omega I - F)^{-1}G] = \frac{\det (j\omega I - F - GK_\rho')}{\det (j\omega I - F)}$$

So as $\rho \to 0$,

$$\left| \frac{\det (j\omega I - F - GK_\rho')}{\det (j\omega I - F)} \right| \to 1$$

That is, as $\rho \to 0$

$$p_{c\rho}(-j\omega)p_{c\rho}(j\omega) \to p_0(-j\omega)p_0(j\omega) \tag{6.2-3}$$

Again, this parallels the single-input case. When $\rho \to 0$ the closed-loop eigenvalues approach the stable open-loop eigenvalues, and the reflections through the imaginary axis of the unstable open-loop eigenvalues. It is not hard to show that the matrix $\bar{P}_\rho$ defining the optimal performance index decreases with ρ and thus $\bar{P}_0 = \lim_{\rho \downarrow 0} \bar{P}_\rho$ exists. Let $K_0 = -\bar{P}_0 GR^{-1}$ be the associated control law. An interesting fact is that if some of the open-loop eigenvalues are unstable, $F + \frac{1}{2}GK_0'$ has this number of eigenvalues on the $j\omega$-axis. (See Problem 6.2-1.)

High state weighting.   As for the single-input case, when the state weighting becomes very high, some of the closed-loop eigenvalues approach left half-plane zeros associated with $G'(-sI - F')^{-1}Q(sI - F)^{-1}G$ and other eigen-

values become arbitrarily large. We shall consider only the most important case here, following [2]. This is also the easiest case to grasp. For more complicated cases, see, for example, [3].

In conjunction with each closed-loop eigenvalue, there is also a closed-loop eigenvector. Thus if $s_{i\rho}$ is a closed-loop eigenvalue, a vector $\mathbf{w}_{i\rho}$ for which

$$(s_{i\rho}I - F - GK'_\rho)\,\mathbf{w}_{i\rho} = 0 \tag{6.2-4}$$

is the associated eigenvector. Then if $\mathbf{w}_{i\rho}$ is the initial state for the closed-loop system $\dot{x} = (F - GK'_\rho)x$, the response will be $\exp(s_{i\rho}t)\mathbf{w}_{i\rho}$. One practical consequence of this is that if $\mathbf{w}_{i\rho}$ has a zero entry in some position, the response retains a zero entry for all time in this position. Suppose F is $n \times n$ and G is $n \times m$. If $m = 1$, specification of $s_{1\rho}, \ldots, s_{n\rho}$ alone determines the gain vector k_ρ and (apart from pathological situations) the eigenvectors. But if $m > 1$, K_ρ is not fully determined and neither are the eigenvectors. Thus they become more important in the multiple input case, since they carry additional degrees of freedom. Hence in the following material, we shall focus attention not just on the closed-loop eigenvalues but also on the closed-loop eigenvectors. A later example will illustrate their importance in design.

Let us now adopt the following assumption.

Assumption 6.2-1

> The weighting matrix is $\rho DD'$, where D is $m \times n$ with $D'G$ nonsingular, and the zeros of $\det D'(sI - F)^{-1}G$ are distinct, have negative real parts, and differ from the eigenvalues of F.

In the single-input case, much of our discussion focused on the case where this assumption held. This assumption underpins [2]. We shall now summarize (without proof) two key results set out with supporting arguments in [2].

Finite eigenvalues and eigenvectors.   The first main result concerns the finite closed-loop eigenvalues and associated eigenvectors: as $\rho \to \infty$, there are $(n - m)$ eigenvalues of the closed-loop system $s_{i\rho}$ which approach the $n - m$ zeros of $\det D'(sI - F)^{-1}G$. Call these zeros s_i^ρ. The associated close-loop eigenvectors $x_{i\rho}$ approach limits defined by

$$x_i^\rho = (s_i^\rho I - F)^{-1}Gv_i^\rho \tag{6.2-5}$$

where v_i^ρ is a null-vector of $D'(s_i^\rho I - F)^{-1}G$:

$$D'(s_i^\rho I - F)^{-1}Gv_i^\rho = 0 \tag{6.2-6}$$

Note the immediate implication, holding for all i:

$$D'x_i^\rho = 0 \tag{6.2-7}$$

Of course, the v_i^ρ and x_i^ρ are defined only to within scaling constants. We show below that given a list of s_i^ρ and m-vectors v_i^ρ, $i = 1, \ldots, n - m$ (and if a complex s_i^ρ occurs,

so must its complex conjugate, and similarly for v_i^p), then one can find a D satisfying the requirements of the assumption. Thus via choice of D, we have approximate control (approximate because we never set $\rho = \infty$) over $(n - m)$ closed-loop eigenvalues, and some control over the associated eigenvectors [through (6.2-5)]. A later example will show how we can exploit this element of choice in the eigenvectors.

The actual construction from prescribed s_i^p and v_i^p of a D to satisfy (6.2-6) is easy. Suppose that with x_1^p defined by (6.2-5),

$$[x_1^0 \, x_2^0 \ldots x_{n-m}^0] = X^0 = T \begin{bmatrix} I \\ 0 \end{bmatrix}$$

where T is nonsingular $n \times n$ and the identity matrix is $(n - m) \times (n - m)$. (In case X^0, which is $n \times m$, has rank less than m, a corresponding change can be made.) Then, to secure $D'X_0 = 0$, clearly we can use

$$D' = [0_{m \times (n-m)} \quad *] \, T^{-1}$$

where $*$ is an arbitrary $(m \times m)$ matrix.

Evidently, there are $(m \times m)$ degrees of freedom in D. This is what one should expect from (6.2-6) and (6.2-7)—premultiplication of D' by any constant nonsingular $m \times m$ matrix will leave the equation unaltered.

Unbounded eigenvalues and their eigenvectors.

The second main result culled from [2] concerns the closed-loop eigenvalues, which become very large as $\rho \to \infty$: as $\rho \to \infty$, there are m eigenvalues s_{jp} which approach quantities $\sqrt{\rho} \, s_j^x$ and associated closed-loop eigenvectors x_{jp} approaching limits

$$x_j^x = G \, v_j^x \tag{6.2-8}$$

where in turn the s_j^x and v_j^x satisfy the generalized eigenproblem

$$[(s_j^x)^2 R - G'DD'G]v_j^x = 0 \qquad j = 1, \ldots, m \tag{6.2-9}$$

Normally, one would think of (6.2-9) as determining s_j^x and v_j^x, given R, G, D. But one could turn the idea around. Having selected D, one could seek to select R to secure desired s_j^x and v_j^x. In case all s_j^x are different, it turns out that there is a unique positive definite R satisfying (6.2-9) if and only if $v_i^x{}'G'DD'Gv_j^x = 0$ for all $i \neq j$ (see Problem 6.2-3). In case all s_j^x are the same, then $R = G'DD'G$ works and the v_j span m-space. In case two or more s_j^x are the same, in-between cases result.

Of course, once Q and R are selected, there remains the task of selecting ρ. In this connection, we note the following facts—see, [4]—for example, paralleling results already stated for the single-input case:

1. (Asymptotic behavior of K_ρ):

$$\rho^{-1/2}R^{1/2}K_\rho' \to WD' \tag{6.2-10}$$

for some orthogonal W.

2. (Asymptotic singular value behavior)

$$\sigma_i\{R^{1/2}[I - K_p'(j\omega I - F)^{-1}G]R^{-1/2}\}$$

$$= \{\lambda_i(I + \rho R^{-1/2}G'(-j\omega I - F')^{-1}DD'(j\omega I - F)^{-1}GR^{-1/2}\}^{1/2}$$

$$\rightarrow \sqrt{\rho}\ \sigma_i[D'(j\omega I - F)^{-1}GR^{-1/2}]$$ 　　　(6.2-11)

3. (High frequency behavior of loop gain)

$$K_p'(j\omega I - F)^{-1}G \rightarrow \frac{K_p'G}{j\omega} \simeq \frac{\rho^{1/2}R^{-1/2}WD'G}{j\omega}$$ 　　　(6.2-12)

4. (Cross-over frequency)

$$\omega_1 \simeq \rho^{1/2}\ \overline{\sigma}\ [R^{-1/2}WD'G]$$ 　　　(6.2-13)

　　　Naturally, when R is a multiple of the identity, simplifications accrue above. Note in particular that the matrix W drops out of (6.2-13).

　　　Example. We shall now look at an example, drawn from [2]. The lateral axis dynamics of an aircraft are

$$\dot{x} = Fx + Gu$$

where

$$x = \begin{bmatrix} p_s \\ r_s \\ \beta \\ \varphi \\ \delta_r \\ \delta_a \end{bmatrix} = \begin{bmatrix} \text{stability axis roll rate} \\ \text{stability axis yaw rate} \\ \text{sideslip angle} \\ \text{bank angle} \\ \text{rudder deflection} \\ \text{aileron deflection} \end{bmatrix}$$

$$u = \begin{bmatrix} \delta_{rc} \\ \delta_{ac} \end{bmatrix} = \begin{bmatrix} \text{rudder command} \\ \text{aileron command} \end{bmatrix}$$

$$F = \begin{bmatrix} -0.746 & 0.387 & -12.9 & 0 & 0.952 & 6.05 \\ 0.024 & -0.174 & 4.31 & 0 & -1.76 & -0.416 \\ 0.006 & -0.999 & -0.0578 & -0.0369 & 0.0092 & -0.0012 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -10 & 0 \\ 0 & 0 & 0 & 0 & 0 & -5 \end{bmatrix}$$

$$G = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 20 & 0 \\ 0 & 10 \end{bmatrix}$$

The last two rows and columns of F capture the actuator dynamics.

The desired finite asymptotic eigenvalues s_i^p available from a specification are

$$
\begin{array}{lll}
s_1^0 & \text{Roll subsidence mode} & -4.0 \\
s_2^0, s_3^0 & \text{Dutch roll mode} & -0.63 \pm j\, 2.42 \\
s_4^0 & \text{Spiral mode} & -0.05
\end{array}
$$

The roll subsidence mode should show up as little as possible in the state elements yaw rate and sideslip. Why? Imagine a correction requiring roll of the aircraft. Then we do not want to introduce yaw or sideslip in making the correction. This suggests that we desire (with * denoting don't care)

$$
x_1^0 = [1 \quad 0 \quad 0 \quad * \quad * \quad *]'
$$

We would expect that the bank angle will change, as well as at least one of the actuators functioning when the aircraft rolls. So we would expect at least two of the don't care entries to end up nonzero. Similarly, physical considerations suggest that it is desirable if the Dutch roll mode (which is oscillatory) not show up in the state elements' roll rate or bank angle. The spiral mode should show up primarily in the bank angle, and not in the side slip. Allowing for normalization of the eigenvalues, we find that this implies

$$
x_2^0 = \bar{x}_3^0 = [0 \quad *+j1 \quad 1+j* \quad 0 \quad * \quad *]'
$$

or

$$
\operatorname{Re} x_2^0 = [0 \quad 1 \quad * \quad 0 \quad * \quad *]'
$$

$$
\operatorname{Im} x_2^0 = [0 \quad * \quad 1 \quad 0 \quad * \quad *]'
$$

Also

$$
x_4^0 = [* \quad * \quad 0 \quad 1 \quad * \quad *]'
$$

In the case of a real vector like x_1^0 we proceed as follows. Let y_1^0 denote the subvector of x_1^0 formed by deleting the don't care entries, and let A_1^0 denote the rows of $(s_1^0 I - F)^{-1} G$ obtained by retaining the same rows as the entries of x_1^0 used to form y_1^0. Then we desire $A_1^0 v_1^0 = y_1^0$. If this is not exactly achievable, then we choose v_1^0 to secure a best fit, that is, minimize $\|y_1^0 - A_1^0 v_1^0\|^2$. Thus

$$
v_1^0 = (A_1^{0\prime} A_1^0)^{-1} A_1^{0\prime} y_1^0
$$

In the case of a complex vector like x_2^0, let y_2^0 and y_3^0 be formed from $\operatorname{Re} x_2^0$ and $\operatorname{Im} x_2^0$ just as y_1^0 was formed from x_1^0. Let A_2^0 and A_3^0 denote the corresponding (complex) submatrices of $(s_2^0 I - F)^{-1} G$. Then

$$
v_2^0 = (A_2^{0*} A_2^0)^{-1} A_2^{0*} y_2^0 + j (A_3^{0*} A_3^0)^{-1} A_3^{0*} y_3^0
$$

$$
v_3^0 = \bar{v}_2^0
$$

Note that in x_4^0 only two entries are constrained. Hence we would expect that by choice of v_4^0 (which is a 2-vector) we should be able to actually attain this x_4^0.

However, there are more than two constraints with x_1^0, and so one could not a priori expect they will be met.

When the v_i^p are chosen in this way, the actual $x_i^p = (s_i^p I - F)^{-1} G v_i^p$ which are secured are

$$x_1^0 = [1 \qquad -0.007 \quad 0 \qquad -0.25 \quad 0.13 \qquad -0.56]'$$

$$x_2^0 = \bar{x}_3^0 = [0 \qquad 15.6 \quad 1 \qquad 0 \qquad 7.88 \qquad -0.103]'$$

$$+ j [0 \qquad 1 \qquad 6.16 \quad 0 \qquad -9.49 \qquad 14.6]'$$

$$x_4^0 = [-0.05 \qquad 0.037 \quad 0 \qquad 1 \qquad -0.0014 \quad -0.0079]'$$

It can be seen that the differences with the desired eigenvectors are not great.

The matrix D can now be found to satisfy Equation (6.2-6), or equivalently (6.2-7), viz. $D'x_i^0 = 0, i = 1, \ldots, 4$. Utilizing the 2×2 freedom in D to force the last two columns of D' to be the identity, we find

$$D' = \begin{bmatrix} -0.131 & -0.612 & 1.64 & 0.0175 & 1.0 & 0 \\ 0.567 & 0.160 & -2.39 & 0.0303 & 0 & 1.0 \end{bmatrix}$$

This choice of normalization of D implies that $D'G$ is diagonal:

$$D'G = \begin{bmatrix} 20 & 0 \\ 0 & 10 \end{bmatrix}$$

The remaining two closed-loop eigenvectors are associated with the large ρ modes. These eigenvectors necessarily lie in the range space of G—see (6.2-8). As such, they are intimately associated with the operation of the actuators, both for this problem and in general. In this problem, the bandwidth of the actuators is explicitly displayed in the F matrix, as 10 and 5 rad/s. Now with control, the associated modes will be $\sqrt{\rho} s_1^x$ and $\sqrt{\rho} s_2^x$. It may make sense to choose s_1^x and s_2^x in the ratio $2:1$, and then later to adjust ρ so that the actuator bandwidths are comfortably employed. Accordingly, set $s_1^x = 1, s_2^x = 0.5$. It also makes sense to associate x_1^x, x_2^x with each of the two actuators, so that $x_1^x = [0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0]', x_2^x = [0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 1]$. Appeal to (6.2-9) and the evaluation of $D'G$ above shows that

$$R = \begin{bmatrix} 400 & 0 \\ 0 & 400 \end{bmatrix}$$

satisfies (6.2-9).

The final step in the design is to choose ρ. A reasonable choice turns out to be $\rho = 400$. (This apparently doubles the actuator bandwidths, which are $\sqrt{\rho} s_1^x$ and $\sqrt{\rho} s_2^x$, and $\sqrt{\rho} s$, but as explained in [2], there is some conservativeness in the use of the earlier bandwidths of 10 and 5.)

When the transient responses are examined, the various transients all display at least qualitatively the desired decay rate and cross-coupling characteristics.

The feedback gain matrix turns out to be

$$K_\rho' = - \begin{bmatrix} -0.132 & -0.882 & 1.576 & 0.026 & 0.681 & -0.026 \\ 0.524 & 0.420 & -2.827 & 0.021 & -0.013 & 0.860 \end{bmatrix}$$

In this case $\rho^{-1/2} R^{1/2} = I$, and formula (6.2-10) is approximately verified with $W = I$:

$$K_\rho' \simeq D'$$

The cross-over frequency predicted from (6.2-13) is $\omega_1 = 20$.

Low control weighting.   We shall conclude this section by commenting briefly on the problem where $V = \int_0^\infty (\mu u'Ru + x'Qx)\, dt$ and $\mu \to 0$. Obviously, the optimal gain is the same whether R is replaced by μR or Q is replaced by ρQ with $\rho = \mu^{-1}$. What of the optimal performance index $x_0' \bar{P}_\mu x_0$? Suppose that $Q = DD'$. In [5], the following results are stated. They generalize the results for single-input plants. See also [6].

1. If dim $D'x >$ dim u (more outputs to control than inputs), then $\lim_{\mu \downarrow 0} \bar{P}_\mu \neq 0$.
2. If $[F, G]$ is controllable, if dim $D'x =$ dim u, and if det $[D'(sI - F)^{-1}G]$ is not identically zero and has all its zeros in Re $[s] < 0$ (minimum phase assumption), then $\lim_{\mu \downarrow 0} \bar{P}_\mu = 0$.
3. If $[F, G]$ is controllable, if dim $D'x <$ dim u, and if there exists M such that $D'(sI - F)GM$ is square, and has a determinant that is not identically zero with all zeros in Re $[s] < 0$ (plant can be squared down to minimum phase), then $\lim_{\mu \downarrow 0} \bar{P}_\mu = 0$.

Not surprisingly, as the control weighting tends to zero, signals can get arbitrarily large. More precisely, near $t = 0$, $x(t)$ will be at least discontinuous and may contain an impulse, doublet, and so on while $u(t)$ contains an impulse, or may contain a doublet, triplet, and so on, see for example [6] for a discussion.

Main points of the section.   With low state weighting, multiple-input plants behave like single-input plants: closed-loop eigenvalues are made up of the stable open-loop eigenvalues and reflection through the imaginary axis of the unstable open-loop eigenvalues. With high state weighting, closed-loop eigenvalue behavior is similar to that for single-input plants. In addition, there is some control over the closed-loop eigenvectors. The closed-loop eigenvalues that are asymptotically infinite and the associated eigenvectors are affected by both ρ and R. The low control weighting result mimics the single-input case; the optimal cost is only zero when a minimum phase "plant" is being controlled.

Problem 6.2-1.   Consider the optimal control problem parametrized by F, G, ρQ, R with F possessing one or more eigenvalues in Re $[s] > 0$. Let $\bar{P}_\rho$, K_ρ be the associated Riccati equation solution and optimal gain, and let $\bar{P}_0$, K_0 be the limits as $\rho \to 0$. Show that $F + \frac{1}{2}GK_0'$ has a pure imaginary eigenvalue. (Assume for convenience that $F + \frac{1}{2}GK_0'$ is diagonalizable.) [*Hint:* Start with the identity $\bar{P}_\rho(F + \frac{1}{2}GK_\rho') + (F' + \frac{1}{2}K_\rho G')\bar{P}_\rho + \rho Q = 0$.]

Problem 6.2-2.   Consider the optimal control problem with F, G, ρQ, R with

$$F = \begin{bmatrix} F_1 & 0 \\ 0 & F_2 \end{bmatrix}$$

where F_1 has all eigenvalues in Re $[s] < 0$ and F_2 has all eigenvalues in Re $[s] > 0$. Show that $\lim_{\rho \to 0} \bar{P}_\rho$ is of the form

$$\bar{P} = \begin{bmatrix} 0 & 0 \\ 0 & \bar{P}_{02} \end{bmatrix}$$

and that $\bar{P}_{02}^{-1}$ satisfies a linear equation. [*Hint:* Find an equation satisfied by $\bar{P}_{02}$ and verify that the form given satisfies the equation and ensures that $F - GR^{-1}G'\bar{P}_0$ is stable.]

Problem 6.2-3.   Consider the equation

$$[(s_j^\infty)^2 R - G'DD'G]v_j^\infty = 0 \qquad j = 1, \ldots, m$$

in which $D'G$ is nonsingular and $m \times m$, the s_j^∞ are distinct, and the v_j^∞ are independent. Show that there exists a unique $R = R' > 0$ satisfying this equation if and only if $(v_i^\infty)'G'DD'Gv_j^\infty = 0$ for all $i \neq j$. [*Hint (only if)*: Premultiply the ith equation by $v_j^\infty{}'$ and postmultiply the jth by v_i^∞. *Hint (if)*: Show the equation is equivalent to $RVS = (G'DD'G)V$, where S and $V'(G'DD'G)V$ are diagonal.]

6.3 FURTHER ISSUES IN Q, R SELECTION

In the preceding two sections, we have given a number of insights into Q, R selection. In particular, we have indicated how some control of closed-loop eigenvalues, eigenvectors, and bandwidth can be exercised. The benefits of choosing R diagonal, and even as a multiple of the identity, have been mentioned. We shall now record a number of miscellaneous points that can also be borne in mind. Four preliminary qualifying remarks ought, however, to be made.

1. Designers need to be prepared to combine Q, R selection with iteration. The translation of specifications into Q, R selection is imprecise, and so often initially chosen Q, R may be inappropriate.
2. Selection of Q, R may interact with the state estimator design process. This course will be addressed in Chapter 8.
3. It is possible to use Q and R that are frequency dependent as we discuss in Chapter 9. This provides another degree of freedom again for the designer.
4. It is generally advisable to describe the plant with a state space coordinate basis in which individual entries of the state vector have physical significance. Then the choice of Q, R entries is more readily reflective of physical insights, especially if diagonal Q, R are used. Moreover, if scaling of each variable can be used so that all are expressed in the same units, rather than one being in

millinewtons and another in kilonewtons, or even as dimensionless quantities, so much the better.

Limiting the magnitude of key variables. Specifications may impose limits on the excursions of a control in bringing certain nonzero states to zero. Specifications may also limit the mean square excursion of state vector entries under certain noise excitations, or the maximum excursion of one state vector entry during the control of another. The general principle is: when a variable of interest takes too high a value in a trial design, increase the weighting given to it in the performance index. Thus if $R = \mathrm{diag}\,(r_1, r_2, \ldots, r_m)$ and u_2 is taking too great a value, then r_2 is increased. Or if $Q = dd'$ and x_3 is taking too great a value, $x'Qx$ is replaced by $x'dd'x + \alpha x_3^2$ for some positive α.

Generally speaking, if one variable is penalized this way, there is a corresponding reduction of penalty on the other variables, and the excursions of these other variables may become correspondingly greater. If state entry x_i is directly affected by control u_j, one should certainly expect increased weighting on x_i to cause greater excursions of u_j.

In reference [7], the following variant on these ideas appears. For a finite-time problem over the interval $[t_0, T]$, choose Q, R diagonal with the ith diagonal entry of Q^{-1} as $n(T - t_0)$ times maximum acceptable value of $x_i^2(t)$ and the ith diagonal entry of R^{-1} as $m(T - t_0)$ times maximum acceptable value of $u_i^2(t)$; the matrix A weighting the terminal state is diagonal, with ith diagonal entry as n times maximum acceptable value of $x_i^2(t_f)$; here, the dimensions of x and u are n and m, respectively. For an infinite time problem (with no A terminal weighting present), one would choose Q, R the same way, neglecting the common factor $(T - t_0)$.

Dealing with variable input cross-coupling in a multi-input system. Suppose that the nominal plant is

$$\dot{x} = Fx + Gu \tag{6.3-1}$$

but G can be varied to GL, where

$$L = \begin{bmatrix} I & X \\ 0 & I \end{bmatrix} \tag{6.3-2}$$

for some X about which we have little information. However, the uncertain input cross-coupling is known to occur only from input block 2 to input block 1. Recall from Chapter 5 that if a design is executed for the nominal plant, with

$$R = \begin{bmatrix} R_1 & 0 \\ 0 & R_2 \end{bmatrix} \tag{6.3-3}$$

then stability is retained when G varies to GL if

$$\overline{\sigma}^2(X) < \frac{\lambda_{\min}(R_2)}{\lambda_{\max}(R_1)} \tag{6.3-4}$$

Greater tolerance of X will be achieved if eigenvalues of R_2 are increased and

eigenvalues of R_1 are decreased. The effect of such changes in R_1, R_2 is to more heavily penalize u_2 (weighted by R_2) than u_1 (weighted by R_1). Since the cross-coupling is Xu_2, the amount of cross-coupling signal will be reduced if X is kept the same, and this provides a qualitative justification of (6.3-4). There is an important principle here, which we separately enumerate.

Dealing with uncertain parameters in the plant.   In the discussion above, we arrived at an adjustment to R which increased the penalty on the term providing the input to the uncertain parameter (here the term X). Generally speaking, this is a sound procedure to cope with any uncertain parameters. Thus if a particular entry, f_{ij} say, of the F matrix is likely to be highly variable, we can increase the penalty on x_j^2 in the index. The result should be that the signal f_{ij} $x_j(t)$ becomes smaller in magnitude, and so the perturbations introduced by varying f_{ij} will become smaller also. This idea is not, however, a universal panacea. Consider, for example, a single-input plant in which only the first entry g_1 of the **g**-vector is variable. Increase of the control weighting will have a number of additional consequences that may not be attractive. Nevertheless, a successful application of this idea arises in an example included in a problem of Chapter 5. The plant is

$$\dot{x} = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix} x + \begin{bmatrix} 1 + \varepsilon \\ 1 \end{bmatrix} u$$

with performance matrix $V = \int_0^\infty [ru^2(t) + (x_1 + x_2)^2]\, dt$. Of course, ε represents uncertainty, and we can assume that $|\varepsilon| \le \varepsilon_0$ for some positive ε_0, with ε otherwise unknown. Design is achieved with $\varepsilon = 0$. It turns out that for any ε_0, there exists an r_0 such that if $r < r_0$, the closed-loop system will be unstable when $\varepsilon = -\varepsilon_0$. The solution is simply to keep the control weighting large enough.

A second interesting example appearing in [8] applies some of these ideas to a pitch-axis control system for a high-performance aircraft. The plant equations are third order and linear, but contain parameter variations arising from different dynamic pressures. With

$$\begin{aligned} x_1 &= \text{angle of attack} \\ x_2 &= \text{rate of pitch angle} \\ x_3 &= \text{incremental elevator angle} \end{aligned}$$

and $u = \text{control input into the elevator actuator}$

the state equations are

$$\dot{x} = \begin{bmatrix} -0.074 & 1 & -0.012 \\ -8.0 & -0.055 & -6.2 \\ 0 & 0 & -6.6667 \end{bmatrix} x + \begin{bmatrix} 0 \\ 0 \\ 6.6667 \end{bmatrix} u$$

for a dynamic pressure of 500, whereas for a dynamic pressure of 22.1, they become

$$\dot{x} = \begin{bmatrix} -0.0016 & 1.0 & -0.0002 \\ -0.1569 & -0.0015 & -0.1131 \\ 0 & 0 & -6.6667 \end{bmatrix} x + \begin{bmatrix} 0 \\ 0 \\ 6.6667 \end{bmatrix} u$$

The performance index minimized is $\int_{t_0}^{\infty} (u^2 + x'Qx)\ dt$ for two different matrices Q—viz.,

$$Q_1 = \begin{bmatrix} 10 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad Q_2 = \begin{bmatrix} 2 \times 10^6 & 0 & 0 \\ 0 & 4 \times 10^5 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Two optimal control laws $u = k_1'x$ and $u = k_2'x$ are calculated, based on the first state equation in both cases, and on performance indices including Q_1 and Q_2, respectively. The control laws are then used not just for the first state equation (for which they are optimal) but for the second state equation, for which they are not optimal. The following table summarizes the results achieved; the sensitivity improvement refers to the sensitivity in x_1.

State Equation	Control Law	Step Response	Sensitivity Improvement
First	k_1	Very good	Moderate
Second	k_1	Very poor	
First	k_2	Good	Huge
Second	k_2	Acceptable	

The weighting Q_2 penalizes x_1, x_2 much more. These are the inputs driving $f_{11}x_1, f_{21}x_1$ and $f_{22}x_2$, all variable on account of the high variability of f_{11}, f_{21}, and f_{22}. Even though f_{13} and f_{23} are highly variable also, suggesting that x_3 might be more heavily weighted, this does not prove necessary.

Use of a root locus.   In the previous two sections, we have suggested approaches for choosing R and Q with a scaling factor ρ on Q being introduced near the end of the procedure. It is common for designs to be evaluated with a range of values of ρ before one is settled on; a root locus of the closed-loop eigenvalues can be generated, and, as earlier noted, ρ can be used as a direct control over bandwidth. The point is that trial and error selection of a single scalar parameter is a comparatively straightforward exercise.

Incidentally, theory implies an indirect loose control over bandwidth in the following way. Observe (in the usual notation) that

$$\begin{aligned} \Sigma\ \lambda_i(F + GK') &= \text{trace}\ (F + GK') \\ &= \text{trace}\ F - \text{trace}\ (GR^{-1}G'\bar{P}) \\ &= \text{trace}\ F - \text{trace}\ (R^{-1/2}G'\bar{P}GR^{-1/2}) \\ &< \Sigma\ \lambda_i(F) \end{aligned} \qquad (6.3\text{-}5)$$

So the center of gravity of the closed-loop eigenvalues must always move left from the center of gravity of the open-loop eigenvalues. One can also show (see Problem 6.3-1) that

$$\Pi|\lambda_i(F + GK')| \geq \Pi|\lambda_i(F)| \qquad (6.3\text{-}6)$$

Both (6.3-5) and (6.3-6) crudely suggest that the bandwidth is pushed out when the loop is closed.

Setting an internal time constant.   Suppose that within the state vector x there are two components x_1 and x_2, with $\dot{x}_1 = x_2$. Suppose also that in the closed-loop control situation, we should like to have a time constant governing decay in x_1 of approximately τ. Then inclusion of a term $\rho(x_1^2 + \tau^2 x_2^2)$ tends to promote this. To understand why this might be so, consider an index

$$V = \int_0^\infty [u^2 + \rho(x_1^2 + \tau^2 x_2^2)] \, dt$$

We shall have

$$(sI - F)^{-1} g = \frac{1}{\det(sI - F)} \begin{bmatrix} \alpha_1(s) \\ s\alpha_1(s) \\ \cdot \\ \cdot \\ \cdot \end{bmatrix}$$

for a certain polynomial $\alpha_1(s)$, and so, with

$$Q = \text{diag}[1, \tau^2, 0, \ldots]$$

$$g'(-sI - F')^{-1} Q (sI - F)^{-1} g = \frac{\alpha_1(-s)\alpha_1(s)(-s^2\tau^2 + 1)}{p_0(-s)p_0(s)}$$

As $\rho \to \infty$, one of the closed-loop eigenvalues will approach $s = -\tau^{-1}$, since this is a root in Re $[s] < 0$ of $\alpha_1(-s)\alpha_1(s)(-s^2\tau^2 + 1)$. Now notice that if Q were replaced by dd' with d chosen so that

$$d'(sI - F)^{-1} g = \frac{\alpha_1(s)(s\tau + 1)}{p_0(s)}$$

We obtain the same closed-loop control law, since $g'(-sI - F')^{-1} Q (sI - F)^{-1} g$ is unaffected by this change. Observe that $d'x = x_1 + \tau x_2$. So we will get the same control law as if we were using the index

$$V = \int_0^\infty [u^2 + \rho(x_1 + \tau x_2)^2] \, dt$$

For large ρ, this will encourage $x_1 + \tau x_2$ to be small, that is, $x_1 + \tau \dot{x}_1$ to be small. This is like saying that x_1 roughly acquires a decay with time constant τ.

An alternative approach to securing certain transient behavior of a particular variable is the following. Suppose that one seeks to have $y = h'x$ behave like $\dot{y} = -\alpha y$ for a certain α. Then one can incorporate an additional term like $\rho(\dot{y} + \alpha y)^2$ in the index. If $h'g = 0$, this amounts to an adjustment of Q, since

$$\rho(\dot{y} + \alpha y)^2 = \rho[(h'F + \alpha h')x]^2$$

and so Q is increased by $\rho(F'h + \alpha h)(h'F + \alpha h')$. On the other hand, if $h'g \neq 0$

then $\dot{y} + \alpha y$ is of the form $h'x + (h'g)u$. On adding the further term to the performance index, one then *loses* the structure of the performance index as the sum of two terms, one penalizing control and the other penalising state, since a cross-product term enters. Such terms can destroy robustness properties.

Control of integral squared error. Suppose one has desired values for $\int_0^\infty x_i^2(t)\,dt$ and $\int_0^\infty u_j^2(t)\,dt$ for a standard initial condition, these quantities being associated with the closed-loop system. One rule of thumb is

1. Select Q as diagonal, with elements equal to the inverses of the desired integral squared state responses.
2. Select R as diagonal, with elements equal to the inverses of the desired integral squared input responses.

This is indeed crude. Consider the scalar system $\dot{x} = bu$, with $V = \int_0^\infty [ru^2 + qx^2]\,dt$, $b \neq 0$, $r > 0$, $q > 0$. Let $x_0 = 1$. Suppose one desires $\int_0^\infty x^2\,dt = \mu$, $\int_0^\infty u^2\,dt = \sigma$. The rule suggests $q = \mu^{-1}$, $r = \sigma^{-1}$. The actual values turn out to be

$$\int_0^\infty u^2\,dt = \frac{1}{2}\frac{\sqrt{\sigma}}{\sqrt{\mu}} \qquad \int_0^\infty x^2\,dt = \frac{1}{2}\frac{\sqrt{\mu}}{\sqrt{\sigma}}$$

There is nevertheless some rationale for this choice. Recall, or recognize (some computation is required) that the optimal u and x lead to

$$\int_0^\infty ru^2\,dt = \int_0^\infty qx^2\,dt$$

Therefore, if

$$q^{-1} = \int_0^\infty x^2\,dt$$

so that

$$\int_0^\infty qx^2\,dt = 1$$

this identity forces

$$\int_0^\infty ru^2\,dt = 1$$

or

$$\int_0^\infty u^2\,dt = r^{-1}$$

In other words, the Q selection having been made in accord with rule 1, rule 2 is forced by the problem structure.

Real part pole-placement by *Q*, α selection.   It turns out to be possible to position individually and arbitrarily the real parts of the poles of an optimal LQ system where there is also a selection of the prescribed degree of stability variable α. Reference [9] introduces and develops results on this topic. The approach requires evaluation of a collection of solutions of the steady state Riccati equation, rather than the single solutions we have used up to now. The additional solutions are used to perturb the *Q* matrix at the same time as the *F* matrix is perturbed by addition of α*I* (to enforce a degree of stability constraint). The net effect is to move a nominated subset of the original closed-loop eigenvalues to the left by an amount α.

Main points of the section.   The choice of *Q*, *R* can be governed by the following factors:

desired closed-loop eigenvalues, eigenvectors, and bandwidth
choosing *R* diagonal to secure robustness properties
the magnitude of key variables
input cross-coupling
uncertain plant parameters
insight provided by a root locus
setting of an internal time constant
control of integral squared error
pole-placement

Problem 6.3-1.   With the usual notation, show that

$$\Pi|\lambda_i(F + GK')| \geq \Pi|\lambda_i(F)|$$

[*Hint:* Consider a return difference inequality for $s = 0$, and take its determinant.]

REFERENCES

[1] L. Weinberg, *Network Analysis and Synthesis.* New York: McGraw-Hill Book Co. Inc., 1962.

[2] C. A. Harvey and G. Stein, "Quadratic Weights for Asymptotic Regulation Properties," *IEEE Trans. Auto. Control,* Vol. AC-23 (1978), pp. 378–387.

[3] G. Stein, "Generalized Quadratic Weights for Asymptotic Regulator Properties," *IEEE Trans. Auto. Control,* Vol. AC-24 (1979), pp. 559–566.

[4] J. C. Doyle and G. Stein, "Multivariable Feedback Design: Concepts for a Classical/ Modern Synthesis," *IEEE Trans. Auto. Control,* Vol. AC-26 (1981), pp. 4–16.

[5] H. Kwakernaak and R. Sivan. *Linear Optimal Control Systems.* New York: Wiley Interscience, 1972.

[6] B. A. Francis, "The Optimal Linear-Quadratic Time-Invariant Regulator with Cheap Control," *IEEE Trans. Auto. Control,* Vol. AC-24, No. 4 (August 1979), pp. 616–621.

[7] A. E. Bryson, Jr., and Y.-C. Ho, *Applied Control.* Waltham, Mass: Ginn and Co., 1969.

[8] E. Kreindler, "Closed-Loop Sensitivity Reduction of Linear Optimal Control Systems," *IEEE Trans. Auto. Control,* Vol. AC-13 (1968), pp. 254–262.

[9] J. Medanic, H. S. Tharp, and W. R. Perkins, "Pole Placement by Performance Criterion Modification," *IEEE Trans. Auto. Control,* Vol. 33, No. 5 (May 1988), pp. 469–472.

7

State Estimator Design

7.1 THE NATURE OF THE STATE ESTIMATION PROBLEM

The implementation of the optimal control laws discussed hitherto depends on the states of the controlled plant being available for measurement. Frequently, in practical situations, this will not be the case, and some artifice to get around this problem is required. Before outlining the various approaches that may be used, let us dismiss upon practical grounds one "theoretically" attractive state estimation procedure.

Given a completely observable system, the state vector may be constructed from linear combinations of the output, input, and derivatives of these quantities, as we now show. Consider for simplicity a single-output system: thus,

$$\dot{x} = Fx + Gu \qquad (7.1\text{-}1)$$

$$y = h'x \qquad (7.1\text{-}2)$$

Differentiation of (7.1-2) and substitution for $\dot{x}$ from (7.1-1) leads to

$$\dot{y} - h'Gu = h'Fx \qquad (7.1\text{-}3)$$

Differentiating again, and substituting again for $\dot{x}$, we get

$$\ddot{y} - h'G\dot{u} - h'FGu = h'F^2x \qquad (7.1\text{-}4)$$

164

Continuing in this way, we can obtain a set of equations that may be summed up as

$$z = \begin{bmatrix} h' \\ h'F \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ h'F^{n-1} \end{bmatrix} x$$

or

$$z' = x'[h \quad F'h \quad \cdots \quad (F')^{n-1}h] \tag{7.1-5}$$

with z' a row vector, the entries of which consist of linear combinations of y, u and their derivatives.

Now one of the results concerning complete observability is that the pair $[F, H]$ is completely observable if and only if the matrix $[H \ F'H \ \cdots \ (F')^{n-1}H]$ has rank n. (See Appendix B.) In the scalar output case, the matrix $[h \ F'h \ \cdots \ (F')^{n-1}h]$ becomes square, and thus has rank n if and only if it is nonsingular.

Therefore, with the system of (7.1-1) and (7.1-2) completely observable, Eq. (7.1-5) implies that the entries of the state vector x are expressible as linear combinations of the system output, input, and their derivatives.

From the strictly theoretical point of view, this idea for state estimation works. From the practical point of view, it will not. The reason, of course, is that the presence of noise in u and y will lead to vast errors in the computation of x, because differentiation of u and y (and, therefore, of the associated noise) is involved. This approach must thus be abandoned on practical grounds.

However, it may be possible to use a slight modification of the preceding idea, because it is possible to build "approximate" differentiators that may magnify noise but not in an unbounded fashion. The rate gyro, normally viewed as having a transfer function of the form Ks, in reality has a transfer function of the form $Kas/(s + a)$, where a is a large constant. Problem 7.1-2 asks the student to discuss how this transfer function might approximately differentiate.

Generally speaking, a less ad hoc approach must be adopted to state estimation. Let us start by stating two desirable properties of a state estimator (also known as an observer).

1. It should be a system of the form of Fig. 7.1-1, with inputs consisting of the system input and output, and output $x_e(t)$ at time t, an *online* estimate of $x(t)$. This should allow implementation of the optimal control law $u = K'x_e$, according to the scheme of Fig. 7.1-2. (A variation on this approach is to seek
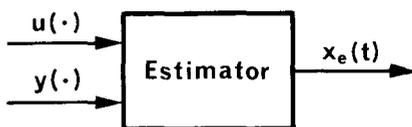
Figure 7.1-1 Desired structure of estimator.
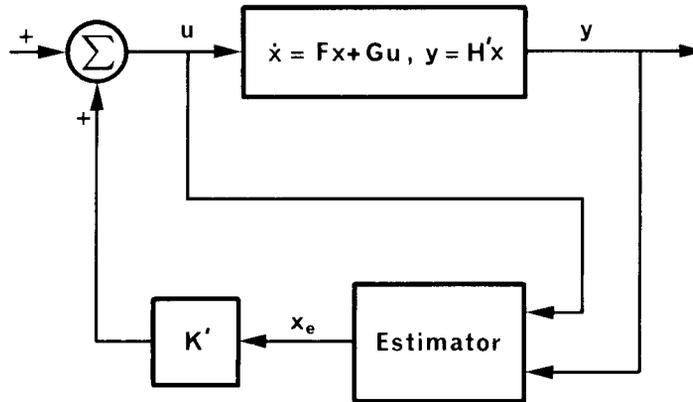
Figure 7.1-2   Use of estimator in implementing a control law.

only an estimate of $K'x$, this being merely a linear function or collection of linear functions of the states, rather than the full state vector x.)

2. It should function in the presence of noise. Preferably, it should be possible to optimize the action of the estimator in a noisy environment—to ensure that the noise has the least possible effect when the estimator is used in connection with controlling a system.

As we shall show, these properties are both more or less achievable. Estimators designed to behave *optimally in certain noise environments* turn out to consist of linear, finite-dimensional systems, if the system whose states are being estimated is linear and finite-dimensional. Moreover, if the dimensions of the system whose states are being estimated and the estimator are the same, and if the system whose states are being estimated is time-invariant and the associated noise is stationary, the estimator is time-invariant. Also, all these properties hold irrespective of the number of inputs and outputs of the system whose states are being estimated.

A further interesting property of such estimators is that their design is independent of the associated optimal feedback law K' shown in Fig. 7.1-2, or of any performance index used to determine K'. Likewise, the determination of K' is independent of the presence of noise, and certainly independent of the particular noise statistics. Yet, use of the control law $u = K'x_e$ turns out to be not merely approximately optimal, but exactly optimal for a modified form of performance index. This point is discussed at length in Chapter 8.

If satisfactory rather than optimal performance in the presence of noise is acceptable, two simplifications can be made. First, the actual computational procedure for designing an estimator (at least for scalar output systems) becomes far less complex. Second, it is possible to simplify the structure of the estimator, if desired. In other words, the designer can stick with the same estimator structure as is used

for optimal estimators (although optimality is now lost), or he or she may opt for a simpler structure. In general, the estimator with simpler structure is more noise prone. The nature of the structural simplification is that the dimension of the estimator is lowered. Even with full state estimation, the dimension can be lowered by one for a single-output system and sometimes by a larger number for a multiple-output system. With estimation of $K'x$, further dimension reductions may be possible.

Because they are simpler to understand, we shall discuss first those estimators that are not designed on the basis of optimizing their behavior in the presence of noise; this class of estimators may be divided into two subclasses, consisting of those estimators with the same dimension as the system and those with lower dimension. Then the class of estimators offering optimal noise performance will be discussed. The next chapter will consider the use of estimators in implementing control laws, and the various properties of the associated plant-controller arrangement.

One point to keep in mind is that the output of the estimator at time t, $x_e(t)$, is normally an estimate rather than an exact replica of the system state $x(t)$, even when there is no noise present. In other words, no estimator is normally perfectly accurate, and thus feedback laws using an estimate rather than the true state will only approximate the ideal situation. Despite this, in many cases little or no practical difficulty arises in controller design as a result of the approximation.

Main points of the section.   There is a need for state estimators that yield estimates of a system state without recourse to differentiation. There can be a trade-off in the design between noise filtering performance, performance in the absence of noise, and estimator complexity.

Problem 7.1-1.   In what sense can the transfer function Ks be approximated by the transfer function $Kas/(s + a)$? What is the significance of the size of a?

Problem 7.1-2.   Show that the circuit of Fig. 7.1-3 can be used as an approximate differentiator. Discuss the effect of noise in e_i.
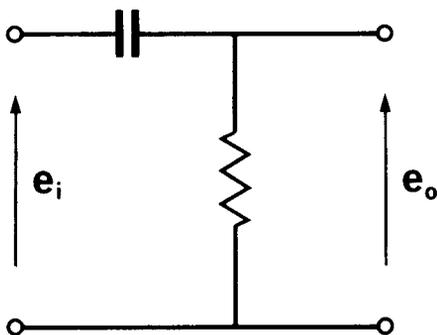
Figure 7.1-3   Circuit for Problem 7.1-2.

7.2 DETERMINISTIC ESTIMATOR DESIGN

In this section, we will first consider the case of deterministic estimator design with estimator dimension equal to that of the plant. Next, the design of reduced order estimators, exploiting the outputs as state information, is studied.

Full-order estimator design.   We consider the design of estimators for systems of the form

$$\dot{x} = Fx + Gu \qquad (7.2\text{-}1)$$

$$y = H'x \qquad (7.2\text{-}2)$$

We assume that F, G, and H are time-invariant, although it is possible to extend the theory to time-varying systems. The estimators will be of the general form

$$\dot{x}_e = F_e x_e + G_{1e} u + G_{2e} y \qquad (7.2\text{-}3)$$

with F_e of the same size as F. Equation (7.2-3) reflects the fact that the inputs to the estimator are the input and output associated with Eqs. (7.2-1) and (7.2-2), hereafter called the *plant input* and *output*. The estimator itself is a linear, finite-dimensional, time-invariant system, the output of which is the estimated state vector of the system.

Before we give a detailed statement specifying how to choose F_e, G_{1e}, and G_{2e}, it is worthwhile to make two helpful observations:

1. It would be futile to think of constructing an estimator using (7.2-3) if the plant equations (7.2-1) and (7.2-2) did not define a detectable pair $[F, H]$, because lack of complete detectability implies the impossibility of determining, by any technique at all, the state of the plant from the plant input and output. Actually, if there is lack of observability we cannot determine the entire state either. But under a detectability assumption, we are assured that the indeterminable part of the state will decay to zero under zero input conditions.

2. One might expect the estimator to be a model for the plant, for suppose that at some time t_0, the plant state $x(t_0)$ and estimator state $x_e(t_0)$ were the same. Then the way to ensure that at some subsequent time $x_e(t)$ will be the same as $x(t)$ is to require that the estimator, in fact, model the plant as

$$\dot{x}_e = Fx_e + Gu \qquad (7.2\text{-}4)$$

Clearly, though, Eq. (7.2-4) is not satisfactory if $x_e(t_0)$ and $x(t_0)$ are different. What is required is some modification of (7.2-4) that reduces to (7.2-4) if $x_e(t_0)$ and $x(t_0)$ are the same, and otherwise tends to shrink the difference between $x_e(t)$ and $x(t)$ until they are effectively the same. Now we may ask what measure we could physically construct of the difference between $x_e(t)$ and $x(t)$. Certainly there is no direct measure, but we do have available $H'x$, and therefore we could physically construct $H'[x_e(t) - x(t)]$. It is to be hoped the complete observability of the plant

would ensure that, over a nonzero interval of time, this quantity contained as much information as $x_e(t) - x(t)$.

These considerations suggest that instead of (7.2-4), we might aim for

$$\dot{x}_e = Fx_e + Gu + K_e H'[x_e - x]$$ (7.2-5)

as an estimator equation. This scheme is shown in Fig. (7.2-1). The equation has the property that if x and x_e are the same at some time t_0, then they will be the same for all $t \geq t_0$, the third term on the right-hand side being zero for all t. Furthermore, it has the property that possible judicious selection of K_e—that is, judicious introduction of a signal into the estimator reflecting the difference between $H'x_e(t)$ and $y(t) = H'x(t)$—may ensure that the error between $x_e(t)$ and $x(t)$ becomes smaller as time advances. Let us now check this latter possibility.

Subtracting (7.2-5) from (7.2-1), we find that

$$\frac{d}{dt}(x - x_e) = F(x - x_e) + K_e H'(x - x_e)$$

$$= (F + K_e H')(x - x_e)$$ (7.2-6)

follows. Consequently, if the eigenvalues of $(F + K_e H')$ have negative real parts, $x - x_e$ approaches zero at a certain exponential rate, and $x_e(t)$ will effectively track $x(t)$ after a time interval determined by the eigenvalues of $F + K_e H'$.
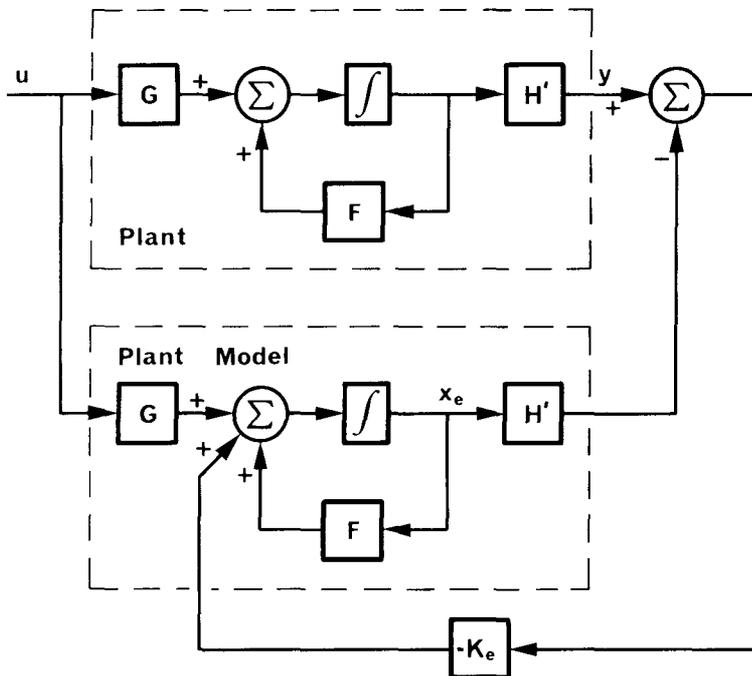
Figure 7.2-1   Estimator, illustrating plant model concept.

Let us recapitulate. We postulated the availability of the input and output of the plant defined by (7.2-1) and (7.2-2), together with complete detectability of the plant. By rough arguments, we were led to examining the possibility of an estimator design of the form of Eq. (7.2-5), which may be rewritten as

$$\dot{x}_e = (F + K_e H')x_e + Gu - K_e y \qquad (7.2\text{-}7)$$

[Figure 7.2-2 shows how this equation can be implemented.] Then we were able to conclude that if K_e could be chosen so that the eigenvalues of $F + K_e H'$ had negative real parts, Equation (7.2-7) did, in fact, specify an estimator, in the sense that $x_e - x$ approaches zero at some exponential rate. Note that at this stage, we have not considered any questions relating to the introduction of noise.

The natural question now arises: When can K_e be found so that $F + K_e H'$ has all eigenvalues with negative parts? The answer is precisely when the pair $[F, H]$ is completely detectable; see Appendix B. If the stronger condition that $[F, H]$ is completely observable is fulfilled, then the eigenvalues of $F + K_e H'$ can be arbitrarily positioned by choice of K_e. In the detectable, unobservable case, certain eigenvalues of $F + K_e H'$ are unaffected by K_e, and one would have to consider whether their presence made the decay of $x - x_e$ unacceptably slow. So far then, aside from the computational details involved in determining K_e, and aside from checking the noise performance, we have indicated one solution to the estimator problem.

The scheme of Fig. 7.2-1, earlier regarded as tentative, has now been shown to constitute a valid estimator. The estimator is a model of the plant, with the addition of a driving term reflecting the error between the plant output $y = H'x$ and the variable $H'x_e$, which has the effect of causing x_e to approach x. Figure 7.2-2 shows
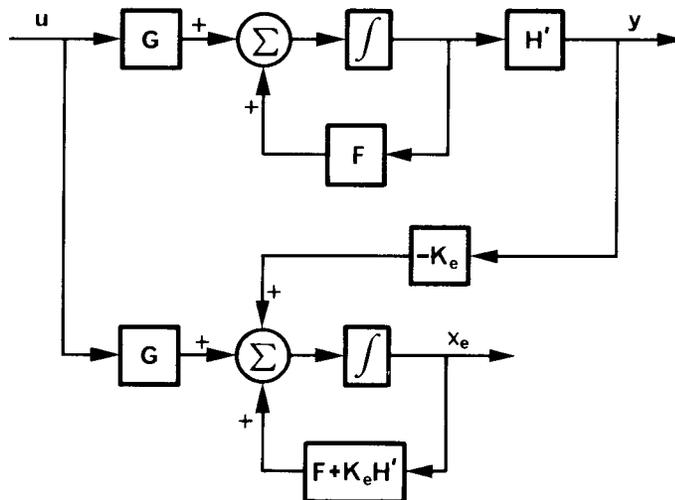
Figure 7.2-2   Estimator with minor simplification.

an alternative valid estimator representation equivalent to that shown in Fig. 7.2-1. In this second figure, however, the concept of the estimator as a model of the plant becomes somewhat submerged.

Let us now consider the question of the effect of noise on estimator operation. If noise is associated with u and y, then inevitably it will be smoothed by the estimator; and if white noise is associated with either u or y (i.e., noise with a uniform power spectrum), the spectrum of the noise in x_e will fall away at high frequencies. In general, the amount of output noise will depend on the choice of K_e, but the problem associated with passing noise into a differentiator will never be encountered. The choice of K_e also affects the rate at which x_e approaches x, because this rate is, in turn, governed by the eigenvalues of $F + K_e H'$. It might be thought that the best way to choose K_e would be to ensure that the eigenvalues of $F + K_e H'$ had real parts as negative as possible, so that the approach of x_e to x would be as rapid as possible. This is so, with one proviso. As the eigenvalues of $F + K_e H'$ get further into the left half-plane, the effective bandwidth of the estimator could be expected to increase, and, accordingly, the noise in x_e due to noise in u and y could be expected to increase. Therefore, noise will set an upper limit on the speed with which x_e might approach x. The situation with using differentiation to estimate x is that the noise becomes infinite, with the estimation time infinitesimally small. The use of estimators of the form just described is aimed at trading off speed of estimation against minimization of loss of performance due to noise. The optimal estimators of Sec. 7.4 essentially achieve the best compromise.

The task of computing K_e to achieve assigned poles is a standard one, [1], though involving somewhat complicated procedures in the multiple-output case. If the plant has a single output, Ackermann's formula achieves a specified closed-loop characteristic polynomial $\Delta(s)$ as

$$K_e = -e_n' [h \quad F'h \quad \cdots \quad (F^{n-1})'h]^{-1} \Delta(F') \tag{7.2-8}$$

where $e_n' = [0 \quad 0 \quad \cdots \quad 1]$, [1]. For the multivariable case, the procedures are considerably more complicated. Of course, linear quadratic *regulator* results can also be used to achieve a stabilizing control law for the system $\dot{w} = F'w + Hv$ to achieve a closed-loop system $\dot{w} = (F' + HK_e')w$ with attractive properties, at least when $[F', H]$ is stabilizable, or equivalently $[F, H]$ is detectable! Indeed, any prescribed degree of stability α can be achieved in the closed-loop system by applying the techniques of Chapter 3, Section 5, so that the eigenvalues of $F' + HK_e'$ can be guaranteed to be to the left of Re $[s] = -\alpha$. Generalizations to the time-varying F, H case can likewise be achieved under uniform observability of $[F, H]$. In the next section, we develop this regulator approach to achieve optimal designs for stochastic noise environments.

We remarked earlier that it is impossible to estimate all the states of a plant that are not observable. However, in general, it is still possible to estimate the observable components of the states of such a plant; if the unobservable components decay extremely rapidly, this partial estimation may be adequate. Problem 7.2-2 asks for an examination of estimator design in these circumstances.

Estimator design with reduced estimator dynamics.

Our aim in this subsection is similar to that of the previous subsection, with a few modifications. We assume that a plant is given with the following equations:

$$\dot{x} = Fx + Gu \tag{7.2-9}$$

$$y = H'x \tag{7.2-10}$$

with F, G, H constant. Without loss of generality, we assume H has full rank, viz. $m \leq n$. Then measurement of y gives us information about m independent linear functionals of x. We seek in addition an $(n - m)$ − dimensional system of the form

$$\dot{w} = F_e w + G_{1e} u + G_{2e} y \tag{7.2-11}$$

such that from w and y together, x can be estimated. In case $[F, H]$ is observable, we would like to be able to control the rate at which the estimation error goes to zero. Figure 7.2-3 illustrates the concept.

Reduced order estimators were first described in [2] by Luenberger for the scalar output case; then the multiple-output case was given in [3]. Our treatment follows that of [1].

The first step is to change the coordinate basis so that entries of y agree with entries of x. Let T be a nonsingular $n \times n$ matrix of the form

$$T = \begin{bmatrix} L' \\ H' \end{bmatrix} \tag{7.2-12}$$

(Here L' is arbitrary, provided that T is nonsingular.) Now change the coordinate basis so that (7.2-9) and (7.2-10) are replaced by

$$\dot{\bar{x}} = TFT^{-1} \bar{x} + TGu \tag{7.2-13}$$

$$y = H'T^{-1} \bar{x}$$

$$= [0 \quad I]\bar{x} \tag{7.2-14}$$

We shall explain how to estimate $\bar{x}$. The estimate will be of the form

$$\bar{x}_e = \begin{bmatrix} w + My \\ y \end{bmatrix} \tag{7.2-15}$$

with w generated as in (7.2-11) and M a certain matrix to be determined. Since

$$x_e = T^{-1} \bar{x}_e = T^{-1} \begin{bmatrix} I & M \\ 0 & I \end{bmatrix} \begin{bmatrix} w \\ y \end{bmatrix} \tag{7.2-16}$$

the linear transformation block in Figure 7.2-3 will be achieved through multiplication by the matrix

$$S \triangleq T^{-1} \begin{bmatrix} I & M \\ 0 & I \end{bmatrix} \tag{7.2-17}$$
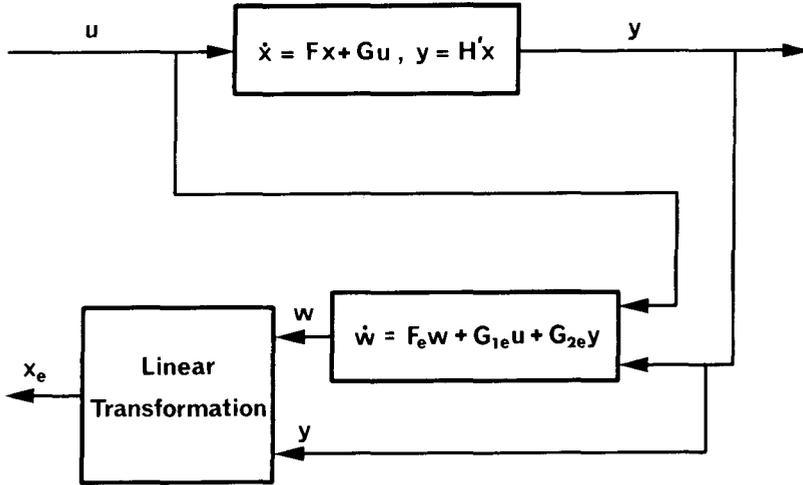
Figure 7.2-3   General structure of the reduced order estimator.

Now with appropriate partitioning,

$$\begin{bmatrix} \dot{\bar{x}}_1 \\ \dot{\bar{x}}_2 \end{bmatrix} = \begin{bmatrix} \overline{F}_{11} & \overline{F}_{12} \\ \overline{F}_{21} & \overline{F}_{22} \end{bmatrix} \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix} + \begin{bmatrix} \overline{G}_1 \\ \overline{G}_2 \end{bmatrix} u \tag{7.2-18}$$

$$y = \begin{bmatrix} 0 & I \end{bmatrix} \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix} = \bar{x}_2 \tag{7.2-19}$$

We claim that $[\overline{F}, \overline{H}]$ completely observable with $\overline{H}' = [0 \quad I]$ implies that $[\overline{F}_{11}, \overline{F}_{21}']$ is completely observable. To see this, assume that $[\overline{F}_{11}, \overline{F}_{21}']$ is not observable so that there exists a nonzero eigenvector v of $\overline{F}_{11}$ such that $\overline{F}_{11} v = \lambda v$, and $\overline{F}_{21} v = 0$. As a consequence

$$\begin{bmatrix} v \\ 0 \end{bmatrix} \neq 0 \qquad \begin{bmatrix} \overline{F}_{11} & \overline{F}_{12} \\ \overline{F}_{21} & \overline{F}_{22} \end{bmatrix} \begin{bmatrix} v \\ 0 \end{bmatrix} = \lambda \begin{bmatrix} v \\ 0 \end{bmatrix} \qquad \begin{bmatrix} 0 & I \end{bmatrix} \begin{bmatrix} v \\ 0 \end{bmatrix} = 0$$

so that $[\overline{F}, \overline{H}]$ is not observable. (Actually, the argument is reversible to establish that the converse claim is also true. Further, $[\overline{F}, \overline{H}]$ is detectable if and only if $[\overline{F}_{11}, \overline{F}_{21}']$ is detectable.)

Now with $[\overline{F}, \overline{H}]$ completely observable, giving $[\overline{F}_{11}, \overline{F}_{21}']$ completely observable as above, there exists some K_e such that $\dot{\xi} = [\overline{F}_{11} + K_e \overline{F}_{21}] \xi$ is asymptotically stable, perhaps with specified eigenvalues. Let $\bar{x}_{1e}$ be defined by

$$\dot{\bar{x}}_{1e} = (\overline{F}_{11} + K_e \overline{F}_{21}) \bar{x}_{1e} - K_e(\dot{y} - \overline{F}_{22} y - \overline{G}_2 u) + \overline{F}_{12} y + \overline{G}_1 u \tag{7.2-20}$$

Of course, there could be implementation difficulties in generating x_{1e}, because of the presence of $\dot{y}$. We will see how to circumvent this problem later. Let us first

observe that $\bar{x}_{1e}$ is indeed an estimate of $\bar{x}_1$. From (7.2-18) and (7.2-19), it is evident that

$$\dot{y} - \bar{F}_{22}y - \bar{G}_2u = \dot{\bar{x}}_2 - \bar{F}_{22}\bar{x}_2 - \bar{G}_2u = \bar{F}_{21}\bar{x}_1 \qquad (7.2\text{-}21)$$

It follows that

$$\frac{d}{dt}(\bar{x}_1 - \bar{x}_{1e}) = (\bar{F}_{11} + K_e\bar{F}_{21})(\bar{x}_1 - \bar{x}_{1e}) \qquad (7.2\text{-}22)$$

Consequently, $\bar{x}_1 - \bar{x}_{1e} \to 0$ at a rate determined by the eigenvalues of $F_e = \bar{F}_{11} + K_e\bar{F}_{21}$.

It remains to be seen how (7.2-20) can be adjusted so that $\bar{x}_{1e}$ is obtainable without introducing $\dot{y}$. Figure 7.2-4(a) illustrates (7.2-20). Simple manipulation in
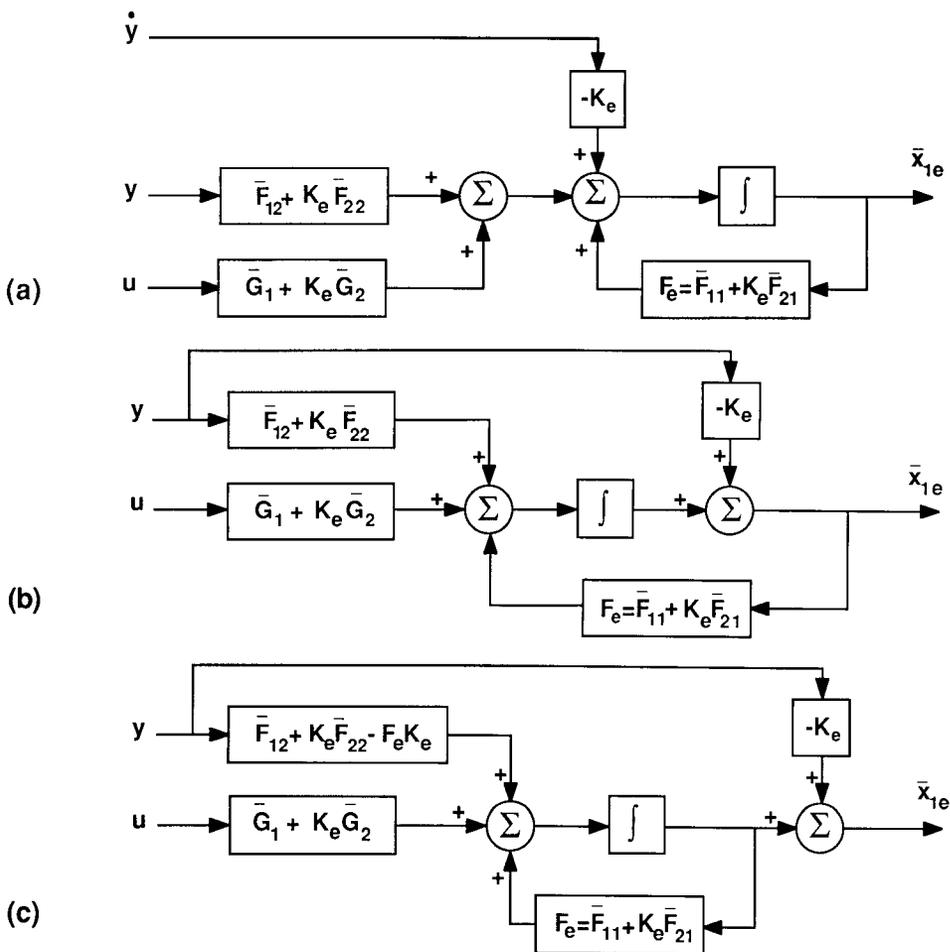
Figure 7.2-4   Generation of $\bar{x}_{1e}$.

the block diagram shows that Figure 7.2-4(c) is equivalent, at least from an input-output point of view. [Initialization of the integrator would have to be varied to secure identical outputs over $(0, \infty)$, rather than just asymptotically.] So the $\bar{x}_{1e}$ of Figure 7.2-4(a) and of Figure 7.2-4(c) may differ by a quantity decaying exponentially fast, at a rate determined by the eigenvalues of F_e. Consequently, (7.2-22) remains valid for the $\bar{x}_{1e}$ of Figure 7.2-4(c). In terms of equations, with $F_e = \bar{F}_{11} + K_e \bar{F}_{21}$,

$$\dot{w} = F_e w + (\bar{G}_1 + K_e \bar{G}_2)u + (\bar{F}_{12} + K_e \bar{F}_{22} - F_e K_e)\, y \qquad (7.2\text{-}23)$$

and

$$\bar{x}_{1e} = w - K_e y \qquad (7.2\text{-}24)$$

Then with

$$S = T^{-1}\begin{bmatrix} I & -K_e \\ 0 & I \end{bmatrix}$$

the overall estimator is depicted in Figure 7.2-5.
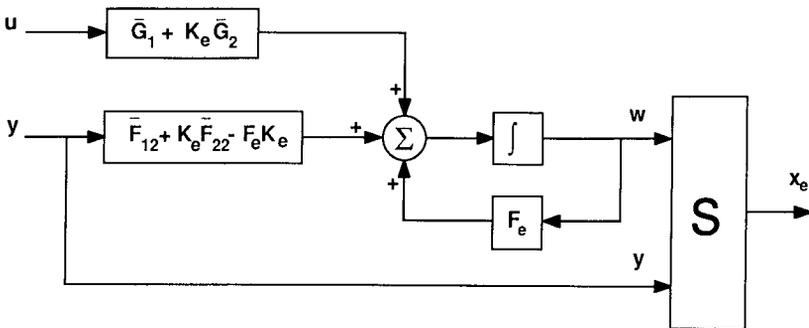
Figure 7.2-5   Overall construction of estimator.

We now give qualitative comments on the effect of noise in the reduced order estimator. As for estimators of the last section, it will be noise that limits the extent to which the eigenvalues of F_e can be made negative, that is, the rapidity with which x_e will approach x. But there is a second factor present in the reduced order estimator which is absent in the full order estimator. Suppose that the plant output y includes white noise. When a full order estimator is used, the resulting noise in x_e is band-limited, essentially because there is always integration between y and x_e. This is not the case with the reduced order estimators. In fact, there will be more non-bandlimited noise in x_e arising from that in y, because x_e is partly determined by a memoryless transformation on y. Consequently, one might expect the performance of a reduced order estimator of this section to be worse than that of a full order estimator.

To complement the above qualitative statements about performance in noise environments, let us remark about the situation when there is uncertainty in the

plant model. Without loss of generality let us assume a coordinate basis where y is a component of the state vector. The estimation of y by a full order estimator will be inaccurate if there is model uncertainty, but for the reduced order estimator there will be no such inaccuracy, since an estimate of the output y in this case is y itself.

As noted earlier, further estimator order reductions can be achieved if only linear functions of the state need to be estimated, such as when estimating a control signal $u = K'x$. Details for such designs are beyond the scope of this text; see, for example, [4] and its references.

We conclude this section by stressing that estimator design is a trade-off between complexity, robustness to plant uncertainty, noise filtering performance, and transient performance.

Main points of the section.   For nth order plants with states x, state estimators of order n can be constructed yielding state estimates x_e with x_e approaching x exponentially fast according to

$$\frac{d}{dt}(x - x_e) = F_e(x - x_e)$$

Here the eigenvalues of F_e completely determine the rate of convergence, and can be chosen arbitrarily in the design given observability, and stably given detectability.

For nth order plants with m independent outputs, reduced order estimators can be constructed of order $(n - m)$. In an appropriate coordinate basis, the plant outputs y comprise elements of the state vector which need not be estimated. The remaining elements comprising an $(n - m)$-vector x_1 are estimated by an $(n - m)^{\text{th}}$ order estimator as x_{1e} where convergence of x_{1e} to x_1 is given from

$$\frac{d}{dt}(x_1 - x_{1e}) = F_e(x_1 - x_{1e})$$

Again the eigenvalues F_e [now $(n - m) \times (n - m)$] can be prescribed arbitrarily given observability.

At the qualitative level, it is clear that in estimator design there is a trade-off in terms of complexity, robustness, noise filtering properties, and transient performance.

Problem 7.2-1.   Design an estimator for

$$\dot{x} = \begin{bmatrix} -2 & 1 \\ 0 & -2 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

$$y = \begin{bmatrix} 1 & 0 \end{bmatrix} x$$

such that $F + k_e h'$ has two eigenvalues of -5.

Problem 7.2-2.   If a system is not completely observable, by a coordinate basis change it may be put into the form

$$\dot{x} = \begin{bmatrix} F_{11} & 0 \\ F_{21} & F_{22} \end{bmatrix} x + \begin{bmatrix} G_1 \\ G_2 \end{bmatrix} u$$

$$y = [H_1' \quad 0] x$$

where $[F_{11}, H_1]$ is completely observable. Show how to estimate certain of the components of x arbitrarily quickly, and show why it is impossible to estimate the remainder arbitrarily fast. What can be done? Distinguish the two cases of detectability and lack of detectability.

Problem 7.2-3.   Devise a computational procedure for estimator design for the following system.

$$\dot{x} = \begin{bmatrix} 0 & 0 & -1 & | & 0 & 0 & 0 & 0 \\ 1 & 0 & -3 & | & 0 & 0 & 0 & 0 \\ 0 & 1 & -3 & | & 0 & 0 & 0 & 0 \\ - & - & - & + & - & - & - & - \\ 4 & 1 & 9 & | & 0 & 0 & 0 & -1 \\ 2 & 1 & 3 & | & 1 & 0 & 0 & -4 \\ 0 & 0 & 0 & | & 0 & 1 & 0 & -6 \\ 1 & 1 & 1 & | & 0 & 0 & 1 & -4 \end{bmatrix} x + \begin{bmatrix} 1 \\ 2 \\ 3 \\ -- \\ 1 \\ 1 \\ 2 \\ 1 \end{bmatrix} u$$

$$y = \begin{bmatrix} 0 & 0 & 1 & | & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & | & 0 & 0 & 0 & 1 \end{bmatrix} x$$

[*Hint:* Examine K_e of the form

$$K_e' = \begin{bmatrix} \alpha & \beta & \gamma & | & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & | & \delta & \epsilon & \varphi & \zeta \end{bmatrix}$$

where Greek letters denote elements that are, in general, nonzero.]

Problem 7.2-4.   Consider the scheme of Fig. 7.2-6. Write down state space equations for the individual blocks, and design a state estimator, assuming availability of the signals labeled u, y_1, and y_2. Discuss the situation where u and y_2 alone are available.

Problem 7.2-5.   Consider the scheme of Fig. 7.2-6. Write down state space equations for the individual blocks, and design a reduced order state estimator, assuming availability of the signals u, y_1 and y_2.

Problem 7.2-6.   Design a second-order estimator for

$$\dot{x} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix} x + \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} u$$

$$y = [1 \quad 1 \quad 1] x$$
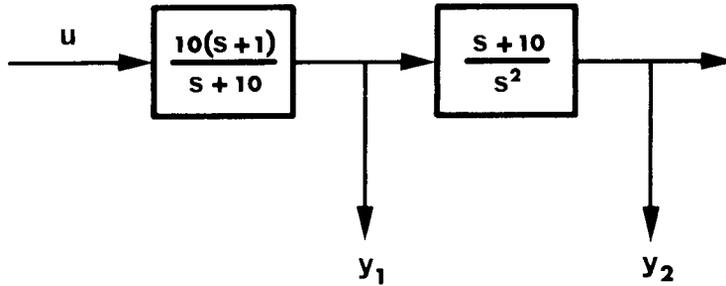
Figure 7.2-6   System for Problems 7.2-4 and 7.2-5.

Problem 7.2-7.   Consider the plant

$$\dot{x} = \begin{bmatrix} -2 & 0 \\ 1 & -2 \end{bmatrix} x + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u$$

$$y = \begin{bmatrix} 0 & 1 \end{bmatrix} x$$

and assume that adding on to y is white noise, that is, noise with power spectral density $M(\omega) = \sigma$. Design a first-order estimator with the F_e matrix equal to $-\alpha$ for a positive constant α. The estimator is driven by u, and y plus the noise. Compute the spectral density of noise in both components of x_e as a function of α. [*Hint:* Obtain the transfer functions relating y to each component of x_e. If these are $t_1(j\omega)$ and $t_2(j\omega)$, the spectral density of the noise in the components is $\sigma|t_1(j\omega)|^2$ and $\sigma|t_2(j\omega)|^2$.]

Problem 7.2-8.   Repeat Problem 7.2-7 with replacement of the estimator by a second-order estimator with $F + k_e h'$ possessing two eigenvalues at $-\alpha$. Compare the resulting noise densities with those of Problem 7.2-7.

7.3 STATISTICAL ESTIMATOR DESIGN (THE KALMAN–BUCY FILTER)

In this section, we touch upon an enormous body of knowledge perhaps best described by the term *filtering theory*. Much of this theory is summarized in the two books [5] and [6]. The particular material covered here is discussed in the important paper [7], although to carry out certain computations, we make use of a method discussed in [8] and [9].

The authors have presented a more complete treatment of optimal filtering, its properties, and applications in a companion text written for discrete-time systems [10].

Broadly speaking, we shall attempt to take quantitative consideration of the noise associated with measurements on a plant when designing an estimator. This means that the design of the estimator depends on probabilistic data concerning the

noise. We shall also aim at building the best possible estimator, where by "best possible" we mean roughly the estimator whose output is closest to what it should be, despite the noise. In other words, we shall be attempting to solve an *optimal filtering,* as distinct from optimal control, problem.

We warn the reader in advance of two things:

1. The treatment we shall give will omit many insights, side remarks, and so on in the interests of confining the discussion to a reasonable length.
2. The discussion will omit some details of mathematical rigor. We shall perform integration operations with integrands involving random variables, and the various operations, although certainly valid for deterministic variables, need to be proved to be valid for random variables. However, we shall omit these proofs. Moreover, we shall interchange integration and expectation operations without verifying that the interchanges are permissible.

In outline, we shall first describe the optimal estimation or filtering problem—that is, we shall describe the systems considered, the associated noise statistics, and a specific estimation task. Then, by the introduction of new variables, we shall convert the filtering problem into a deterministic optimal regulator problem, of the sort we have been discussing all through this book. The solution of this regular problem will then yield a mathematical solution of the filtering problem. A technique for physical implementation of the solution will then be found, leading to an estimator structure of the same form as that considered in Sec. 7.2, except that noise is present at certain points in the plant, and as a consequence in the estimator; see Fig. 7.3-1. (The figure does not show the structure of the plant, which is assumed
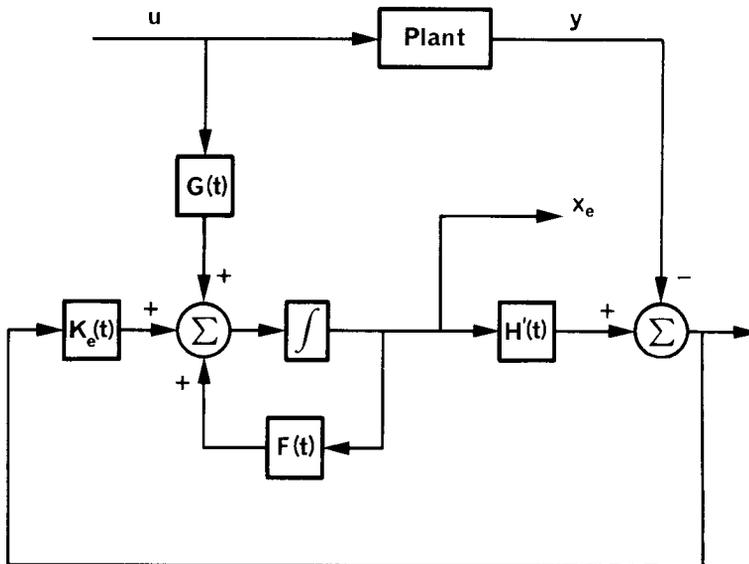
Figure 7.3-1   Structure of optimal estimator.

to be of the standard form $\dot{x} = Fx + Gu$, $y = H'x$, with additional terms in these equations representing noise, to be indicated precisely later.)

Since the structures of the optimal estimator and the estimator of Sec. 7.2 are the same, we can regard the present section as describing a technique for *optimally* designing one of the estimators of Sec. 7.2. The computations required for optimal design are a good deal more involved than those for the earlier design for a time-invariant single-output plant. However, for a multiple-output plant it is possible that the calculations to be presented here might even be simpler than the appropriate multiple-output plant generalization of the calculations of Sec. 7.2. The calculations here also extend to time-varying plants.

Description of plants and noise statistics.   The plants we shall consider are of the form

$$\frac{dx(t)}{dt} = F(t)x(t) + G(t)u(t) + v(t) \tag{7.3-1}$$

$$y(t) = H'(t)x(t) + w(t) \tag{7.3-2}$$

Here, $v(t)$ and $w(t)$ represent noise terms, which will be explained shortly. The dependence of F, G, and H on t, indicated in the equations, is to emphasize that, at least for the moment, these quantities are not necessarily time-invariant. However, for infinite-time interval problems, which are considered later, we shall specialize to the time-invariant case. Without further comment, we assume that the entries of $F(\cdot)$, $G(\cdot)$, and $H(\cdot)$ are all continuous. There is no restriction on the dimensions of u and y in these equations, and the subsequent calculations will not, in fact, be simplified significantly by an assumption that either u or y is scalar.

The properties of the noise terms will now be discussed. First note that the model of (7.3-1) and (7.3-2) assumes additive noise only, and it also assumes that noise is injected at only two points (see Fig. 7.3-2). The latter restriction is not so severe as might at first appear. Thus, for example, any noise entering with $u(t)$ [and passing through the $G(t)$ block] is equivalent to some other noise entering at the same point as $v(t)$.

In the case of both $v(t)$ and $w(t)$, the noise is assumed to be white, gaussian,
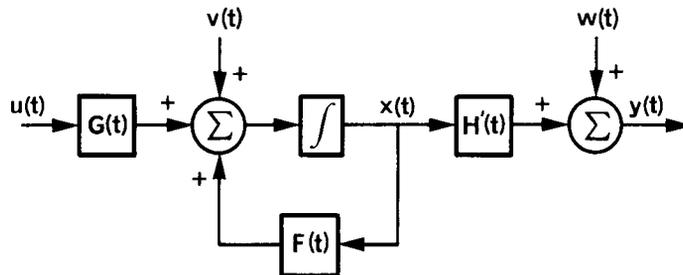
Figure 7.3-2 Plant with additive noise.

and to have zero mean. The first property implies that it is uncorrelated from instant to instant; if it were also stationary, it would have a constant power spectrum. The second property implies that all probabilistic information about the noise is summed up in the covariance of the noise—viz., $E[v(t)v'(\tau)]$ for $v(t)$ and likewise for $w(t)$. This convenient mathematical assumption is fortunately not without physical basis, for many naturally occurring noise processes are, indeed, gaussian, and such processes normally have zero mean. Therefore, in mathematical terms,

$$E[v(t)v'(\tau)] = Q(t)\delta(t - \tau) \qquad E[v(t)] \equiv 0 \qquad (7.3\text{-}3)$$

$$E[w(t)w'(\tau)] = R(t)\delta(t - \tau) \qquad E[w(t)] \equiv 0 \qquad (7.3\text{-}4)$$

for some matrices $Q(\cdot)$ and $R(\cdot)$, which we assume without further comment to have all entries continuous. The presence of the $\delta(t - \tau)$ term guarantees the whiteness property. Precisely because the quantities on the left sides of (7.3-3) and (7.3-4) are *covariances,* the matrices $Q(t)$ and $R(t)$ must be symmetric and non-negative definite. But we shall make the additional assumption that $R(t)$ is positive definite for all t. If this were not the case, there would be some linear combination of the outputs that was entirely noise free. Then, in an appropriate coordinate basis, one entry of the state vector would be known without filtering. As a result, the optimal estimator would not have the structure of Fig. 7.3-1 and a different optimal estimator design procedure would be required. We shall omit consideration of this difficult problem here; the interested reader should consult references [10] and [11].

Because it is often the case physically, we shall assume that the noise processes $v(t)$ and $w(t)$ are independent. This means that

$$E[v(t)w'(\tau)] = 0 \quad \text{for all } t \text{ and } \tau \qquad (7.3\text{-}5)$$

The final assumptions concern the initial state of (7.3-1). State estimation is assumed to commence at some time t_0, which may be minus infinity or may be finite. It is necessary to assume something about the state $x(t_0)$, and the assumptions that prove of use are that $x(t_0)$ is a gaussian random variable, of mean m, and covariance P_{e0}—that is,

$$E\{[x(t_0) - m][x(t_0) - m]'\} = P_{e0} \qquad E[x(t_0)] = m \qquad (7.3\text{-}6)$$

Furthermore, $x(t_0)$ is independent of $v(t)$ and $w(t)$—that is,

$$E[x(t_0)v'(t)] = E[x(t_0)w'(t)] = 0 \quad \text{for all } t \qquad (7.3\text{-}7)$$

Notice that the case where $x(t_0)$ has a known (deterministic) value is included: If $P_{e0} = 0$, then $x(t_0) = m$, rather than just $E[x(t_0)] = m$.

Let us now summarize the plant and noise descriptions.

Assumption 7.3-1

1. The plant is described by the equations (7.3-1) and (7.3-2).
2. The noise processes $v(t)$ and $w(t)$ are white, gaussian, of zero mean, and independent, and have known covariances [see Eqs. (7.3-3), (7.3-4), and (7.3-5)]. The matrix $R(t)$ in (7.3-4) is nonsingular.

3. The initial state of the plant is a gaussian random variable, of known mean and covariance [Eq. (7.3-6)]. It is independent of $v(t)$ and $w(t)$ [see Eq. (7.3-7)].

Notice that an assumption of detectability or observability of the plant has not been made. Such will be made later in considering infinite-time problems.

Although all the assumptions made often have some physical validity, there will undoubtedly be many occasions when this is not the case. Many associated extensions of the preceding problem formulation have, in fact, been considered, and optimal estimators derived, but to consider these would be to go well beyond the scope of this book.

Statement of the optimal estimation problem.   We shall now define the precise task of estimation. The information at our disposal consists of the plant input $u(t)$ and output $y(t)$ for $t_0 \le t \le t_1$, and the probabilistic descriptions of $x(t_0)$, $v(t)$, and $w(t)$. To obtain a first solution to the estimation problem, it proves convenient to make two temporary simplifications.

Temporary Assumption 7.3-2

The external input to the plant $u(t)$ is identically zero, and the mean m of the initial state $x(t_0)$ is zero

and

Temporary Assumption 7.3-3

The initial time t_0 is finite.

These assumptions will be removed when we have derived the optimal estimator for the special case implicit in the assumptions. With Temporary Assumptions 7.3-2 and 7.3-3 in force, the data at our disposal is simply the plant output $y(t)$ for $t_0 \le t \le t_1$, and our knowledge of the covariances of $v(\cdot)$, $w(\cdot)$, and $x(t_0)$.

The sort of estimate for which we shall aim is a minimum variance estimate—that is, we want to construct from a measurement of $y(t)$, $t_0 \le t \le t_1$, a certain vector, call it x_e such that

$$\text{Error variance} = E\{[x(t_1) - x_e(t_1)]'[x(t_1) - x_e(t_1)]\} \qquad (7.3\text{-}8)$$

is minimum. Then $x_e(t_1)$ is the minimum variance estimate of $x(t_1)$. It turns out that because all the random processes and variables are gaussian, and have zero mean, the vector x_e can be derived by linear operations on $y(t)$, $t_0 \le t \le t_1$—that is, there exists some matrix function $M(t; t_1)$, $t_0 \le t \le t_1$ such that

$$x_e(t_1) = \int_{t_0}^{t_1} M'(t; t_1) y(t) \, dt \qquad (7.3\text{-}9)$$

(This is a deep result which we shall not prove here.) The introduction of $M(\cdot; t_1)$ now allows the following formal statement of the optimal estimation problem.

> ***Optimal estimation problem.***   Given the plant of Eqs. (7.3-1) and (7.3-2), suppose that Assumptions 7.3-1, 7.3-2, and 7.3-3 hold. Then, for a fixed but arbitrary value $t_1 \geq t_0 > -\infty$, find a matrix function of time $M(t; t_1)$, $t_0 \leq t \leq t_1$, such that the index defined by (7.3-8) and (7.3-9)—a form of performance index—is minimized, the expectation being over all possible realizations of the two noise processes and over the random variable $x(t_0)$. [A minimum variance estimate of $x(t_1)$ is then provided by $\int_{t_0}^{t_1} M'(t; t_1) y(t) \, dt$.]

A further problem is to state how this estimate might be physically implemented to produce an on-line estimate $x_e(t_1)$ at time t_1 of $x(t_1)$, which is continuously updated, rather than a single estimate of the vector random variable $x(t_1)$ for fixed t_1.

Without further comment, we shall use the notation $M(t)$ and $M(\cdot)$ as shorthand for $M(t; t_1)$, provided no ambiguity occurs.

At this stage, let us pause to review what we have done so far, and what we shall do in the next part of this section. So far, we have accomplished the following:

1. We have described the plants considered, together with the noise associated with the plants.
2. We have posed a problem of estimating the state vector at a *particular* time [i.e., $x(t_1)$ for fixed t_1], using input and output measurements till t_1. The estimate is to be a minimum variance one.
3. We have posed the problem of constructing a device which at every time t_1 will produce an estimate of $x(t_1)$. That is, we have posed the problem of constructing an on-line estimate of $x(t_1)$.

In the remainder of this section, we shall do the following:

1. We shall show how the first of the preceding optimal estimation problems can be reformulated as an optimal control problem. *We caution the reader that the existence of such a reformulation is probably not intuitively reasonable, and the parameters appearing in the regulator problem are only suggested by hindsight.* Therefore, the reader will have to suppress such natural questions as "Why pick such-and-such set of system equations?" and be content that justification for the choice of such-and-such system equation lies in the fact that it works, somewhat surprisingly.
2. Using the optimal regulator problem reformulation, and our knowledge of the general regulator problem solution, we shall solve the specific optimal regulator problem.

3. Next, we shall do a natural thing—use the solution of the regulator problem to write down a solution of the optimal estimation problem associated with estimating $x(t_1)$ for specific t_1.

4. We shall then show how to obtain an estimate of $x(t_1)$ on line.

5. Elimination of restrictive assumptions, examples, and some extensions will then follow.

Reformulation of the optimal estimation problem as an optimal regulator problem.

To carry through the reformulation, we introduce a new square matrix function of time $Z(\cdot)$, of the same row dimension as $x(\cdot)$. This function is defined from $M(\cdot)$ via the equation

$$\frac{d}{dt} Z(t) = -F'(t) Z(t) + H(t) M(t) \qquad Z(t_1) = I. \tag{7.3-10}$$

Observe that Eq. (7.3-10) has a similar structure to the familiar vector state equation $\dot{x} = Fx + Gu$, with prescribed boundary condition $x(t_0)$ except that it involves a matrix, and that we shall be interested in the solutions of (7.3-10) for $t \leq t_1$ rather than for $t \geq t_1$.

We now rewrite the performance index (7.3-8) as a quadratic performance index involving $Z(\cdot)$ and $M(\cdot)$. From (7.3-1) and (7.3-10), we have

$$\frac{d}{dt}[Z'(t) x(t)] = \dot{Z}'(t) x(t) + Z'(t) \dot{x}(t)$$

$$= -Z'Fx + M'H'x + Z'Fx + Z'v$$

$$= M'y - M'w + Z'v$$

Integrating this equation from t_0 to t_1, using the boundary condition on Z, leads to

$$x(t_1) - Z'(t_0) x(t_0) = \int_{t_0}^{t_1} M'(t) y(t) \, dt - \int_{t_0}^{t_1} M'(t) w(t) \, dt + \int_{t_0}^{t_1} Z'(t) v(t) \, dt$$

or

$$x(t_1) - \int_{t_0}^{t_1} M'(t) y(t) \, dt = Z'(t_0) x(t_0) - \int_{t_0}^{t_1} M'(t) w(t) \, dt + \int_{t_0}^{t_1} Z'(t) v(t) \, dt$$

The next step is to essentially square each side, and take the expectation. Because of the independence of $x(t_0)$, $w(t)$, and $v(t)$, there results

$$E \left\{ \left[x(t_1) - \int_{t_0}^{t_1} M'(t) y(t) \, dt \right] \left[x(t_1) - \int_{t_0}^{t_1} M'(t) y(t) \, dt \right]' \right\}$$

$$= E[Z'(t_0) x(t_0) x'(t_0) Z(t_0)] + E \left[\int_{t_0}^{t_1} \int_{t_0}^{t_1} M'(t) w(t) w'(\tau) M(\tau) \, dt \, d\tau \right]$$

$$+ E \left[\int_{t_0}^{t_1} \int_{t_0}^{t_1} Z'(t) v(t) v'(\tau) Z(\tau) \, dt \, d\tau \right]$$

Now $Z(\cdot)$ and $M(\cdot)$, although unknown, are deterministic. So they can be pulled out of the expectations. The first summand on the right is then, by using (7.3-6),

$$E[Z'(t_0)\,x(t_0)\,x'(t_0)\,Z(t_0)] = Z'(t_0)\,E[x(t_0)\,x'(t_0)]\,Z(t_0)$$

$$= Z'(t_0)\,P_{e0}\,Z(t_0)$$

To handle the second summand, we must first interchange integration and expectation. The validity of this interchange ought to be established, but we ask the reader to accept it. Then, using (7.3-4), we have

$$E\left[\int_{t_0}^{t_1}\int_{t_0}^{t_1} M'(t)\,w(t)\,w'(\tau)\,M(\tau)\,dt\,d\tau\right]$$

$$= \int_{t_0}^{t_1}\int_{t_0}^{t_1} M'(t)\,E[w(t)\,w'(\tau)]\,M(\tau)\,dt\,d\tau$$

$$= \int_{t_0}^{t_1}\int_{t_0}^{t_1} M'(t)\,R(t)\delta(t-\tau)\,M(\tau)\,dt\,d\tau$$

$$= \int_{t_0}^{t_1} M'(t)\,R(t)\,M(t)\,dt$$

The third summand is similarly evaluated. Recalling (7.3-9), we see that

$$E\{[x(t_1) - x_e(t_1)][x(t_1) - x_e(t_1)]'\} = Z'(t_0)\,P_{e0}\,Z(t_0)$$

$$+ \int_{t_0}^{t_1} [M'(t)\,R(t)\,M(t) + Z'(t)\,Q(t)\,Z(t)]\,dt \qquad (7.3\text{-}11)$$

Now take the trace of both sides. Thus

$$E\{[x(t_1) - x_e(t_1)]'[x(t_1) - x_e(t_1)]\}$$

$$= \operatorname{tr}\left\{Z'(t_0)\,P_{e0}\,Z(t_0) + \int_{t_0}^{t_1} [M'(t)\,R(t)\,M(t) + Z'(t)\,Q(t)\,Z(t)]\,dt\right\} \qquad (7.3\text{-}12)$$

Now all quantities on the right of (7.3-12) are deterministic, with $M(\cdot)$ free to be chosen and $Z(\cdot)$ related to $M(\cdot)$ via (7.3-10). Therefore, the problem of choosing $M(\cdot)$ to minimize the left side of (7.3-12) is the same as the *deterministic* problem of choosing $M(\cdot)$ to minimize the right side of (7.3-12), subject to (7.3-10) holding.

Let us summarize the reformulation of the optimal estimation problem as follows.

> ***Reformulated optimal estimation problem.***   Suppose we are given the plant of Eqs. (7.3-1) and (7.3-2), and suppose that Assumptions 7.3-1, 7.3-2, and 7.3-3 hold. Let $t_1 > t_0$ be an arbitrary time. Find a function $M(t; t_1)$ for $t_0 \le t \le t_1$ such that the (deterministic) performance index (7.3-12) is minimized, subject to Eq. (7.3-10) holding.

We shall now comment upon this reformulated problem. Recall that R is positive definite for all t; Q is nonnegative definite for all t; and P_0, being the

covariance of a vector random variable, is also nonnegative definite. Therefore, the only factors distinguishing the problem of finding the function $M(\cdot)$, which minimizes (7.3-12) subject to (7.3-10), from the usual optimal regulator problem are (1) that the boundary condition of $Z(t)$ occurs at the final time t_1 rather than the initial time t_0, (2) that the term $Z'(t_0)P_{e0}Z(t_0)$ represents an initial rather than a final value term, (3) that the state variable $Z(\cdot)$ and control variable $M(\cdot)$ are matrices rather than vectors, and (4) that the index is the trace of a nonnegative definite symmetric matrix. In a sense, the problem we face here is a regulator problem for matrix state equations with time running backward.

Explicit solution of the optimal estimation problem.
We claim that the optimal "control" $M^*(\cdot)$ has the form

$$M^*(t) = R^{-1}(t)H'(t)P_e(t)Z(t) \tag{7.3-13}$$

where $P_e(t)$ is the solution of the Riccati equation

$$\dot{P}_e(t) = P_e(t)F'(t) + F(t)P_e(t) - P_e(t)H(t)R^{-1}(t)H'(t)P_e(t) + Q(t) \tag{7.3–14}$$

This is solved forward in time with boundary condition $P_e(t_0) = P_{e0}$. The proof of optimality is almost immediate. It turns out that one can reorganize (7.3-11) as

$$E\{[x(t_1) - x_e(t_1)][x(t_1) - x_e(t_1)]'\} = Z'(t_1)\,P_e(t_1)\,Z(t_1)$$

$$+ \int_{t_0}^{t_1} [M(t) - R^{-1}(t)\,H'(t)\,P_e(t)\,Z(t)]'R(t)\,[M(t) - R^{-1}(t)\,H'(t)\,P_e(t)\,Z(t)]\,dt$$

$$\tag{7.3-15}$$

Obviously, with $M(\cdot)$ free to be chosen and $Z(t_1)$ fixed (as the identity matrix), we make the trace of the left side (and indeed the matrix left side) as small as possible by selecting $M(t) = R^{-1}(t)H'(t)P_e(t)Z(t)$. To actually compute the optimal $M(t)$, we must solve $\dot{Z} = -F'Z + HR^{-1}H'P_eZ$ backwards from t_1 and use the solution in (7.3-13).

Notice that this choice of $M(\cdot)$ leads [because $Z(t_1) = I$] to

$$E\{[x(t_1) - x_e(t_1)][x(t_1) - x_e(t_1)]'\} = P_e(t_1) \tag{7.3-16}$$

The reorganization yielding (7.3-15) follows the pattern of that explored for the standard regulator in Problem 2.3-2. The fact that Z, M are matrices rather than vectors, as in the standard regulator problem, has only a trivial effect on the manipulations. Problem 7.3-7 seeks details.

The time reversal in the formulation of the regulator problem is also but a mild variation of the standard theory. Note that this time reversal reflects itself as sign changes in the time derivatives, in the Riccati equation, for example.

In the above argument, we have implicitly assumed that the Riccati equation (7.3-14) necessarily has a solution for $t \geq t_0$. It is possible to argue this by observing

that it is a Riccati equation associated with a regulator problem, albeit one in disguise. Alternatively, one can argue as follows. We have asserted that

$$Z'(t_0) P_{e0} Z(t_0) + \int_{t_0}^{t_1} [M'(t) R(t) M(t) + Z'(t) Q(t) Z(t)] \, dt$$

$$= Z'(t_1) P_e(t_1) Z(t_1) + \int_{t_0}^{t_1} [M(t) - R^{-1}(t) H'(t) P_e(t) Z(t)]' R(t)$$

$$\times [M(t) - R^{-1}(t) H'(t) P_e(t) Z(t)] \, dt$$

Noting that $Z(t_1) = I$ and that the second summand on the right is nonnegative, we see that it follows that the left side, for all $M(\cdot)$, is an upper bound for $P_e(t_1)$. Also, setting $M(t) = R^{-1}(t) H'(t) P_e(t) Z(t)$ shows that $P_e(t_1)$ is a nonnegative matrix; that is, 0 is a lower bound. [In fact, the identification of $P_e(t_1)$ as a covariance of the optimal estimate via (7.3-16) indicates also that $P_e(t_1) \geq 0$. Thus $\|P_e(t_1)\|$ for any finite $t_1 \geq t_0$ cannot become infinite. Hence the Riccati equation has no finite escape time.

To this point, we have explained how $x_e(t_1)$ can be computed, for fixed t_1. The steps are

1. Solve the Riccati equation (7.3-14) forward to t_1.
2. Solve $\dot{Z} = -F'Z + HR^{-1}H'P_e Z$ backwards to t_0 from $Z(t_1) = I$
3. Set $x_e(t_1) = \int_{t_0}^{t_1} M'(t; t_1) y(t) \, dt$ where $M(t, t_1) =$
   $R^{-1}(t) H'(t) P_e(t) Z(t; t_1)$. Here $Z(t; t_1)$ denotes the solution of $\dot{Z} = -F'Z + HR^{-1}H'P_e Z$ with $Z(t_1; t_1) = I$. Thus

$$x_e(t_1) = \int_{t_0}^{t_1} Z'(t; t_1) P_e(t) H(t) R^{-1}(t) H'(t) y(t) \, dt \qquad (7.3\text{-}17)$$

The question now to be addressed is: How may one obtain an on-line estimate $x_e(t_1)$? Notice that $Z(t; t_1)$ is a transition matrix. Hence $Z(t; t_1) = Z^{-1}(t_1; t)$ and

$$\frac{d}{dt_1} Z(t; t_1) = \frac{d}{dt_1} Z^{-1}(t_1; t)$$

$$= -Z^{-1}(t_1; t)\{[-F'(t_1) + H(t_1) R^{-1}(t_1) H'(t_1) P_e(t_1)] Z(t_1; t)\} Z^{-1}(t_1; t)$$

$$= -Z(t; t_1)[-F'(t_1) + H(t_1) R^{-1}(t_1) H'(t_1) P_e(t_1)]$$

So

$$\frac{d}{dt_1} Z'(t; t_1) = [F(t_1) - P_e(t_1) H(t_1) R^{-1}(t_1) H'(t_1)] Z'(t; t_1)$$

Now it is easy to differentiate (7.3-17). There results

$$\frac{d}{dt_1} x_e(t_1)$$

$$= \int_{t_0}^{t_1} [F(t_1) - P_e(t_1) H(t_1) R^{-1}(t_1) H'(t_1)] Z'(t; t_1) P_e(t) H(t) R^{-1}(t) H'(t) y(t) \, dt$$

$$+ P_e(t_1) H(t_1) R^{-1}(t_1) H'(t_1) y(t_1)$$

$$= [F(t_1) - P_e(t_1) H(t_1) R^{-1}(t_1) H'(t_1)] x_e(t_1)$$

$$+ P_e(t_1) H(t_1) R^{-1}(t_1) H'(t_1) y(t_1)$$

The initial condition, from (7.3-17), is $x_e(t_0) = 0$. We have therefore established the following.

Construction of an on-line optimal estimate.   Let $P_e(t)$ be the solution of the Riccati equation (7.3-14) with $P_e(t_0) = P_{e0}$. Then the optimal estimate $x_e(t)$ of $x(t)$ is defined by

$$\frac{d}{dt} x_e(t) = F(t) x_e(t) + K_e(t)[H'(t) x_e(t) - y(t)] \qquad x_e(t_0) = 0 \qquad (7.3\text{-}18)$$

where

$$K_e(t) = -P_e(t) H(t) R^{-1}(t) \qquad\qquad (7.3\text{-}19)$$

Moreover

$$E\{[x(t) - x_e(t)][x(t) - x_e(t)]'\} = P_e(t) \qquad\qquad (7.3\text{-}20)$$

Figure 7.3-3(a) shows a realization of this equation, with the identification

$$F_e(t) = F(t) + K_e(t) H'(t) \qquad\qquad (7.3\text{-}21)$$

Figure 7.3-3(b) shows a rearrangement. We reiterate that the equation is valid under the various provisos, including $u(t) \equiv 0$, $E[x(t_0)] = m = 0$ and t_0 finite.

We have now covered the basic optimal estimation problem, and our task for the remainder of this section consists in tidying up some loose ends, and presenting examples. Here, in order, is what we shall do.

1. We shall eliminate the restrictive Assumption 7.3-2, which required a zero plant input and zero mean initial state—the resulting change in the estimator is very minor. Then we shall present a simple example.

2. We shall show how to cope with the case $t_0 = -\infty$, drawing special attention to time-invariant problems, and including one example.

Elimination of Assumption 7.3-2.   We wish to consider situations where $u(t)$ can be nonzero, and $E[x(t_0)] = m$ can be nonzero. The effect of nonzero values of either of these quantities will be to leave the plant state covariance and output covariance the same as before, but to change the mean of the plant
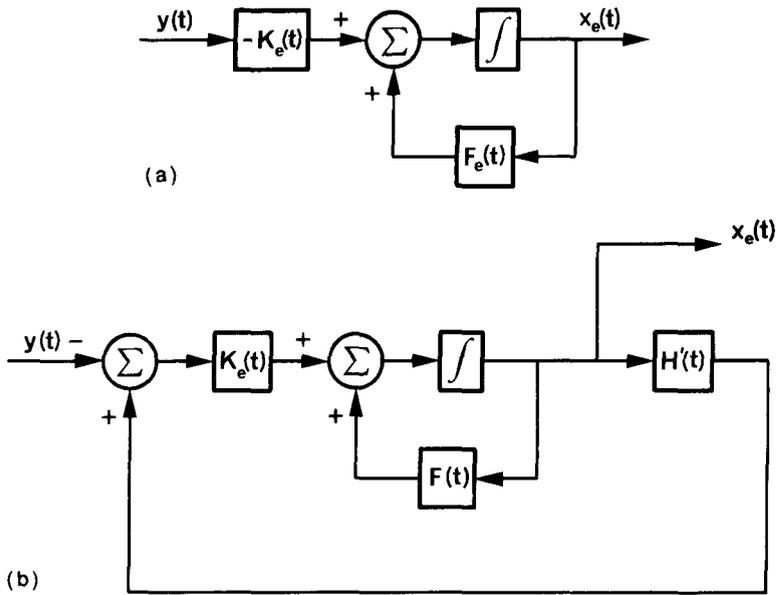
Figure 7.3-3   (a) First estimator structure $[u(t) \equiv 0, m \equiv 0]$; (b) Second estimator structure $[u(t) \equiv 0, m \equiv 0]$.

state and plant output. Without carrying out a direct derivation, we merely state the modified equations:

$$\frac{d}{dt} x_e(t) = F_e(t) x_e(t) - K_e(t) y(t) + G(t) u(t) \qquad x_e(t_0) = m \qquad (7.3-22)$$

or

$$\frac{d}{dt} x_e(t) = F(t) x_e(t) + K_e(t)[H'(t) x_e(t) - y(t)] + G(t) u(t) \qquad x_e(t_0) = m$$

$$(7.3-23)$$

where, as before, $K_e(t)$ is given by (7.3-19) and $P_e(t)$ satisfies the Riccati equation (7.3-14). Figure 7.3-4(a) shows plant and estimator according to (7.3-22) and Fig. 7.3-4(b) shows plant and estimator according to (7.3-23). Part of the estimator in Fig. 7.3-4(b) is enclosed in dotted lines, to emphasize the fact that the estimator is a model of the plant with additions.

Of course, the estimation structure obtained is just like that of the previous section. We find interesting also to note, using the plant equation (7.3-1) and estimation equation (7.3-22), that

$$\frac{d}{dt} [x(t) - x_e(t)] = F_e(t)[x(t) - x_e(t)] + v(t) + K_e(t) w(t) \qquad (7.3-24)$$

This error equation for $x(t) - x_e(t)$ is driven by zero mean processes, viz. $v(t)$ and $w(t)$, and $E[x(t_0) - x_e(t_0)] = 0$. As a consequence $E[x(t) - x_e(t)] = 0$ for all t; that
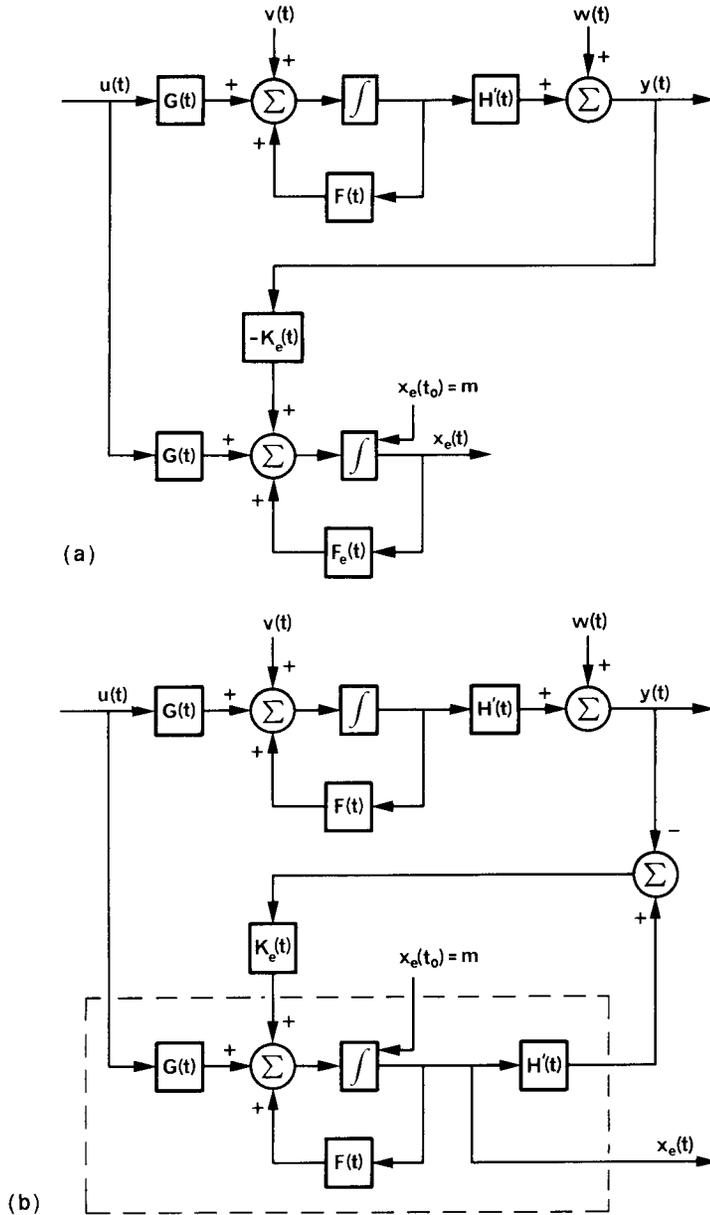
Figure 7.3-4   (a) Full plant estimator structure; (b) Redrawn estimator.

is, the estimate $x_e(t)$ is an unbiased estimate of $x(t)$. Also, for readers who understand how covariance matrices propagate, (7.3-24) provides insight into the optimality of $P_e(t)$. Suppose that $K_e(t)$ in (7.3-23) is replaced by a not necessarily

optimum matrix $K_e^\#$. Then we obtain, using (7.3-24), an error covariance $P_e^\#(t)$ for $E\{x(t) - x_e(t)][x(t) - x_e(t)]'\}$ defined by (see Appendix B)

$$\dot{P}_e^\# = (F + K_e^\# H') P_e^\# + P_e^\# (F' + HK_e'^\#) + Q + K_e^\# RK_e'^\#   P_e^\#(t_0) = P_{e0}   (7.3\text{-}25)$$

Now $P_e^\#$ and indeed tr $[P_e^\#(t)]$ is minimized for all t over $K_e^\#(\cdot)$, see Problem 7.3-8, when $K_e^\#$ is set equal to $-P_e HR^{-1}$. Then

$$\dot{P}_e^\# = (F - P_e HR^{-1}H') P_e^\# + P_e^\# (F' - HR^{-1}H'P_e) + Q + P_e HR^{-1}H'P_e$$

$$P_e^\#(t_0) = P_{e0}   (7.3\text{-}26)$$

As we know, $P_e^\# \equiv P_e$ then solves this equation, as comparison with (7.3-14) shows.

The argument just developed does not tell us that the optimal full-order estimator is the best estimator among all possible estimators for the stochastic signal model here, since the argument presumes an estimator structure of the form (7.3-23). This must come from the earlier main results.

To summarize, we have

Optimal Estimator Construction.   Given the plant equations

$$\dot{x}(t) = F(t)x(t) + G(t)u(t) + v(t)   (7.3\text{-}1)$$

$$y(t) = H'(t)x(t) + w(t)   x(t_0)   (7.3\text{-}2)$$

with initial time t_0 finite, suppose that $v(\cdot)$, $w(\cdot)$, $x(t_0)$ are independent and gaussian with

$$E[v(t)v'(\tau)] = Q(t)\,\delta(t - \tau)   E[v(t)] \equiv 0   (7.3\text{-}3)$$

$$E[w(t)w'(\tau)] = R(t)\,\delta(t - \tau)   E[w(t)] \equiv 0   (7.3\text{-}4)$$

$$E\{[x(t_0) - m][x(t_0) - m]'\} = P_{e0}   E[x(t_0)] = m   (7.3\text{-}6)$$

Then an on-line unbiased estimate $x_e(t)$ of $x(t)$ is provided at time t by the arrangement of Fig. 7.3-4 with equations (7.3-22) or (7.3-23) where the filter gain K_e and estimation system matrix F_e are given from

$$K_e(t) = -P_e(t)H(t)R^{-1}(t)   (7.3\text{-}19)$$

and

$$F_e(t) = F(t) + K_e(t)H'(t)   (7.3\text{-}21)$$

where $P_e(t)$ is the solution of the matrix Riccati equation

$$\dot{P}_e(t) = P_e(t)F'(t) + F(t)P_e(t) - P_e(t)H(t)R^{-1}(t)H'(t)P_e(t) + Q(t)   (7.3\text{-}14)$$

which is solved forwards in time with boundary condition $P_e(t_0) = P_{e0}$. The matrix $P_{e0}(t)$ exists for all $t \geq t_0$, and is symmetric nonnegative definite, being the minimum error covariance $E\{[x(t) - x_e(t)][x(t) - x_e(t)]'\}$.

By way of example, we consider the following problem. The plant is time-invariant, with transfer function $1/(s + 1)$, and there is input noise of covariance

$1\delta(t - \tau)$, and additive noise at the output of covariance $2\delta(t - \tau)$. At time zero, the initial state has a known value of 1. The problem is to design an optimal state estimator. In state-space terms, the plant equations are

$$\dot{x} = -x + u + v \qquad x(0) = 1$$

$$y = x + w \tag{7.3-27}$$

with $E[v(t)v(\tau)] = Q\delta(t - \tau) = 1\delta(t - \tau)$ and $E[w(t)w(\tau)] = R\delta(t - \tau) = 2\delta(t - \tau)$. The initial state covariance matrix P_0 is zero. From Eq. (7.3-14), we have

$$\dot{P}_e = -2P_e - \tfrac{1}{2}P_e^2 + 1$$

This equation yields

$$\int_0^{P(t)} \frac{dP_e}{P_e^2 + 4P_e - 2} = -\frac{1}{2} \int_0^t dt$$

whence

$$\left[\frac{1}{2\sqrt{6}} \ln \frac{P_e + 2 - \sqrt{6}}{P_e + 2 + \sqrt{6}} \right]_0^{P(t)} = -\frac{1}{2}t$$

or

$$P_e(t) = \frac{(\sqrt{6} - 2)[1 - \exp(-\sqrt{6}\,t)]}{1 + [(\sqrt{6} - 2)/(\sqrt{6} + 2)] \exp(-\sqrt{6}\,t)} \tag{7.3-28}$$

The gain matrix $K_e(t)$ for the optimal estimator, here a scalar, is given from (7.3.19) as $-\tfrac{1}{2}P_e(t)$. Figure 7.3-5 shows the plant and estimator. Notice that although the plant is time-invariant and $v(\cdot)$, $w(\cdot)$ are stationary, the finite initial time leads to a time-varying estimator (7.3.19) as $-\tfrac{1}{2}P_e(t)$.

The following example is discussed in [12]. The position and velocity of a satellite are to be estimated; the measurements available are, first, a position measurement including a constant random variable measurement error, accounting for drift, and the like. The motion of the satellite is linear and one-dimensional, and there is a constant gaussian random acceleration, independent of the measurement error in the acceleration measurement.

With x_1 denoting position and x_2 velocity, the equations of motion are

$$\dot{x}_1 = x_2$$

$$\dot{x}_2 = a$$

where a is the constant acceleration and is a gaussian random variable. Now the problem data concerning measurements and noise do not immediately allow construction of the standard system equations; therefore, we proceed as follows. The measurement of acceleration is $y_2 = a + b$, where b is a constant gaussian random variable. Then

$$\dot{x}_2 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_b^2} y_2 + \left(\frac{\sigma_b^2}{\sigma_a^2 + \sigma_b^2} y_2 - b \right)$$
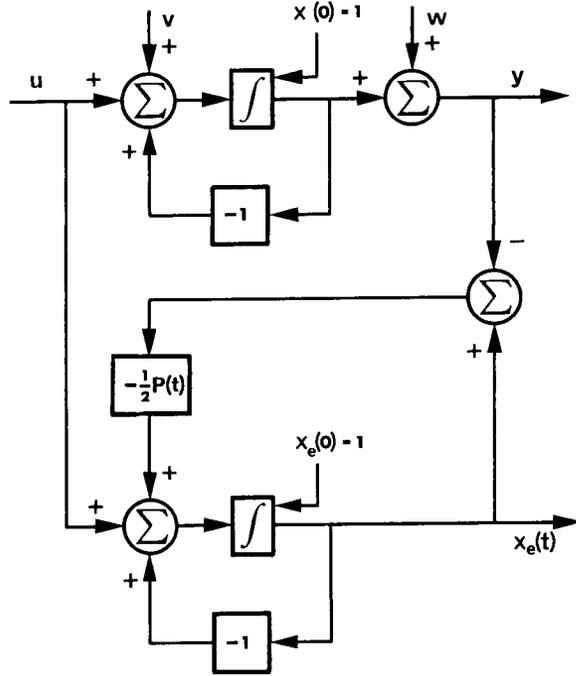
Figure 7.3-5   Plant and optimal estimator.

where $\sigma_a = E[a^2]$ and $\sigma_b = E[b^2]$. Observe that $[\sigma_a^2/(\sigma_a^2 + \sigma_b^2)]y_2$ and $[\sigma_b^2/(\sigma_a^2 + \sigma_b^2)]y_2 - b$ are independent, for

$$E\left\{\left[\frac{\sigma_a^2}{\sigma_a^2 + \sigma_b^2}\, y_2\right]\left[\frac{\sigma_b^2}{\sigma_a^2 + \sigma_b^2}\, y_2 - b\right]\right\} = \frac{\sigma_a^2\,\sigma_b^2}{(\sigma_a^2 + \sigma_b^2)^2}\, E(y_2^2) - \frac{\sigma_a^2}{\sigma_a^2 + \sigma_b^2}\, E(y_2\, b)$$

$$= \frac{\sigma_a^2\,\sigma_b^2}{(\sigma_a^2 + \sigma_b^2)^2}\,(\sigma_a^2 + \sigma_b^2) - \frac{\sigma_a^2}{\sigma_a^2 + \sigma_b^2}\,\sigma_b^2$$

$$= 0$$

So this equation is of the form

$$\dot{x}_2 = u + x_3$$

where u is known and independent of x_3. Moreover, x_3 is a constant gaussian random variable whose variance is easily checked to be $[\sigma_a^2\sigma_b^2/(\sigma_a^2 + \sigma_b^2)]$, a quantity that we shall call ρ from now on.

The full system equations thus become

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} u$$

$$y = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} x + w$$

We assume that u is known and that the initial values of x_1 and x_2 are known, whereas $E[x_3^2(0)] = \rho$, and $E[w(t)w(\tau)] = r\delta(t - \tau)$.

The Riccati equation becomes

$$\dot{P}_e = P_e \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} P_e - P_e \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} r^{-1}[1 \quad 0 \quad 0]P_e$$

with initial condition

$$P_e(0) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \rho \end{bmatrix}$$

The fact that the right side of the differential equation for P_e contains no term independent of P_e underpins the analytic solvability of the equation. The solution to this equation turns out to be

$$P_e(t) = \frac{r}{t^5/20 + r/\rho} \begin{bmatrix} \dfrac{t^4}{4} & \dfrac{t^3}{2} & \dfrac{t^2}{2} \\ \dfrac{t^3}{2} & t^2 & t \\ \dfrac{t^2}{2} & t & 1 \end{bmatrix}$$

and the optimal gain vector is

$$k_e' = -\left[\frac{t^4}{4(t^5/20 + r/\rho)} \quad \frac{t^3}{2(t^5/20 + r/\rho)} \quad \frac{t^2}{2(t^5/20 + r/\rho)} \right]$$

In the limit as $t \to \infty$, $P_e(t) = 0$ and $k_e(t) = 0$. Essentially what happens is that x_3 is exactly identified at $t \to \infty$. Since $x_1(0)$ and $x_2(0)$ are both known, this means that x_1 and x_2 become exactly identified as $t \to \infty$. In consequence, the error covariance approaches zero—that is, $P_e(t) \to 0$. Simultaneously, the noisy measurements become of no use, and thus $k_e(t) \to 0$. (This is actually not a good property for a practical filter to have; see, for example [10].)

Duality.   The reader will by now have observed the close parallels between the optimal estimator and control problems. They are often termed duals of one another. The duality can be summed up in the following way:

Estimator-Regulator Duality.   Consider a regulator problem defined by $F(t)$, $G(t)$, $Q(t)$, $R(t)$, terminal time t_1, and initial time t_0. Let $P(t)$ be the associated Riccati equation solution, and $K(t)$ the control law gain. Define matrices $\hat{F}(t) = F'(-t)$, $\hat{H}(t) = G(-t)$, $\hat{Q}(t) = Q(-t)$, $\hat{R}(t) = R(-t)$, and define a filtering problem using $\hat{F}$, $\hat{H}$, $\hat{Q}$, $\hat{R}$ with initial time $-t_1$. Let $\hat{P}_e(t)$ be the associated Riccati equation solution, and $K_e(t)$ the associated filter gain. Then

$$\hat{P}_e(t) = P(-t) \qquad K_e(t) = K(-t) \qquad\qquad (7.3\text{-}29)$$

How does this come about? Observe that

$$\dot{P}_e = \hat{P}_e(t)\hat{F}'(t) + \hat{F}'(t)\hat{P}_e(t) - \hat{P}_e(t)\hat{H}(t)\hat{R}^{-1}(t)\hat{H}'(t)\hat{P}_e(t) + \hat{Q}(t)$$

$$\hat{P}_e(-t_1) = 0$$

Since this holds for all t, it holds when we replace t by $-t$. Thus

$$-\frac{d}{dt}\hat{P}_e(-t) = \hat{P}_e(-t)\hat{F}'(-t) + \hat{F}(-t)\hat{P}_e(-t)$$

$$-\hat{P}_e(-t)\hat{H}(-t)\hat{R}^{-1}(-t)\hat{H}'(-t)\hat{P}_e(-t) + \hat{Q}(-t) \qquad \hat{P}_e(-t_1) = 0$$

that is,

$$-\frac{d}{dt}\hat{P}_e(-t) = \hat{P}_e(-t)F(t) + F'(t)\hat{P}_e(-t)$$

$$-\hat{P}_e(-t)G(t)R^{-1}(t)G'(t)\hat{P}_e(-t) + Q(t) \qquad \hat{P}_e(-t_1) = 0$$

Obviously, $\hat{P}_e(-t) = P(t)$ satisfies this equation. Also, $K_e(t) = -\hat{P}_e(t)\hat{H}(t)\hat{R}^{-1}(t) = -P(-t)G(-t)R^{-1}(-t) = K(-t)$.

It is not hard to check a further consequence of duality, namely that $[F(t), G(t)]$ is controllable at all times if and only if $[\hat{F}(t), \hat{H}(t)]$ is observable at all times. This is done as follows. From the relation between $F(t)$ and $\hat{F}(t)$, one can establish that the associated transition matrices satisfy $\hat{\Phi}(t, s) = \Phi'(-s, -t)$ (see Problem 7.3-9). It follows that

$$\int_{t_1}^{t_2} \hat{\Phi}'(s, t_1)\hat{H}(s)\hat{H}'(s)\hat{\Phi}(s, t_1)\, ds$$

$$= \int_{t_1}^{t_2} \Phi(-t_1, -s)G(-s)G'(-s)\Phi'(-t_1, -s)\, ds$$

$$= \int_{-t_2}^{-t_1} \Phi(-t_1, -s)G(-s)G'(-s)\Phi'(-t_1, -s)\, d(-s)$$

Positive definiteness of the first quantity is equivalent to $[\hat{F}, \hat{H}]$ being observable at t_1. Positive definiteness of the second is equivalent to $[F, G]$ being controllable at $-t_2$.

Initial times in the infinite past.

The interpretation of $P(t)$ as the error covariance will now be used in a discussion of the estimation problem for $t_0 = -\infty$. At this stage, therefore, we drop Assumption 7.3-3, but introduce a new one.

Assumption 7.3-4

For all t, the pair $[F(t), H(t)]$ is completely observable, or in the case F, H constant, the pair $[F, H]$ is detectable.

To get a rough idea of the reason for this assumption, consider a time-invariant plant having all unobservable, unstable states. Suppose also that at time $t_0 = -\infty$, the initial state of the plant has mean zero. Since any estimator can deduce absolutely no information about the plant state from the available measurements, the only sensible estimate of the plant state is $x_e(t_1) = 0$. Now the covariance of the plant state vector will be infinite at any finite time, and therefore so will the error covariance. Consequently, if $P_e(t)$ retains the significance of being the error covariance in the $t_0 = -\infty$ case, it is infinite. The gain K_e of the optimal filter will certainly have some, if not all, entries infinite also. The purpose of Assumption 7.3-4 is to prevent this sort of difficulty. In fact, under this assumption, and, to simplify matters, under the assumption $P_{e0} = 0$, we claim that $P_e(t)$ exists as the solution of (7.3-14) (and is finite, of course) when (7.3-14) has the boundary condition

$$\lim_{t_0 \to -\infty} P_e(t_0) = 0$$

This result is an immediate consequence of the estimator–regulator duality, and the known results for the regulator.

One feature that distinguishes the filter from the regulator should be noted. It is an issue in the time-varying and time-invariant case, but most easily understood for the latter. Suppose Re $\lambda_i(F) > 0$ for some F and $[F, G]$ is completely controllable. Then $E[x(t) x'(t)]$ will be infinite when $t_0 \to -\infty$. We are asking the estimator to estimate a variable with an infinitely large variance! It follows from this remark that one does *not* normally contemplate $t_0 \to -\infty$ in conjunction with an unstable signal model.

If F and H are constant (i.e., the plant is time-invariant) and if Q and R are constant (i.e., the noise is stationary), it follows (again, by examining the dual regulator problem) that the value of $P_e(t)$ obtained by letting $t_0 \to -\infty$ is independent of time, and can therefore be computed by evaluating $\lim_{t \to \infty} P_e(t)$ where $P_e(t)$ satisfies (7.3-14) with the initial condition $P_e(0) = 0$. Also the constant matrix P_e is a solution of the quadratic matrix equation

$$P_e F' + F P_e - P_e H R^{-1} H' P_e + Q = 0 \qquad (7.3\text{-}30)$$

The gain of the optimal estimator is then constant, being given by

$$K_e = -P_e H R^{-1} \qquad (7.3\text{-}31)$$

and if G is constant, the optimal estimator is a time-invariant system:

$$\dot{x}_e = F x_e + G u + K_e[H' x_e - y] \qquad (7.3\text{-}32)$$

However, this equation is of little importance from the practical point of view unless it represents an asymptotically stable system or, equivalently, $F_e + K_e H'$ has eigenvalues with negative real parts. The way to ensure this is to require the following:

Assumption 7.3-5

With constant F and Q, and D any matrix such that $DD' = Q$, the pair $[F, D]$ is stabilizable.

One way to prove the asymptotic stability of the optimal estimator under this assumption (with also G, H, and R constant, $t_0 = -\infty$) is to examine the associated regulator problem and to apply the results of earlier chapters. The assumption serves to provide a detectability assumption for the regulator problem, which guarantees asymptotic stability for the optimal regulator. This carries over to the optimal estimator.

For the time-invariant plant, under Assumptions 7.3-4 and 7.3-5, $P_e = \lim_{t \to \infty} P_e(t)$ is the only solution of (7.3-30) that is nonnegative definite, and the only solution of (7.3-30) for which $F - P_e H R^{-1} H'$ has all eigenvalues in Re $[s] < 0$. In case Assumption 7.3-5 is strengthened to requiring $[F, D]$ reachable (which roughly says that the input noise affects all states), P_e is positive definite (which states that no state can be estimated with zero error).

Since the time-invariant problem is so important, we shall summarize the results.

> **The time-invariant optimal estimator.**   Given the plant equations (7.3-1) and (7.3-2), suppose that F, G, and H are constant. Suppose also that Assumptions 7.3-1 and 7.3-4 hold, that Q and R are constant, and that $t_0 = -\infty$. The matrix P_e, which is the error covariance, is constant and satisfies (7.3-30); it may be computed by taking a limiting boundary condition of the form $\lim_{t_0 \to -\infty} P_e(t_0) = 0$ for (7.3-14) and evaluating $P_e(t)$ for any t, or by evaluating $\lim_{t \to \infty} P_e(t)$ where $P_e(t)$ satisfies (7.3-14) but with boundary condition $P_e(0) = 0$. The optimal estimator is time-invariant. Moreover, if Assumption 7.3-5 holds, the optimal estimator is asymptotically stable.

The result is, of course, closest to the ideas of Sec. 7.2. There, we confined our attention to time-invariant systems, and time-invariant estimators. There, also, the estimators were asymptotically stable, as a result of the complete observability assumption.

To illustrate the nature of the optimal estimator, let us consider a first-order example. Suppose

$$\dot{x} = fx + v, \qquad y = hx + w$$

with associated noise intensities $r > 0$, $q \geq 0$. The transfer function is $h/(s - f)$. The steady state Riccati equation is $2p_e f - p_e^2 h^2 r^{-1} + q = 0$ and the positive solution of this equation is

$$p_e = \frac{fr}{h^2} + \sqrt{\frac{f^2 r^2}{h^4} + \frac{qr}{h^2}}$$

The filter gain is

$$k_e = -p_e h r^{-1} = -h^{-1}[f + \sqrt{f^2 + h^2 q r^{-1}}]$$

and the filter is

$$\dot{x}_e = -\sqrt{f^2 + h^2 q r^{-1}}\, x_e - k_e y$$

Notice that the filter gain depends only on the process noise to measurement noise ratio q/r, rather than on the individual covariance levels, although of course the error covariance $\bar{P}_e$ depends also on r.

For the case when there is zero process noise ($q = 0$), then

$$\dot{x}_e = -|f|x_e - k_e y$$

and the estimator pole is the reflection of the plant pole into its image in the left half-plane, if it is not there already. As (q/r) increases, the filter pole moves left and approaches the plant zero at $-\infty$ as $(q/r) \to \infty$. Correspondingly, the filter gain magnitude $|k_e|$ increases to infinity. Notice that if the plant is stable and $q = 0$, then $\bar{P}_e = 0$, $k_e = 0$; a stable plant and $q = 0$ imply $x(t) \equiv 0$ and the best estimate is $x_e(t) \equiv 0$; that is, one should not use the measurements. When $f = 0$, $q = 0$ then clearly $[f, q^{1/2}]$ is not detectable and the estimator is unacceptably only neutrally stable. In such a situation, the idealized process model should be made more realistic by including a process noise term, even if very small.

It is clear that the noise intensities q, r are akin to the state and control weighting parameters in the dual control problem, in that adjustments to these allow adjustments to a filter design. In practice, the noise characteristics may not be known precisely, but a desired filter bandwidth is known. Then q, r can be used as tuning parameters to achieve the desired bandwidths. If in a first-cut design the bandwidth is too high (low), then (q/r) should be decreased (increased). There are obvious qualitative implications for the case of high-order multivariable filters.

Let us consider a further example. We suppose white noise of covariance $q\delta(t - \tau)$ is the input to a plant of transfer function $1/[s(s + 1)]$, starting from time $t_0 = -\infty$. There is additive output noise of covariance $r\delta(t - \tau)$. Thus, we take

$$F = \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix} \qquad g = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad h' = \begin{bmatrix} 0 & 1 \end{bmatrix}$$

Figure 7.3-6(a) shows the scheme. To put this in the standard form, we set

$$E[v(t)v'(\tau)] = \begin{bmatrix} q & 0 \\ 0 & 0 \end{bmatrix} \delta(t - \tau)$$

The input noise is not available to the estimator as an input. Denoting by p_{ij} the entries of the P_e matrix, the quadratic equation (7.3-30) becomes

$$\begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix}$$

$$-\frac{1}{r}\begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix}\begin{bmatrix} 0 & 1 \end{bmatrix}\begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} + \begin{bmatrix} q & 0 \\ 0 & 0 \end{bmatrix} = 0$$

or
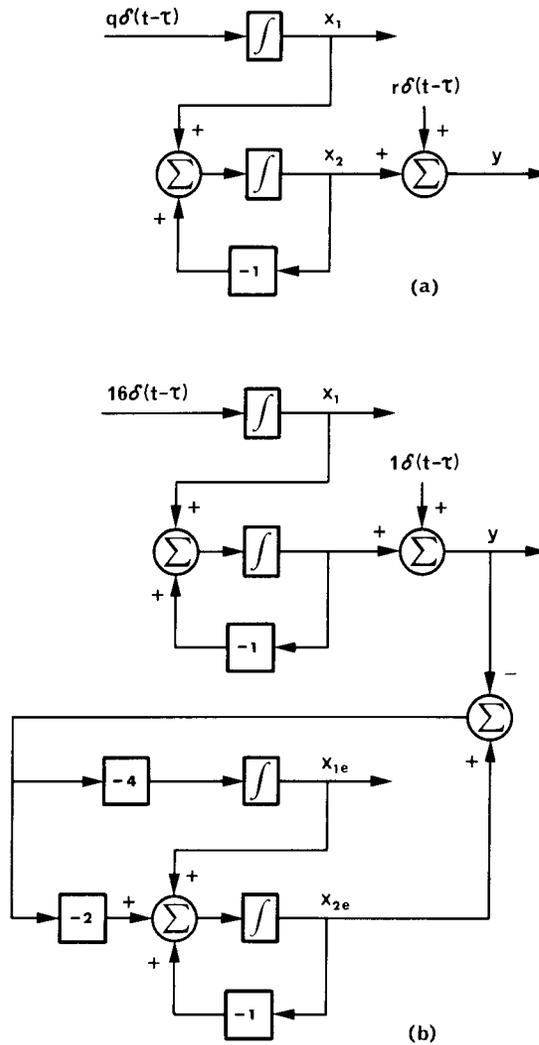
Figure 7.3-6   A specific plant and state estimator.

$$-\frac{1}{r}p_{12}^2 + q = 0$$

$$p_{11} - p_{12} - \frac{1}{r}p_{12}p_{22} = 0$$

$$2p_{12} - 2p_{22} - \frac{1}{r}p_{22}^2 = 0$$

It follows that $p_{12} = \sqrt{rq}$ from the first equation, $p_{22} = r[\sqrt{1 + 2\sqrt{q/r}} - 1]$ from the third equation, and $p_{11} = \sqrt{rq + 2q}$ from the second equation. Alternative solutions are ruled out as not leading to positive definite P_e. The estimator gain from Eq. (7.3-31) is

$$k_e = \begin{bmatrix} -\sqrt{q/r} \\ -\sqrt{1 + 2\sqrt{q/r}} + 1 \end{bmatrix}$$

If, for example, $q = 16$, $r = 1$, we have $k_e' = [-4 \quad -2]$. The matrix F_e is $F + k_e h'$, or

$$\begin{bmatrix} 0 & -4 \\ 1 & -3 \end{bmatrix}$$

It is readily checked to have eigenvalues with negative real parts. The plant and estimator are shown in Fig. 7.3-6(b) for the case $q = 16$, $r = 1$.

Spectral factorization and the innovations process.

Consider the time-invariant signal model and filter under Assumptions 7.3-4, 7.3-5. The steady state Riccati equation (7.3-30) gives, dualizing the result for regulators, the spectral factorization

$$R + H'(j\omega I - F)^{-1}Q(-j\omega I - F)')^{-1}H$$

$$= [I - H'(j\omega I - F)^{-1}K_e] R[I - K_e'(-j\omega I - F')^{-1}H] \qquad (7.3\text{-}33)$$

The left side is the spectrum $\Phi_{yy}(j\omega)$ of the measurements y arising from the signal model (7.3-1) through (7.3-5) in the case $t_0 = -\infty$. The spectral factor on the right side has the property that $[I - H'(sI - F)^{-1}K_e]^{-1} = [I + H'(sI - F - K_eH')^{-1}K_e]$ is asymptotically stable. As for the regulator problem, it is in effect uniquely specified by this requirement.

The spectral factorization identity allows a quick derivation of an important property of the random process

$$v(t) = y(t) - H'(t)x_e(t) \qquad (7.3\text{-}34)$$

which is termed the *innovations process*. This process can be thought of as the error in estimating $y(t)$ using $y(s)$ for $s < t$, or as the new information in $y(t)$ not carried by that in $y(s)$, $s < t$ (hence the name innovations). The property is

$$E[v(t)v'(s)] = R(t)\delta(t - s) \qquad (7.3\text{-}35)$$

So the innovations process is white, with the noise intensity as the measurement noise. This property is much more easily seen in the time-invariant case, and we shall content ourselves with a demonstration for that case. The innovation process is illustrated in Fig. 7.3-7. A deterministic external input to the plant plays no real role in the estimation process (other than, in effect, to reset the mean), and so is suppressed. Now the transfer function matrix from y to v is $I + H'(sI - F - K_eH')^{-1}K_e$.
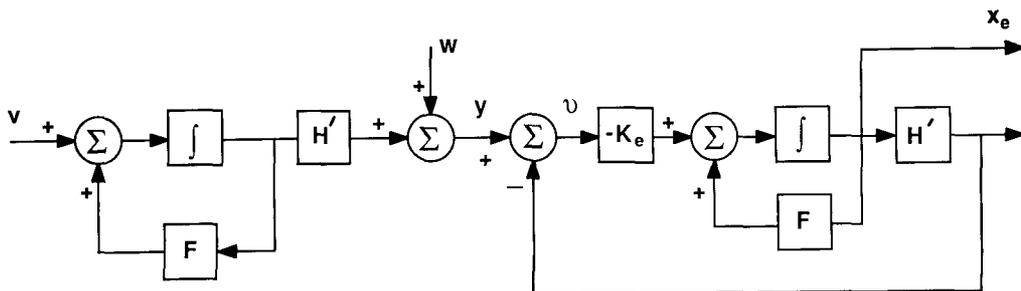
Figure 7.3-7   The innovations process ν is white.

It follows that the spectrum of ν is

$$\Phi_{\nu\nu}(j\omega) = [I + H'(j\omega I - F - K_e H')^{-1} K_e]\Phi_{yy}(j\omega)[I + K_e'(-j\omega I - F' - HK_e')^{-1}H]$$

$$= R \qquad (7.3\text{-}36)$$

by (7.3-33). This is the frequency domain equivalent of (7.3-35).

One way of testing whether an estimator is optimal is to measure the innovations process and see if it is white [10]. Note too that the spectral factorization identity and the innovations process allow ready connection with Wiener filtering concepts; see [10].

Discrete-time estimator.   The corresponding discrete-time optimal estimator algorithm is now summarized without a separate derivation. For further details, see [10].

Consider the signal model

$$x(t+1) = F(t)x(t) + G(t)u(t) + v(t) \qquad (7.3\text{-}37)$$

$$y(t) = H'(t)x(t) + w(t) \qquad (7.3\text{-}38)$$

where the process $v(\cdot)$, $w(\cdot)$ are white, with $v(t)$, $w(s)$, and $x(t_0)$ independent and zero mean for all t, s, with $E[x(t_0)x'(t_0)] = \Sigma_0$, $E[v(t)v'(\tau)] = Q(t)\delta(t-\tau)$, $E[w(t)w'(\tau)] = R(t)\delta(t-\tau)$ where $\delta(t-\tau) = 1$ if $t = \tau$ and is zero otherwise.

Two forms of the optimal filter are in use. One makes explicit the optimal one-step-ahead prediction $x_e(t|t-1)$, which minimizes the covariance

$$\Sigma(t|t-1) = E\{[x(t) - x_e(t|t-1)][x(t) - x_e(t|t-1)]'\} \qquad (7.3\text{-}39)$$

Here, $x_e(t|t-1)$ is a minimum variance estimate of $x(t)$, conditioned on measurements $y(t_0)$, $y(t_0+1)$, $\ldots$, $y(t-1)$. The other form of optimal filter gives, in obvious notation, $x_e(t|t)$ with conditioned error covariance $\Sigma(t|t)$. We have, for the one-step predictor,

$$x_e(t+1|t) = [F(t) + K_e(t)H'(t)]x_e(t|t-1) - K_e(t)y(t) + G(t)u(t) \qquad (7.3\text{-}40)$$

with

$$K_e(t) = F(t) \Sigma (t|t - 1)H(t)[H'(t)\Sigma (t|t - 1)H(t) + R(t)]^{-1} \qquad (7.3\text{-}41)$$

and

$$\Sigma (t + 1|t) = F(t)\{\Sigma (t|t - 1) - \Sigma (t|t - 1)H(t)[H'(t) \Sigma (t|t - 1)H(t) + R(t)]^{-1}$$
$$\times H'(t) \Sigma (t|t - 1)\}F'(t) + Q(t) \qquad (7.3\text{-}42)$$

initialized by $\Sigma (t_0|t_0-1) = \Sigma_0$. It is assumed that the inverse in the Riccati equation exists, this being guaranteed of course if $R(t) > 0$ for all t. Moreover,

$$x_e(t|t) = x_e(t|t - 1) + \Sigma (t|t - 1)H(t)[H'(t) \Sigma (t|t - 1)H(t) + R(t)]^{-1}$$
$$\times [y(t) - H'(t)x_e(t|t - 1)] \qquad (7.3\text{-}43)$$

and

$$\Sigma (t|t) = \Sigma (t|t - 1) - \Sigma (t|t - 1)H(t)[H'(t) \Sigma (t|t - 1)H(t) + R(t)]^{-1}$$
$$\times H'(t) \Sigma (t|t - 1) \qquad (7.3\text{-}44)$$

If F, H, Q, and R are constant and $[F, H]$ is detectable, the limiting Riccati solutions $\lim_{t \to \infty} \Sigma (t|t - 1)$, $\lim_{t \to \infty} \Sigma(t|t)$ exist, and the estimator is asymptotically (as $t \to \infty$) time-invariant. If in addition, $[F, D]$ is stabilizable for any D with $DD' = Q$, then there is asymptotic stability of the optimal estimator.

The innovations process

$$v(t) = y(t) - H'(t)x_e(t|t - 1) \qquad (7.3\text{-}45)$$

is white.

We conclude this section with two comments. First, an optimally designed estimator may be optimal for noise covariances differing from those assumed in its design. This holds for precisely the same reasons that a control law resulting from an optimal design may be optimal for more than one performance index, a point we discussed earlier.

Second, in the interests of achieving an economical realization of an estimator, it may be better to design a suboptimal one that is time-invariant, rather than an optimal one that is varying. For example, suppose the plant whose states are being estimated is time invariant, and that estimation is to start at time zero. Suppose also that at this time, the initial state of the plant is known to be zero. Then there would not normally be a great deal of loss of optimality if, instead of implementing the optimal estimator that would be time-varying, a time-invariant estimator were implemented, designed perhaps on the assumption that the initial time t_0 were $-\infty$ and not zero.

Main points of the section.   Under certain specific assumptions (gaussianness, independence, and whiteness of the noise processes), full order estimators can be designed which yield a minimum variance of the plant state.

Riccati equations are involved, and the whole construction is dual to regulator design. With initial time at $-\infty$ and with a time-invariant plant and constant noise statistics, detectability and stabilizability assumptions yield a time-invariant, asymptotically stable estimator.

Increase of Q or decrease of R results in general in the estimator bandwidth increasing, so that less filtering of the measurement noise takes place.

The results for the stationary model case connect with spectral factorization.

Problem 7.3-1. Consider the plant $\dot{x} = x + v$, $y = x + w$, with $E[v(t)v(\tau)] = E[w(t)w(\tau)] = \delta(t - \tau)$ and v and w independent. Suppose that at time zero, $x(0)$ is known to be zero. Design an optimal estimator.

Problem 7.3-2. Repeat Problem 7.3-1 for the case when the initial time t_0 is $-\infty$. Then with $t_0 = -\infty$, suppose that an additional measurement becomes available; that is, suppose now

$$y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

where $y_1 = x + w_1$, $y_2 = x + w_2$ and $E[w_1(t)w_1(\tau)] = E[w_2(t)w_2(\tau)] = \delta(t - \tau)$ with v, w_1, and w_2 independent. Design an optimal estimator for this case. Compare the error covariances for the single- and multiple-output cases.

Problem 7.3-3. Suppose that $\dot{x} = ax + v$, $y_1 = x + w_1$, $y_2 = x + w_2$, where v, w_1, and w_2 all are independent, with covariances $q\delta(t - \tau)$, $r_{11}\delta(t - \tau)$, and $r_{22}\delta(t - \tau)$, respectively. For the $t_0 = -\infty$ case, derive analytic expressions for the error covariance, assuming that y_1 alone is available and that both y_1 and y_2 are available. Observe that the former case can be obtained from the latter by setting $r_{22} = \infty$.

Problem 7.3-4. Given $\dot{x} = Fx + Gu + v$, $y = H'x + w$, with F, G, and H constant and v and w stationary white noise, suppose that t_0 is finite. Show that the estimator will be time-invariant if $E[x(t_0)x'(t_0)]$ takes on a particular value—in general, nonzero. Assume $v(\cdot)$, $w(\cdot)$ independent for simplicity.

Problem 7.3-5. Consider the standard estimation problem with $u(t) \equiv 0$, $E[x(t_0)] = 0$, and t_0 finite, save that $v(\cdot)$ and $w(\cdot)$ are no longer independent: rather, $E[v(t)w'(\tau)] = S(t)\delta(t - \tau)$ for some matrix $S(t)$. Show that the problem of finding a minimum variance estimate of $x(t_1)$ for arbitrary t_1 is again equivalent to a quadratic regulator problem, with a cross-product term between state and control in the loss function. Attempt to solve the complete estimation problem.

Problem 7.3-6. (The smoothing problem). Let $x_e(t_0|t_1)$ denote the minimum variance estimate of $x(t_0)$, given measurements $y(t)$ up till time $t_1 \geq t_0$, where, as

usual, $\dot{x} = Fx + Gu + v$, $y = H'x + w$, $E[v(t)v'(\tau)] = Q(t)\delta(t - \tau)$, $E[w(t)w'(\tau)] = R(t)\delta(t - \tau)$, $E[x(t_0)x'(t_0)] = P_{e0}$, $E[x(t_0)] = m$, and v, w, and $x(t_0)$ are independent and gaussian, the first two also being zero mean. It is desired to define a procedure for computing $x_e(t_0|t_1)$.

Consider the scheme of Fig. 7.3-8. Observe that $x_{1e}(t|t) = x_e(t|t)$ and $x_{2e}(t|t) = x_e(t_0|t)$, implying that the smoothing problem may be viewed as a filtering problem. Show that

$$\dot{x}_{1e} = (F - P_{11}HR^{-1}H')x_{1e} + P_{11}HR^{-1}y \qquad x_{1e}(t_0|t_0) = m$$

$$\dot{x}_{2e} = -P'_{12}HR^{-1}H'x_{1e} + P'_{12}HR^{-1}y \qquad x_{2e}(t_0|t_0) = m$$

where

$$\dot{P}_{11} = P_{11}F' + FP_{11} - P_{11}HR^{-1}H'P_{11} + Q \qquad P_{11}(t_0) = P_{e0}$$

$$\dot{P}_{12} = (F - P_{11}HR^{-1}H')P_{12} \qquad P_{12}(t_0) = P_{e0}$$

Show also how to find $x_e(t_2|t_1)$ for arbitrary $t_2 < t_1$. [*Hint:* Represent the system of Fig. 7.3-8 as a n-dimensional system with augmented matrix F_a, and so forth, and let

$$P_e = \begin{bmatrix} P_{11} & P_{12} \\ P'_{12} & P_{22} \end{bmatrix}$$

be the solution of the associated filtering Riccati equation. Use P_e to define the optimal $2n$-dimensional filter, and show that the two differential equations shown follow from the optimal filter description. This technique is used in [9].]
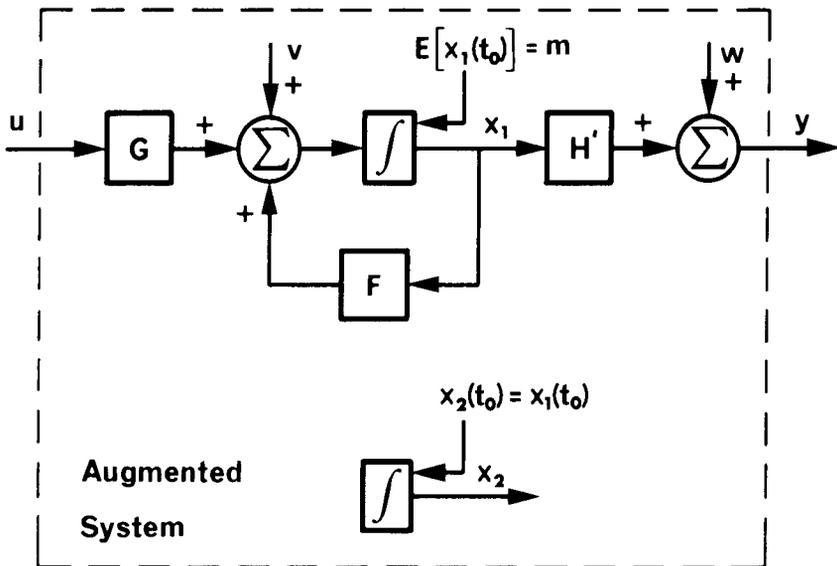
Figure 7.3-8   System augmented with integrators.

Problem 7.3-7.   Suppose that

$$\frac{d}{dt} Z(t) = -F'(t) Z(t) + H(t) M(t) \qquad Z(t_1) = Z_1$$

and suppose that $P_e(t)$ exists as the solution of

$$\dot{P}_e(t) = P_e(t)F'(t) + F(t)P_e(t) - P_e(t)H(t)R^{-1}(t)H'(t)P_e(t) + Q(t)$$

$$P_e(t_0) = P_{e0}$$

with $R(t)$ positive definite symmetric. Show that for all $M(\cdot)$

$$Z'(t_0) P_{e0} Z(t_0) + \int_{t_0}^{t_1} [M'(t) R(t) M(t) + Z'(t) Q(t) Z(t)] \, dt$$

$$= Z'(t_1) P_e(t_1) Z(t_1) + \int_{t_0}^{t_1} [M(t) - R^{-1}(t) H'(t) P_e(t) Z(t)]' R(t)$$

$$\times [M(t) - R^{-1}(t) H'(t) P_e(t) Z(t)] \, dt$$

[*Hint:* Evaluate $-\dfrac{d}{dt}[Z'P_e Z] + M'RM + Z'QZ.$]

Problem 7.3-8.   Consider the linear system (7.3-24), but with K_e replaced by $K_e^{\#}$, and with (7.3-25) holding. Show that $P_e^{\#}(t)$ is minimized for all t with a selection of $K_e^{\#}$ given by $-P_e HR^{-1}$, P_e being the filtering Riccati equation solution. [*Hint:* Write the equation for P_e as

$$\dot{P}_e = (F + K_e^{\#} H') P_e + P_e(F + K_e^{\#} H')' + Q$$

$$- K_e^{\#} H' P_e - P_e H K_e^{\#'} - P_e HR^{-1}H' P_e$$

Set $\Delta = P_e^{\#} - P_e$ and show that

$$\dot{\Delta} = (F + K_e^{\#} H') \Delta + \Delta (F + K_e^{\#} H')' + \hat{Q}$$

for some $\hat{Q} \geq 0$, with $\Delta(t_0) = 0$. By using an explicit solution for Δ, show that $\Delta(t) \geq 0$ for all t, with $\equiv \Delta(t)$ 0 if and only if $K_e^{\#} = -P_e HR^{-1}$].

Problem 7.3-9.   Suppose that $\hat{F}(t) = F'(-t)$. Show that the corresponding transition matrices satisfy $\hat{\Phi}(t, s) = \Phi'(-s, -t)$.

REFERENCES

[1] T. Kailath, *Linear Systems*. Englewood Cliffs, New Jersey: Prentice-Hall, 1980.

[2] D. G. Luenberger, "Observing the State of a Linear System," *IEEE Trans. Military Electron.*, Vol. MIL-8, No. 2 (April 1964), pp. 74–80.

[3] D. G. Luenberger, "Observers for Multivariable Systems," *IEEE Trans. Auto. Control*, Vol. AC-11, No. 2 (April 1966), pp. 190–197.

[4] J. O'Reilly, *Observers for Linear Systems.* New York: Academic Press, 1983.

[5] N. Wiener, *Extrapolation, Interpolation and Smoothing of Stationary Time Series.* Cambridge, Mass: M.I.T. Press, 1949.

[6] A. H. Jazwinski, *Stochastic Processes and Filtering Theory.* New York: Academic Press, 1970.

[7] R. E. Kalman, and R. S. Bucy, "New Results in Linear Filtering and Prediction Theory," *Trans. ASME Ser. D: J. Basic Eng.,* Vol. 83 (March 1961), pp. 95–108.

[8] A. E. Bryson and M. Frazier, "Smoothing for Linear and Nonlinear Dynamic Systems," *Proc. Optimum Systems Synthesis Conf.,* USAF Tech. Rep. ASD-TDR-063-119, February 1963.

[9] L. E. Zachrisson, "On Optimal Smoothing of Continuous Time Kalman Processes," *Inform. Sci.,* Vol. 1 (1969), pp. 143–172.

[10] B. D. O. Anderson and J. B. Moore, *Optimal Filtering.* Englewood Cliffs, New Jersey: Prentice-Hall, 1979.

[11] A. E. Bryson and D. E. Johansen, "Linear Filtering for Time-Varying Systems Using Measurements Containing Colored Noise," *IEEE Trans. Auto. Control,* Vol. AC-10, No. 1 (January 1965), pp. 4–10.

[12] R. E. Kalman and T. S. Englar, "A User's Manual for the Automatic Synthesis Program," *NASA Contractor Rep. NASA CR-475,* June 1966.

8

System Design Using State Estimators

8.1 CONTROLLER DESIGN—BASIC VERSIONS AND VARIATIONS

This chapter is concerned with tying together the notions of state-variable feedback and estimation. In other words, we consider controllers of the sort shown in Fig. 8.1-1, where state estimates x_e are used in lieu of the states x in a full-state feedback design. Attention in this chapter is focused on the time-invariant plant/controller case.

In this section, we concentrate primarily on transfer functions or transfer function matrices from u_{ext} to y in Fig. 8.1-1 and the associated closed-loop eigenvalue locations. We show that the closed-loop transfer function matrices are the same for the state-estimator feedback design as for the full-state design. Also, the closed-loop eigenvalues consist of those of the full-state regulator together with those of the estimator—this is termed an *eigenvalue separation property*. Variations to the scheme of Fig. 8.1-1 are studied based on classical compensator configurations. These will achieve the same closed-loop transfer function matrices but different closed-loop eigenvalues (some of which are uncontrollable), and usually, but not always, the same loop gain or open-loop transfer function matrix.

In the remainder of the chapter, optimality of the arrangement of Fig. 8.1-1 is studied for the stochastic case, leading to the *Separation Theorem* (also known as the Certainty Equivalence Principle). It is also pointed out that with the introduction of observers, input or output robustness properties, such as attractive gain and phase
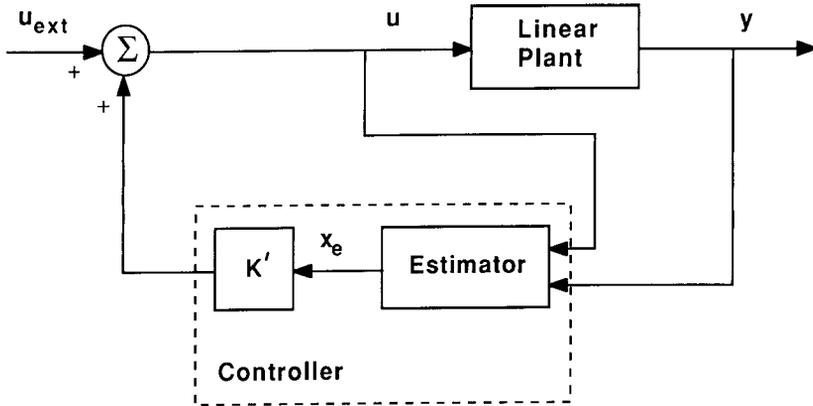
Figure 8.1-1 The basic controller.

margins, can evaporate. However, roll-off rates can improve, enhancing high-frequency robustness. The so-called loop recovery technique is outlined as a method to recover robustness properties associated with the state feedback design. On a different tack, the class of all stabilizing controllers is introduced as a vehicle to achieve variations of the basic state estimate feedback structure. The variations involve feedback of residuals $(y - y_e)$, and allow robustness of a state estimate feedback design to be varied or "optimized" while preserving performance properties of a nominal design. We do not give explicit optimization procedures in this text, except where the optimum is given analytically in a straightforward manner.

Basic controller design.   As our basic plant, we take the time-invariant system

$$\dot{x} = Fx + Gu \tag{8.1-1}$$

$$y = H'x \tag{8.1-2}$$

For the observer, we shall take for the moment the full-order structure of Chapter 7, Secs. 7.2 and 7.3—viz.,

$$\dot{x}_e = (F + K_e H')x_e + Gu - K_e y \tag{8.1-3}$$

We assume that K_e is chosen so that all eigenvalues of $F + K_e H'$ have negative real parts. Whether it is optimal for some noise statistics is irrelevant for our considerations here; either the scheme of Sec. 7.2 or that of Sec. 7.3 can be assumed to lead to the choice of K_e. Subsequently in this section we shall consider the use of reduced order observers in controller design.

We further assume that we should like to implement the control law $u = K'x + u_{\text{ext}}$. Here u_{ext} denotes an external input, possibly derived by optimal or suboptimal tracking results as in Chapter 4. It may include a term $-K'\tilde{x}$ where $\tilde{x}$ is a desired state trajectory, or a feedforward term $K_1'r$ where r is a reference trajectory.

Also presumably, K has been selected as an optimal control law. However, for obvious reasons, we implement the following law instead:

$$u = K'x_e + u_{ext} \tag{8.1-4}$$

Another arrangement useful for y to track a reference may have the external input combined with y prior to entering the state estimator. This case is discussed later in the section.

Equations (8.1-1) through (8.1-4) sum up the entire plant-controller arrangement. We shall now study some properties of the arrangement using these equations. From (8.1-1)–(8.1-4),

$$\dot{x} = (F + GK')x - GK'(x - x_e) + Gu_{ext} \tag{8.1-5a}$$

$$\dot{x}_e = (F + GK' + K_eH')x_e - K_ey + Gu_{ext} \tag{8.1-5b}$$

Subtracting and exploiting (8.1-2), we have

$$\frac{d}{dt}(x - x_e) = (F + K_eH')(x - x_e) \tag{8.1-6}$$

which holds independently of u_{ext}. Now we regard the $2n$ vector, whose first n entries are x and whose second n entries are $x - x_e$, as a new state vector for the overall plant-controller scheme. (It would, of course, be equally valid to take as a state vector a $2n$ vector with the first n entries consisting of x and the second n entries consisting of x_e.)

The plant-controller arrangement, then, has the following description—the first equation following from (8.1-5) and (8.1-6), the second from (8.1-2):

$$\frac{d}{dt}\begin{bmatrix} x \\ x - x_e \end{bmatrix} = \begin{bmatrix} F + GK' & -GK' \\ 0 & F + K_eH' \end{bmatrix}\begin{bmatrix} x \\ x - x_e \end{bmatrix} + \begin{bmatrix} G \\ 0 \end{bmatrix}u_{ext} \tag{8.1-7}$$

$$y = \begin{bmatrix} H' & 0 \end{bmatrix}\begin{bmatrix} x \\ x - x_e \end{bmatrix} \tag{8.1-8}$$

With input u_{ext} and output y, the plant-controller arrangement has the following transfer function matrix, derivable by manipulating (8.1-7) and (8.1-8):

$$W(s) = H'[sI - (F + GK')]^{-1}G \tag{8.1-9}$$

This is exactly the transfer function matrix that would have resulted if true state-variable feedback were employed. The poles of the open-loop plant, corresponding to the zeros of $\det(sI - F)$, are shifted to the zeros of $\det[sI - (F + GK')]$. The zeros of a scalar $W(s)$ are unaltered.

Thus, from the steady-state (or zero initial state) point of view, use of the estimator as opposed to use of true state-variable feedback makes no difference. This is, of course, what should be expected. For the case in which the steady state has been reached, $x - x_e$ has approached zero and $x = x_e$, or, in the case of zero initial state, $x = 0$ and $x - x_e = 0$, so that again $x = x_e$. Clearly, with $x = x_e$, the control used is precisely that obtained with true state-variable feedback.

From the transient point of view, the plant-controller scheme of Eqs. (8.1-7) and (8.1-8) will behave differently from a scheme based on true state-variable feedback. Equation (8.1-7) defines a $2n$-dimensional system, whereas state-variable feedback yields an n-dimensional system. However, the $2n$-dimensional system is still asymptotically stable; inspection of (8.1-7) shows that the eigenvalues of $F + GK'$ and of $F + K_eH'$ determine the characteristic modes. (The eigenvalues of $F + K_eH'$ are, of course, associated with the additional new modes, which are evidently uncontrollable from u_{ext}.)

The open-loop transfer functions depend on where the loop is opened. Of course, if the loop is opened just after the summing function in Fig. 8.1-1, then the open-loop transfer function is simply $K'(sI - F)^{-1}G$, being identical to that with full-state feedback. However, to analyze input robustness it is usual to open the loop at the plant input. Then the loop gain is the product of the plant transfer function matrix $H'(sI - F)^{-1}G$ and that of the controller $K'(sI - F - GK' - K_eH')^{-1}(-K_e)$. The latter is the transfer function from the controller input y to its output $K'x_e$, being most easily obtained from (8.1-5b) with $u_{ext} \equiv 0$. The open-loop gain transfer function matrix is thus

$$W_{OL}(s) = [K'(sI - F - GK' - K_eH')^{-1}(-K_e)][H'(sI - F)^{-1}G] \qquad (8.1\text{-}10)$$

The situation is depicted in the regulator of Fig. 8.1-2 ($u_{ext} \equiv 0$), where the loop is opened at the point X.
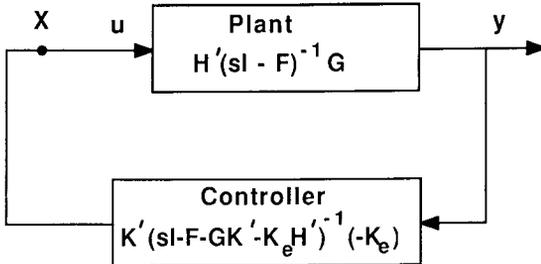
Figure 8.1-2 Regulator structure.

This is not the same as for full-state feedback when the loop gain is $K'(sI - F)^{-1}G$. Notice that in view of this fact, we cannot expect the return difference inequality to hold, with its associated guaranteed robustness properties, for the new return difference. Notice also that the roll-off rate of the loop gain now is necessarily faster than 6 dB/octave, the roll-off rate for the full-state feedback case. Thus there is the potential to gain in high-frequency robustness and to lose passband robustness by introducing state estimation.

Controller with a reduced-order observer.

We now discuss briefly the equations of the plant-controller arrangement when a reduced-order observer is used. We first observe that there is no loss of generality involved in describing the plant and controller by state-space equations with a special coordinate basis when we are seeking to compute the overall transfer function matrix, or

the qualitative stability properties of the arrangement. Overbars will indicate this special coordinate basis. Thus $\dot{\bar{x}} = \bar{F}\bar{x} + \bar{G}u$, $u = \bar{K}'\bar{x} + u_{\text{ext}}$, where

$$\bar{x} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix} = \begin{bmatrix} \bar{x}_1 \\ y \end{bmatrix} \qquad \bar{F} = \begin{bmatrix} \bar{F}_{11} & \bar{F}_{12} \\ \bar{F}_{21} & \bar{F}_{22} \end{bmatrix} \qquad \bar{K} = \begin{bmatrix} \bar{K}_1 \\ \bar{K}_2 \end{bmatrix}$$

The reduced-order estimator of Sec. 7.2 gives state estimates

$$\bar{x}_e = \begin{bmatrix} \bar{x}_{1e} \\ \bar{x}_{2e} \end{bmatrix} = \begin{bmatrix} \bar{x}_{1e} \\ y \end{bmatrix}$$

A state vector for the overall plant-controller arrangement is, therefore, provided by the vector

$$\begin{bmatrix} x \\ \bar{x}_{1e} \end{bmatrix}$$

However, a more appropriate, but equally acceptable, choice for state vector proves to be

$$\begin{bmatrix} \bar{x} \\ \bar{x}_1 - \bar{x}_{1e} \end{bmatrix}$$

Recall from Sec. 7.2 that the errors $(\bar{x}_1 - \bar{x}_{1e})$ satisfy the differential equation, see (7.2-22),

$$\frac{d}{dt}(\bar{x}_1 - \bar{x}_{1e}) = F_e(\bar{x}_1 - \bar{x}_{1e}), \qquad F_e = \bar{F}_{11} + K_e\bar{F}_{21}$$

where K_e is the state estimator gain. There results the overall closed-loop system equations

$$\frac{d}{dt}\begin{bmatrix} \bar{x} \\ \bar{x}_1 - \bar{x}_{1e} \end{bmatrix} = \begin{bmatrix} \bar{F} + \overline{GK}' & -G\bar{K}_1' \\ 0 & F_e \end{bmatrix}\begin{bmatrix} \bar{x} \\ \bar{x}_1 - \bar{x}_{1e} \end{bmatrix} + \begin{bmatrix} \bar{G} \\ 0 \end{bmatrix}u_{\text{ext}} \qquad (8.1\text{-}11)$$

$$y = \bar{H}'\bar{x} \qquad (8.1\text{-}12)$$

These two equations then imply again that the transfer function matrix relating u_{ext} to y is $H'[sI - (F + GK')]^{-1}G$, and that the zero-input response of the plant-controller arrangement is determined by the eigenvalues of $F + GK'$ and of F_e. The only difference between the reduced-order estimator and that considered earlier is that the $n \times n$ matrix $F + K_eH'$ is replaced by the reduced-order matrix F_e.

There are reduced-order controller designs, not studied here, which seek to estimate the control signal $K'x$ directly, rather than by means of a state estimate x_e. In such methods, a state estimate x_e cannot usually be extracted from the controller. See, for example, [1, 2] and their references.

Classical controller structures.   The aim of what follows is to indicate some parallels with classical control ideas. This will be done by exhibiting some variants on the controller arrangement of Fig. 8.1-1, the structures of which will be familiar from classical control.

We will derive controller structures as indicated in Fig. 8.1-3. We now list some general properties of these controller structures.

1. The controller structures will be derived by manipulations on the controller structure of Fig. 8.1-1. These manipulations do not affect the input-output steady-state performance (or zero-state performance) of the overall scheme. In other words, the transfer function or transfer function matrix of the overall scheme is still $H'[sI - (F + GK')]^{-1}G$, where $H'(sI - F)^{-1}G$ is the transfer function (matrix) of the plant and $u = K'x$ is the desired feedback law.

2. There is no restriction to single-input or single-output plants.

3. The dimensions of the compensators are the same as the dimension of the controllers from which they are derived. This means that if, for example, an n-dimensional single-output system employs an $(n - 1)$-dimensional controller, then the series compensator *and* the feedback compensator will each have dimension $(n - 1)$.

4. In view of 3, there are additional modes again in the controller of Figs. 8.1-3(b) and 8.1-3(c) beyond those introduced in the controllers of Figs. 8.1-3(a) and 8.1-3(d). In the case of the controllers of Figs. 8.1-3(c), these

(a)

(b)

(c)

(d)
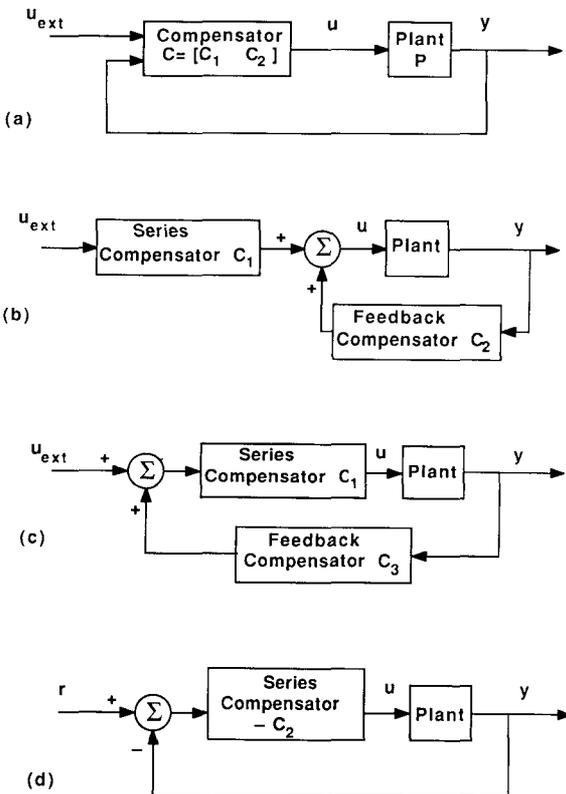
Figure 8.1-3   Controller structures familiar from classical ideas.

additional modes are always asymptotically stable. However, this is not the case with the scheme of Fig. 8.1-3(b), which may thus prevent its use.

We shall now proceed with the derivation of the structures. We start with the controller of Fig. 8.1-1, reorganized as having inputs u_{ext}, y and outputs u, as in Fig. 8.1-4. This has the desired structure of the two-degrees-of-freedom compensator scheme of Fig. 8.1-3(a). For the full-order state estimator case, the state equations are

$$\dot{x}_e = (F + K_e H' + GK')x_e + Gu_{ext} - K_e y$$
$$u = K'x_e + u_{ext}$$

(8.1-13)

Its transfer function matrix, linking $[u'_{ext} \; y']'$ to u, is

$$C(s) = K'(sI - F - K_e H' - GK')^{-1}[G \quad -K_e] + [I \quad 0] \triangleq [C_1(s) \quad C_2(s)] \quad \text{(8.1-14)}$$

There is no guarantee in any of our theory that the compensator of (8.1-13) and (8.1-14) is (open-loop) stable, that is, that $F + K_e H' + GK'$ has all eigenvalues with negative real parts. Indeed, it is known that for a single-input, single-output plant, any stabilizing compensator is (open-loop) unstable unless there is an even number of right half-plane real poles between successive right half-plane real axis zeros [3]. Further, a stabilizing compensator may need to have dimension far greater than that of the plant if it is also to be open-loop stable.

Engineers are understandably hesitant to implement open-loop unstable controllers, and a search for alternative designs may be indicated, perhaps using alternative sensors and actuators to avoid this open-loop instability. Of course, from a theoretical point of view, working with a linear unstable compensator is no different from working with a linear unstable plant. In practice, saturation and a possible desire to maintain some acceptable level of performance in the face of sensor or actuator failure will alter this view.

If the compensator denoted C_1 in Fig. 8.1-3(a) is open-loop stable, then the arrangement of Fig. 8.1-3(b) can be implemented with the series compensator C_1 and the feedback compensator C_2. Notice that the orders of C_1 and C_2 are each the
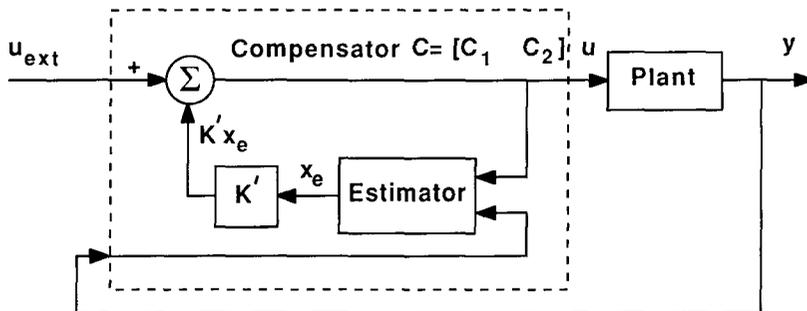
Figure 8.1-4   The basic controller scheme rearranged.

same, in general, as the order of $C = [C_1 \quad C_2]$, so there is a complexity cost increase here. The extra modes introduced, in addition to those of the basic control scheme of Fig. 8.1-3(a) (with modes that are eigenvalues of $F + GK'$, $F + K_eH'$), are the poles of $C_1(s)$, being the eigenvalues of $F + GK' + K_eH'$. Details on this derivation are left to Problem 8.1-4.

The two-compensator scheme of Fig. 8.1-3(b) is not recommended here, but it is of interest to make comparisons with classical designs. The scheme could be of interest with an incorporated tracking series compensator, as in Chapter 4, and compensator order reductions, as in Chapter 10.

In the series-feedback arrangement of Fig. 8.1-3(c), the series compensator is C_1, as in the previous arrangements. The feedback compensator is the subsystem of the Fig. 8.1-4 scheme with input y and output feeding into the summing node. Its transfer function is

$$C_3(s) = K'(sI - F - K_eH')(-K_e) \qquad (8.1\text{-}15)$$

Notice that there must hold $C_1(s)C_3(s) = C_2(s)$, given the equivalence of the arrangements of Figs. 8.1-3(a) and 8.1-3(c). This is readily checked via algebraic manipulations. Again, there is an increase in total compensator complexity over the basic control scheme in such an arrangement. The additional modes beyond those for the basic control scheme are those of the estimator, which are known to be stable, and are under the designer's control. Details are explored in Problem 8.1-4.

The classical compensator structures studied so far are all direct variations of Fig. 8.1-1 with u_{ext} as indicated. A further variation is easily organized to achieve the perhaps most familiar classical arrangement of all—the unity feedback scheme of Fig. 8.1-3(d). The idea is that y should track an external reference r and indeed, y will track r in the bandwidth of the closed-loop system. Because of this property, the unity feedback arrangement is perhaps the main alternative to that of Fig. 8.1-3(a) for linear quadratic-based design.

Of course, the same rearrangements as in Fig. 8.1-3 can be achieved by using reduced-order estimators, whether or not there is an explicit state estimation. Again there is the possible additional complexity and extra modes that must be weighed against any conceivable gains in the rearrangements. The following example is instructive.

Illustrative example.   We consider a second-order position controller for a plant with transfer function s^{-2}. Suppose that

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

$$y = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Also let the control be $u = [-1 \quad -1]x$. Let us use a reduced-order estimator with

a single eigenvalue at $s = -\alpha$. Then $x_{1e} = y$, and $x_{2e} = w + \alpha y$, where $\dot{w} = -\alpha w - \alpha^2 y + u$. Also

$$u = -y - (w + \alpha y) + u_{\text{ext}}$$

Inserting this into the expression for w yields

$$\dot{w} = -\alpha w - \alpha^2 y - y - (w + \alpha y) + u_{\text{ext}}$$

$$= -(\alpha + 1)w - (1 + \alpha + \alpha^2)y + u_{\text{ext}}$$

The transfer function from u_{ext} to u is $(s + \alpha)(s + \alpha + 1)^{-1}$, and from y to u it is $-[(1 + \alpha)s + \alpha](s + \alpha + 1)^{-1}$.

The associated two-degrees-of-freedom controller, with input u_{ext} where $u = -[1 \quad 1]x + u_{\text{ext}}$, is shown in Fig. 8.1-5(a). The closed-loop eigenvalues are the zeros of $(s^2 + s + 1)(s + \alpha)$. Alternative arrangements are depicted in Figs. 8.1-5(b) and 8.1-5(c). Notice that the product of the series and feedback compensator

(a)

(b)

(c)

(d)
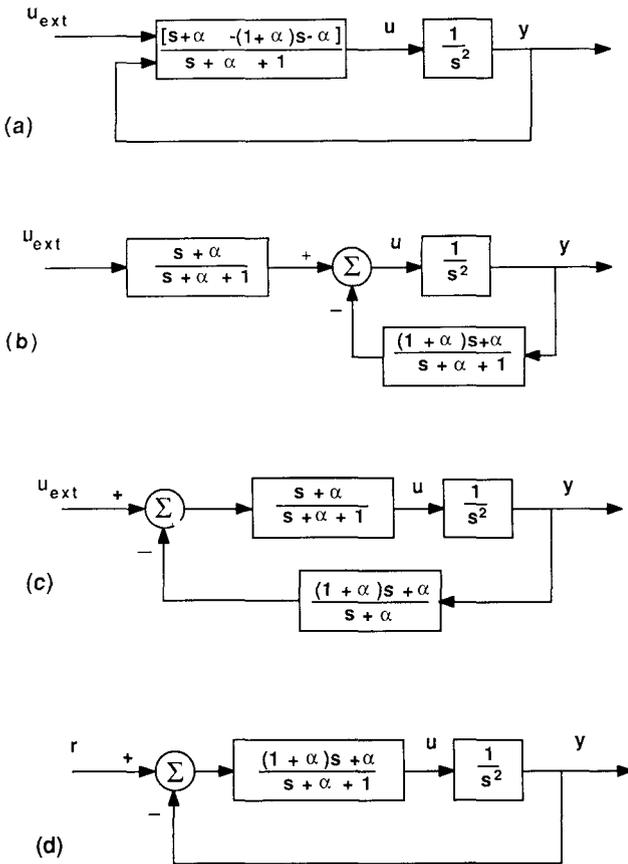
Figure 8.1-5 Controllers discussed in example.

transfer functions in Fig. 8.1-5(c) agrees with the feedback compensator in Fig. 8.1-5(b). The eigenvalues associated with the arrangements of Fig. 8.1-5(b) and 8.1-5(c) are the zeros of $(s^2 + s + 1)(s + \alpha)(s + \alpha + 1)$ and $(s^2 + s + 1)$ $(s + \alpha)^2$, respectively.

A classical design, set up with a different external input point, might actually have led to the scheme of Fig. 8.1-5(d). [Compare with Fig. 8.1-5(c).] With $\alpha = 10$, which would be suggested by the requirement that the estimator have a much faster time constant than that of the closed-loop transfer function, the compensator is

$$\frac{11s + 10}{s + 11}$$

This is evidently introducing a phase-lead. Since it emphasizes high frequencies more than low frequencies, too large an α will give problems with noise. The choice $\alpha = 10$ is in accord with classical control. The associated root locus diagram is shown in Fig. 8.1-6, parametrized with the gain K. The value $K = 1$ corresponds to the nominal value, and leads to closed-loop poles at $-10, -0.5 \pm j\sqrt{3/2}$, whereas pure state-feedback would have given poles at $-0.5 \pm j\sqrt{3/2}$. Observe that the infinite gain margin property is retained here; in Section 8.3, we shall return to the question of gain margins when state estimators are used.

Notice that the open-loop transfer function

$$W_{OL}(s) = \frac{(1 + \alpha)s + \alpha}{s + \alpha + 1} \frac{1}{s^2}$$

has a roll-off rate of 12 dB/octave.
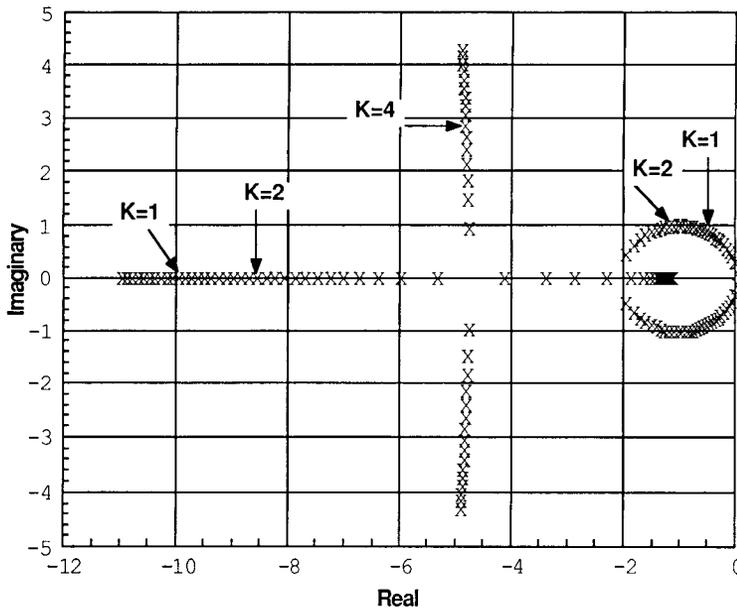
Figure 8.1-6 Root locus plot for system of Fig. 8.1-5(d).

Main points of the section.   State estimate feedback designs have the same transfer functions as full-state feedback designs. Their closed-loop eigenvalues consist of those for a full-state feedback design together with those of the estimator. When reorganized in terms of classical compensator configurations, the compensators are of high order relative to classical compensators, introducing uncontrollable and unobservable modes into the closed-loop system equations.

Problem 8.1-1.   Consider the plant

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u \qquad y = [1 \quad 0]x$$

Calculate an optimal state feedback control law that will minimize the performance index $\int_{t_0}^{\infty} (u^2 + 2x_1^2 + 3x_2^2)\, dt$. Using the concepts of this section, design dynamic output feedback controllers of both dimension 2 and dimension 1 with poles that have real parts with significantly larger magnitude than the closed-loop system poles associated with the state feedback design. Present the controllers as for a unity negative feedback scheme.

Problem 8.1-2.   Consider the first-order plant

$$\dot{x} = x + u, \qquad y = x$$

The control law $u = -3x$ is an optimal law for this plant. Design state estimators of dimension 1 with poles at -1, -5, -10. Then sketch the response of $x(t)$, given $x(0) = 1$ for the following eight cases. In case 1, no estimator is used, and in cases 2 through 8 an estimator is used.

1. The feedback law is $u = -3x$.
2. Estimator pole is at -1 and $x_e(t_0) = 0$.
3. Estimator pole is at -1 and $x_e(t_0) = \frac{1}{2}$.
4. Estimator pole is at -1 and $x_e(t_0) = -1$.
5. Estimator pole is at -5 and $x_e(t_0) = 0$.
6. Estimator pole is at -10 and $x_e(t_0) = 0$.
7. Estimator pole is at -10 and $x_e(t_0) = \frac{1}{2}$.
8. Estimator pole is at -10 and $x_e(t_0) = -1$.

Comment.

Problem 8.1-3.   Consider the $2n$-dimensional system defined by Eqs. (8.1-1) through (8.1-3) and let $u = K'x$ be an optimal control law resulting from minimization of a performance index $\int_{t_0}^{\infty} (u'u + x'Qx)\, dt$. Show that if Q is positive definite, there exists a positive definite $\hat{Q}$ such that $u = K'x_e$ is the optimal control for a performance index $\int_{t_0}^{\infty} [u'u + (x' \quad x' - x_e')\hat{Q}(x' \quad x' - x_e')']\, dt$. Indicate difficulties in extending this result to the case where Q is singular. (The conclusions of this problem are also valid when reduced-order estimators are used.) [*Hint:* Take as the state vector z, where $z' = (x' \quad x' - x_e')$.]

Problem 8.1-4.   Consider the basic control scheme Fig. 8.1-1 for the full-state estimator case. Show that additional modes in the reorganizations of Fig. 8.1-3(b), (c) are, respectively, eigenvalues of $F + GK' + K_e H'$ and of $F + K_e H'$. [*Hint:* In the second case, use a state vector consisting of x, x_1 (the state of C_1), and $(x - x_3 - x_1)$, where x_3 is the state of C_3.]

8.2 THE SEPARATION THEOREM AND PERFORMANCE CALCULATION

This section is confined to a brief treatment of a theoretical result known as the *Separation Theorem* or *Certainty Equivalence Principle*. Performance calculations are also studied for state estimate feedback LQG design, and a design example is included.

We assume that we are given a linear system with additive input noise:

$$\dot{x} = F(t)x + G(t)u + v \tag{8.2-1}$$

The input noise v is white, gaussian, of zero mean, and has covariance $\hat{Q}(t)\delta(t - \tau)$, where $\hat{Q}$ is nonnegative definite symmetric for all t. The output y of the system is given by

$$y = H'(t)x + w \tag{8.2-2}$$

where w is white gaussian noise of zero mean, and has covariance $\hat{R}(t)\delta(t - \tau)$, where $\hat{R}(t)$ is positive definite for all t. The processes v and w are independent. The initial state $x(t_0)$ at time t_0 is a gaussian random variable of mean m and covariance P_0, and is independent of the processes v and w. The matrices $F, G, H, \hat{Q}$, and $\hat{R}$ are all assumed to have continuous elements.

It is not possible to pose an optimal control problem requiring minimization of

$$V = \int_{t_0}^{T} [x'Q(t)x + u'R(t)u]\, dt \tag{8.2-3}$$

where $Q(t)$ is nonnegative definite symmetric and $R(t)$ is positive definite symmetric, even if one restricts the optimal $u(t)$ to being derived from the measurement $y(\cdot)$, because the performance index V must actually be a random variable, taking values depending on $v(\cdot)$, $w(\cdot)$, and $x(t_0)$, which, of course, are random.

To eliminate this difficulty, one can replace (8.2-3) by

$$V = E\left\{\int_{t_0}^{T} [x'Q(t)x + u'R(t)u]\, dt\right\} \tag{8.2-4}$$

where the expectation is over $x(t_0)$ and the processes $v(\cdot)$ and $w(\cdot)$ on the interval $[t_0, T]$. It is understood that at time t, the measurements $y(\tau)$, $t_0 \leq \tau < t$ are available along with the initial statistics of $x(t_0)$ (in principle m and P_0, but in practice only m is needed). Equivalently (but not altogether obviously), $u(t)$ is allowed to depend on $y(\tau)$ and $u(\tau)$ for $t_0 \leq \tau < t$, as well as the initial statistics of $x(t_0)$. *Note:* The optimal $u(t)$ is *not* required to be an instantaneous function of $y(t)$.

The solution of this problem, which has come to be known as the Separation Theorem for reasons that will be obvious momentarily, is deceptively simple. It falls into three parts:

1. Compute a causal minimum variance estimate $x_e(t)$ of $x(t)$ at time t, using $u(\tau)$, $t_0 \leq \tau < t$ and $y(\tau)$, $t_0 \leq \tau < t$. As we know, this problem has a solution wherein $x_e(t)$ is the output of a linear system excited by $u(\cdot)$ and $y(\cdot)$. This linear system is independent of the matrices $Q(t)$ and $R(t)$—that is, the same linear system generates $x_e(t)$, irrespective of what $Q(t)$ and $R(t)$ are.

2. Compute the optimal control law $u(t) = K'(t)x(t)$, which would be applied if there were no noise, if $x(t)$ were available, and if (8.2-3) were the performance index.

3. Replace x by its estimate x_e. That is, use the control law $u(t) = K'(t)x_e(t)$, where $x_e(t)$ is obtained as in (1). This law is optimal for the noisy problem. Notice that the calculation of $K(t)$ is independent of $H(t)$, and the statistics of the noise.

Evidently, the calculation of $x_e(t)$ and of the control law gain matrix $K(t)$ are separate problems that can be tackled independently. Hence the name "Separation Theorem." Figure 8.2-1 shows the optimal controller.
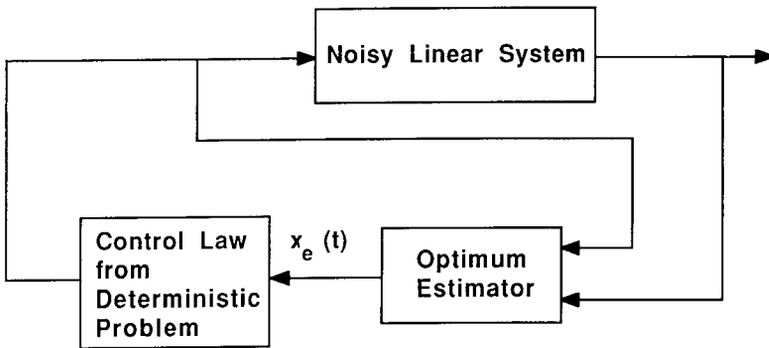
Figure 8.2-1   Illustration of the Separation Theorem.

A proof of the Separation Theorem is given at the end of the section.

The Separation Theorem does not extend to arbitrary nonlinear stochastic optimal control problems. In a sense, the result is surprising. One might have expected that in the face of the additional uncertainty caused by the presence of noise, a more cautious form of control would be used, perhaps with lower gains, as in the case of plant uncertainty.

A most important special case of the Separation Theorem deserves to be highlighted. Suppose that (8.2-1) holds, while $x(t)$ is available for measurement, with no noise. Thus we have a linear estimation problem with white input disturbances, with output equal to state, and with no measurement noise. In this case

the Separation Theorem states that $u(t) = K'(t)x(t)$ is again optimal, of course in the sense of minimizing the index (8.2-4), not (8.2-3).

In case $T \to \infty$, (8.2-4) needs replacement. It is clear that there is no way that u and x will decay to zero as $t \to \infty$, if in (8.2-1), the random process $v(\cdot)$ remains active; that is, $\hat{Q}(t) \not\to 0$ as $t \to \infty$. As a result, the index (8.2-4) will diverge to ∞, no matter what control is employed. Restricting consideration for simplicity to the time-invariant case, the natural replacement for (8.2-4) becomes

$$V = \lim_{T \to \infty} \frac{1}{T} E \left\{ \int_0^T [x'Qx + u'Ru]\, dt \right\} \tag{8.2-5}$$

with the same restrictions as before on $u(t)$ [viz. $u(t)$ depends either on past $y(\cdot)$, $u(\cdot)$, and x_0 statistics, or on current $x(\cdot)$] together with one new restriction: $u(t)$ must be generated by a time-invariant controller, or state feedback law. Without this latter restriction, it would be possible to change the controller on a finite interval arbitrarily without affecting the value of the performance index.

The Separation Theorem continues to apply to the minimization of (8.2-5), with the steady state control gain K and filter gain K_e defining the controller. [Of course, if $x(\cdot)$ is available for measurement, K_e does not enter the picture.]

In any properly posed LQG problem, the optimal closed-loop will be stable, and consequently, $x(\cdot)$ and $u(\cdot)$ will be stationary random processes. There results the following rewriting of (8.2-5):

$$V = E[x'Qx + u'Ru] \tag{8.2-6}$$

The discrete-time case.   Not surprisingly, the Separation Theorem applies also in discrete time. If the admissible control strategies restrict $u(t)$ so that it is a function of only $y(0), y(1), \ldots, y(t-1)$, then $x_e(t|t-1)$ is used in lieu of $x(t)$—leading to strictly proper controllers. If there is only a properness restriction rather than a strict properness restriction, then $x_e(t|t)$ is used in lieu of $x(t)$. This issue is explored in Problem 8.2-2.

Performance calculations.   It is important to know how to compute the optimal performance index when noise is present. Let us begin with the case when $x(t)$ is available for measurement. Then we implement $u(t) = K'(t)x(t)$, so (8.2-1) becomes

$$\dot{x} = [F(t) + G(t)K'(t)]x + v(t) \tag{8.2-7}$$

Let $M(t)$ be the solution of

$$\dot{M} = (F + GK')M + M(F + GK')' + \hat{Q}$$
$$M(t_0) = P_0 + mm' \tag{8.2-8}$$

Then $M(t) = E[x(t)x'(t)]$; see Appendix B. Further

$$V = \int_{t_0}^{T} \{E[x'(t)Qx(t)] + E[x'(t)K(t)R(t)K'(t)x(t)]\}\, dt$$
$$= \int_{t_0}^{T} \operatorname{tr}[(Q + KRK')M]\, dt \tag{8.2-9}$$

[It is not hard to verify that if $\hat{Q} = 0$, $P_0 = 0$, and $m = x(t_0)$, then this formula is consistent with the earlier deterministic result.] In case $T \to \infty$ and we use index (8.2-6), we have as a replacement for (8.2-8),

$$(F + GK')M + M(F + GK')' + \hat{Q} = 0 \qquad (8.2\text{-}10)$$

and, replacing (8.2-9),

$$V = \text{tr}\,[(Q + KRK')M] \qquad (8.2\text{-}11)$$

The initial statistics of $x(t_0)$ play no role in this expression; they are forgotten, or become irrelevant, in an infinite time interval.

In case $x(t)$ must be estimated, the style of calculation above still applies, but is more complicated. We shall set out the details for the infinite time case only. Then we have

$$\begin{bmatrix} \dot{x} \\ \dot{x}_e \end{bmatrix} = \begin{bmatrix} F & GK' \\ -K_e H' & F + K_e H' + GK' \end{bmatrix} \begin{bmatrix} x \\ x_e \end{bmatrix} + \begin{bmatrix} v & 0 \\ 0 & -K_e w \end{bmatrix}$$

or

$$\begin{bmatrix} \dot{x} \\ \dot{x}_e - \dot{x} \end{bmatrix} = \begin{bmatrix} F + GK' & GK' \\ 0 & F + K_e H' \end{bmatrix} \begin{bmatrix} x \\ x_e - x \end{bmatrix} + \begin{bmatrix} I & 0 \\ -I & -K_e \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} \qquad (8.2\text{-}12)$$

The covariance matrix

$$E\left\{\begin{bmatrix} x(t) \\ x_e(t) - x(t) \end{bmatrix} [x'(t) \quad x_e'(t) - x'(t)]\right\} = \begin{bmatrix} S_{11} & S_{12} \\ S_{12}' & S_{22} \end{bmatrix} \qquad (8.2\text{-}13)$$

can now be evaluated from

$$\begin{bmatrix} F + GK' & GK' \\ 0 & F + K_e H' \end{bmatrix} \begin{bmatrix} S_{11} & S_{12} \\ S_{12}' & S_{22} \end{bmatrix} + \begin{bmatrix} S_{11} & S_{12} \\ S_{12}' & S_{22} \end{bmatrix} \begin{bmatrix} (F + GK')' & 0 \\ KG' & (F + K_e H')' \end{bmatrix}$$

$$+ \begin{bmatrix} \hat{Q} & -\hat{Q} \\ -\hat{Q} & \hat{Q} + K_e \hat{R} K_e' \end{bmatrix} = 0 \qquad (8.2\text{-}14)$$

(It is easily verified that $S_{22} = P_e$.) Also, from (8.2-13), it follows that

$$E[x_e(t)x_e'(t)] = (S_{11} + S_{12} + S_{12}' + S_{22}) \qquad (8.2\text{-}15)$$

Finally,

$$V = E[x'Qx + u'Ru]$$
$$= \text{tr}\,\{QE(xx')\} + \text{tr}\,\{RK'E(x_e x_e')K\} \qquad (8.2\text{-}16)$$
$$= \text{tr}\,[QS_{11}] + \text{tr}\,[RK'(S_{11} + S_{12} + S_{12}' + S_{22})K]$$

Notice that this calculation does not appeal to the optimality of K, K_e (except in our side remark that $S_{22} = P_e$). It could therefore be used with suboptimal controllers.

Lastly, we comment on an alternative way of expressing (8.2-16). Call $T(j\omega)$

and $T_e(j\omega)$ the transfer function matrices from the vector $[v'\ w']'$ to x and x_e respectively. Then, the second equality in (8.2-16) can be reorganized by using

$$E(xx') = \frac{1}{2\pi} \int_{-\infty}^{\infty} T(j\omega) \begin{bmatrix} \hat{Q} & 0 \\ 0 & \hat{R} \end{bmatrix} T'(-j\omega)\, d\omega \qquad (8.2\text{-}17a)$$

$$E(x_e x_e') = \frac{1}{2\pi} \int_{-\infty}^{\infty} T_e(j\omega) \begin{bmatrix} \hat{Q} & 0 \\ 0 & \hat{R} \end{bmatrix} T_e'(-j\omega)\, d\omega \qquad (8.2\text{-}17b)$$

(see Appendix B). We shall now illustrate the use of this type of frequency domain formulation as an aid to comparing the efficacy of various controllers.

A resonance suppression design example.   We consider here a low-order model of an aircraft in level flight subjected to wind gust turbulence. The objective is to suppress horizontal vibrations at the flight deck and in the tail section by means of LQG (yaw-damper) control. Measurements are the forward and aft acceleration readings y_f, y_a, and the scalar control is the rudder position u. In this design example, it makes sense to work with modal coordinates.

Consider a stochastic plant model $\dot{x} = Fx + Gu + \Gamma v$, $y = H'x + w$ in the usual form where $x = [x_1\ \ x_2\ \ x_3\ \ x_4\ \ x_5\ \ x_6]'$, $y = [y_f\ \ y_a]'$ and

$$F = \begin{bmatrix} -2.82 \times 10^{-8} & & & 0 \\ & -6.76 \times 10^{-3} & & \\ & & \begin{bmatrix} -0.122 & 1.57 \\ -1.57 & -0.122 \end{bmatrix} & \\ 0 & & & \begin{bmatrix} -0.71 & 21.3 \\ -21.3 & -0.71 \end{bmatrix} \end{bmatrix}$$

$$[G\ \ \Gamma] = \begin{bmatrix} -4.06 \times 10^{-4} & -1.65 \times 10^{-7} \\ -2.20 \times 10^{-1} & 7.61 \times 10^{-5} \\ 8.84 \times 10^{-2} & 3.05 \times 10^{-2} \\ -3.08 \times 10^{-1} & -1.05 \times 10^{-2} \\ 6.39 \times 10^{-2} & 4.46 \times 10^{-3} \\ -1.08 & -1.02 \times 10^{-2} \end{bmatrix}$$

$$H = \begin{bmatrix} -3.5\ \ \times 10^{-4} & -2.05 \times 10^{-4} \\ -1.57 \times 10^{-1} & -1.54 \times 10^{-1} \\ -1.56 \times 10^{-1} & 2.81 \times 10^{-1} \\ 4.61 \times 10^{-3} & -9.85 \times 10^{-2} \\ -3.37 \times 10^{-1} & -1.02 \\ -2.40 \times 10^{-2} & -7.68 \times 10^{-2} \end{bmatrix}$$

Figure 8.2-2 shows the open-loop resonances at 1.5 rad/s, 21 rad/s with process and measurement noise invariances $\hat{Q} = 1$, $\hat{R} = 10^{-6}I$. The figure actually plots the power spectral densities of y_f, y_a. With H_1' the first row of H', that of y_f is
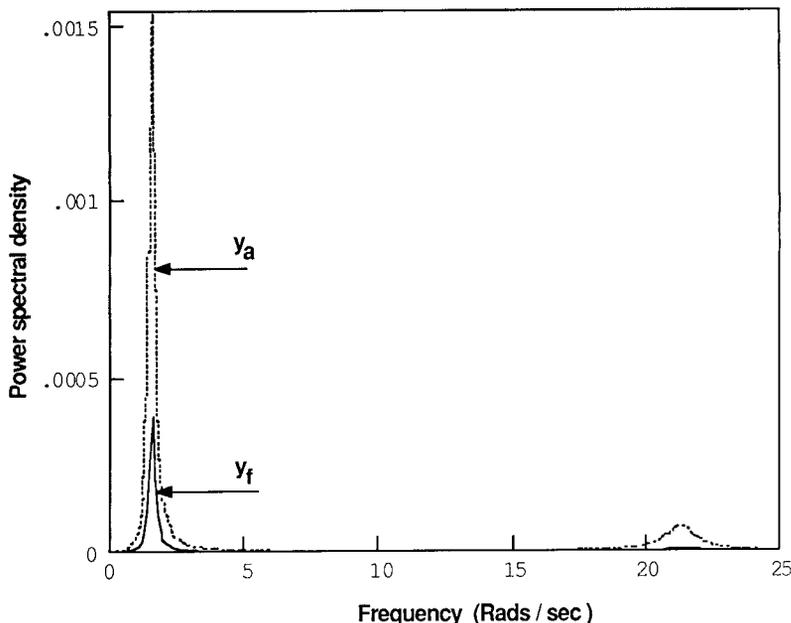
Figure 8.2-2   Open-loop performance showing resonances.

$[H_1'(j\omega I - F)^{-1}\Gamma]^2 \hat{Q}$, due to v, plus the 1-1 entry of $\hat{R}$, due to w. That of y_a is obtained similarly. Figure 8.2-3 shows the closed-loop responses when there is an LQG controller design included. The index is chosen reasonably as

$$V_1 = E[y_f^2 + y_a^2 + 0.2u^2]$$

Notice the reduction in vibration both fore and aft. We comment that further adjustments to the relative magnitude of the individual cost terms in V_1 do not appear to give any improvement. However, since the high resonances are clearly associated with the states x_3, x_4, it makes sense to consider penalizing directly these terms by using an index such as

$$V_2 = E[y_f^2 + y_a^2 + 4x_3^2 + 4x_4^2 + u^2]$$

The dramatic further improvement thereby achieved is shown in Figure 8.2-4.

The use of the frequency response (power spectral density) plots in performance index selection is clearly a crucial ingredient to achieve a good LQG design in this case. We can call the final design a frequency-shaped LQG design. This concept of frequency shaping is developed further in Chapter 9. This example has not, of course, illustrated directly the calculation of a performance index, though the fact that the performance index has a frequency domain interpretation underpins the entire approach of the example.

In the remainder of this section, we turn to a proof of the Separation Theorem.
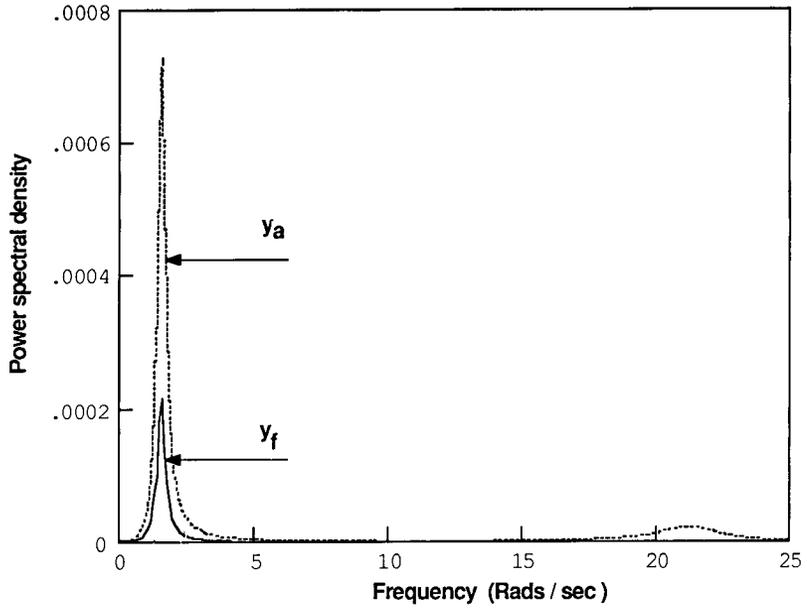
Figure 8.2-3   Closed-loop performance with LQG cost terms ($y_f^2 + y_a^2 + 0.2u^2$).
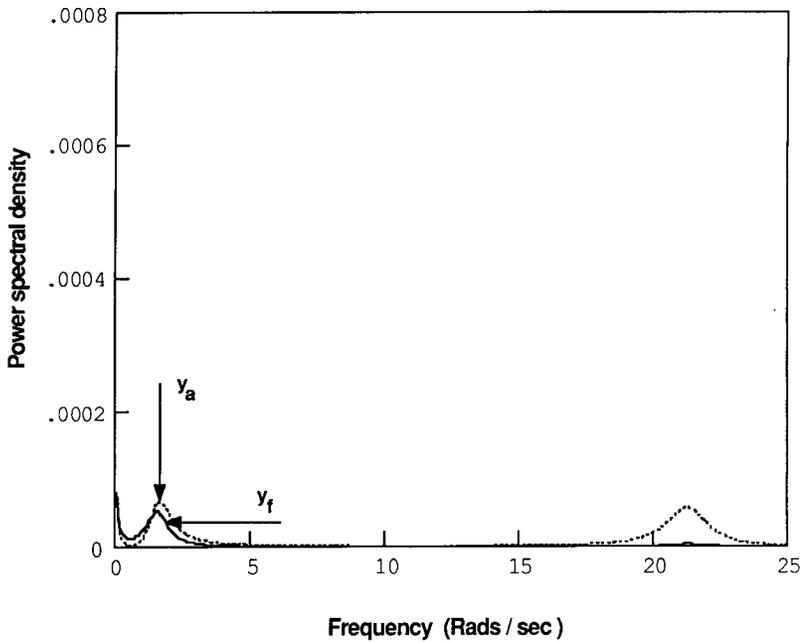
Figure 8.2-4   Closed-loop performance with LQG cost terms ($y_f^2 + y_a^2 + 4x_3^2 + 4x_4^2 + u^2$).

The material is not essential for later concepts in the book, is quite difficult, and can of course be omitted.

Proof of Separation Theorem—Part I.

We appeal to two related results beyond the scope of this text to establish the Separation Theorem in a rigorous manner; see [4]. The first is that the optimal state estimation errors are orthogonal to the optimal state estimate in the sense that

$$E[(x - x_e)x_e'] = 0 \tag{8.2-18}$$

The second is that the innovations process $v(t) = y(t) - H'(t)x_e(t)$ is white, satisfying

$$E[v(t)] = 0, \qquad E[v(t)x_e(t)] = 0, \qquad E[v(t)v'(\tau)] = \hat{R}(t)\delta(t - \tau) \tag{8.2-19}$$

Recall that (8.2-19) is established in Chapter 7, Section 3, for the stationary time-invariant case. Note also that it is readily shown that (8.2-18) holds if a linear control law $u(t) = L(t)x_e(t)$ is used—see Problem 8.2-1. The orthogonality property (8.2-18) follows from what is known as the *Projection Theorem* [4], which is depicted in Fig. 8.2-5. Here $x_e(t)$ is viewed as the orthogonal projection on the space S_t of all random variables obtained through all causal operations in $u(\tau)$, $y(\tau)$ over the interval $[0, t)$.
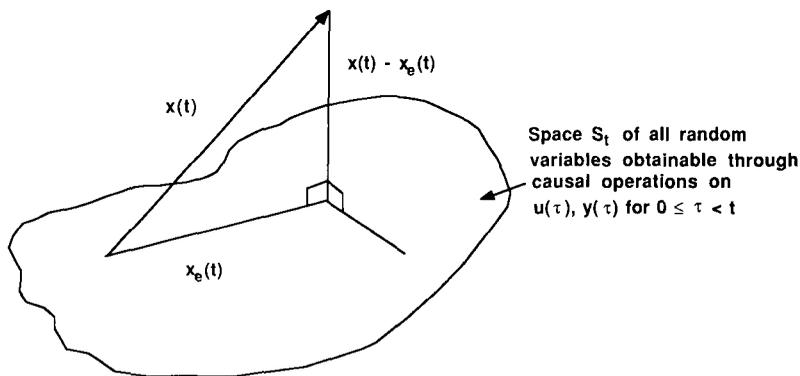
Figure 8.2-5   Projection Theorem.

The above two concepts can, incidentally, be connected. Notice that the innovations $v(t)$ is $y(t) - y_e(t)$, where $y_e(t) = H'x_e(t)$ is the orthogonal projection of $y(t)$ onto the space S_t, so that then $v(t) = y(t) - y_e(t)$ is orthogonal to S_t. Now S_t includes $v(\tau)$ for any $\tau \in [0, t)$, since $y_e(\tau)$ is obtainable by operations on $u(\sigma)$, $y(\sigma)$, $\sigma < \tau$, and $v(\tau) = y(\tau) - y_e(\tau)$. Since $v(\tau) \in S_t$, there holds $E[v(t)v'(\tau)] = 0$ for $\tau < t$, and $v(t)$ is seen to be white, as claimed in (8.2-19).

Part II.

Let us proceed with the Separation Theorem proof. First consider the following reorganization

$$E[x'Qx] = E\{[x_e + (x - x_e)]'Q[x_e + (x - x_e)]\}$$
$$= E[x_e'Qx_e] + E[(x - x_e)'Q(x - x_e)] + 2E[x_e'Q(x - x_e)]$$
$$= E[x_e'Qx_e] + E\{\text{tr}[Q(x - x_e)(x - x_e)']\}$$
$$+ 2E\{\text{tr}[Q(x - x_e)x_e']\}$$

The trace and expectation operators can be interchanged, so that with $P_e \triangleq E[(x - x_e)(x - x_e)']$ and $E[(x - x_e)x_e'] = 0$, then

$$E[x'Qx] = E[x_e'Qx_e] + \text{tr}(QP_e) \qquad (8.2\text{-}20)$$

This result allows a reformulation of the optimal control task, assuming that we can interchange expectation operators and integrations, as follows. Minimize over $u(\cdot)$ the index

$$V = E \int_{t_0}^{T} (x_e'Qx_e + u'Ru)\, dt + \int_{t_0}^{T} \text{tr}(QP_e)\, dt \qquad (8.2\text{-}21)$$

subject to

$$\dot{x}_e = Fx_e + Gu - K_e v \qquad (8.2\text{-}22)$$

where $v = y - H'x_e$ is the white innovations process satisfying (8.2-19). Note that (8.2-22) is just a rewrite of the Kalman filter equation, exploiting the fact that $v(t) = y(t) - H'(t)x_e(t)$. Now (8.2-21) and (8.2-22) together define a stochastic regulation task, but with complete state information, in that x_e is available. It is a special case of the more general study of this section.

Part III. We proceed to the solution of the complete-state information stochastic regulation task (8.2-21), (8.2-22), and we shall make use of (8.2-19) and a result concerning linear systems driven by white noise, which can be derived by using the Ito differential rule. Such derivations are beyond the scope of this text, so we merely quote that with $E[x_e(t)x_e'(t)]$ denoted by $W_e(t)$, and with the control $u(t)$ integrable in that $\int_0^T |u(t)|\, dt < \infty$ almost surely for all finite T, then under (8.2-22).

$$W_e(t_0) = x_e(t_0)x_e'(t_0) = E[x(t_0)]E[x'(t_0)]$$
$$\dot{W}_e = FW_e + W_eF' + K_e\hat{R}K_e' + E[Gux_e' + x_eu'G'] \qquad (8.2\text{-}23)$$

For the case $u \equiv 0$ this result specializes to a well-known result in Appendix B. Also with $u(t) = Lx_e(t)$ this appendix result generalizes immediately to (8.2-23) with $E[Gux_e' + x_eu'G'] = GLW_e + W_eL'G'$. The integrability condition on $u(t)$ is rather weak and will be satisfied by most practical control laws derived from external inputs and state estimates.

Recall that the state estimate feedback gain is, when we delete time arguments,

$$K' = -R^{-1}G'P \qquad (8.2\text{-}24a)$$

$$-\dot{P} = PF + F'P - PGR^{-1}G'P + Q, \qquad P(T) = 0 \qquad (8.2\text{-}24b)$$

Observe now that

$$\frac{d}{dt}(PW_e) = \dot{P}W_e + P\dot{W}_e = (PGR^{-1}G'P - Q - PF - F'P)W_e$$
$$+ P\{FW_e + W_eF' + K_e\hat{R}K_e' + E[Gux_e' + x_eu'G']\}$$

Taking the trace, recalling that tr $(AB) = $ tr (BA), and applying (8.2-24a) yield

$$\mathrm{tr}\left[\frac{d}{dt}(PW_e)\right] = \mathrm{tr}\,[KRK'W_e - QW_e + PK_e\hat{R}K_e' - 2E[x_e'KRu]\}$$

Consequently,

$$E[x_e'Qx_e] = \mathrm{tr}\,(QW_e)$$
$$= \mathrm{tr}\left[-\frac{d}{dt}(PW_e) + PK_e\hat{R}K_e' + KRK'W_e\right] - 2E\,[x_e'KRu] \qquad (8.2\text{-}25)$$

while also

$$E[(u - K'x_e)'R(u - K'x_e)] = E[u'Ru - 2x_e'KRu + x_e'KRK'x_e]$$
$$= E[u'Ru] - E[2x_e'KRu] + \mathrm{tr}\,[KRK'W_e] \qquad (8.2\text{-}26)$$

Now using (8.2-25) and (8.2-26), we can organize the index (8.2-21) as

$$V = \int_{t_0}^{T} E[(u - K'x_e)'R(u - K'x_e) + \mathrm{tr}\,(QP_e) + \mathrm{tr}\,(PK_e\hat{R}K_e')]\,dt - [\mathrm{tr}\,(PW_e)]_{t_0}^{T}$$

which is clearly minimized by

$$u^* = K'x_e$$

with optimal index

$$V^* = \int_{t_0}^{T} \mathrm{tr}\,(QP_e + PK_e\hat{R}K_e')\,dt - [\mathrm{tr}\,(PW_e)]_{t_0}^{T}$$

This completes the proof.

Remark.   We have avoided explicit use of an advanced tool termed the Ito differential rule in this proof, although it is implicit in the Appendix B result on linear systems driven by white noise. This level of technical difficulty is avoided in the discrete-time separation theorem.

Main points of the section.   The practical suggestion that state estimates x_e be used in lieu of states x in a control law is in fact the optimal strategy in the linear quadratic Gaussian (LQG) case of this text. This result is known as the Separation Theorem or Certainty Equivalence Principle. Its proof is based on the Projection Theorem, which tells us that the errors $x - x_e$, v are orthogonal to the measurements, controls, and optimal estimates x_e.

Convenient performance index calculations for the stochastic case, and fre-

quency domain interpretations for the infinite time stationary stochastic cases are available.

Problem 8.2-1.   Consider the stochastic signal model of (8.2-1) and (8.2-2) and optimal minimum variance state estimator of Chapter 7, Section 3 giving optimal estimates x_e with $x_e(t_0) = E[x(t_0)]$. Show that (8.2-18) holds when the control $u = Lx_e$ for some $L(t)$. [*Hint:* Apply the Appendix B output covariance formula for a linear system driven by white noise. Work with an augmented system with states x_e, $x - x_e$ and establish that a certain submatrix of the covariance matrix is zero.]

Problem 8.2-2.   Consider a discrete-time version of (8.2-21) and (8.2-22), as

$$V = E \sum_{t_0}^{T} [x_e'(t)Qx_e(t) + u'(t-1)Ru(t-1)]$$

$$x_e(t+1|t) = Fx_e(t|t-1) + Gu(t) - K_e v(t)$$

(1) Show that with a strict causality constraint on the controller, that is, $u(t)$ can depend only on $y(s)$, $u(s)$ for $s < t$, then the optimal control law is $u^*(t|t-1) = K'x_e(t|t-1)$. (2) Show also that when $u(t)$ can depend on $y(0), y(1), \ldots, y(t)$, and $u(s), s < t$, then

$$u^*(t|t) = K'[x_e(t|t-1) - K_e v(t)]$$

$$= K'[(I + K_e H')x_e(t|t-1) - K_e y(t)]$$

$$= K'x_e(t|t)$$

Note that $v(t) = y(t - H'x_e(t|t-1))$.

8.3 LOSS OF PASSBAND ROBUSTNESS WITH OBSERVERS

Linear quadratic controllers using state estimate feedback have a certain notoriety in the control community. Although optimal for the nominal model, the performance may be far from satisfactory in a real-life situation in which the plant differs somewhat from the model. The guaranteed passband robustness properties established in Chapter 5 for all full-state feedback designs can simply evaporate with the introduction of a state estimator. In this section, we give some details concerning possible loss of robustness, with approaches to recovering robustness being developed more fully in the next sections. Before proceeding, we should note that with the introduction of a full-order state estimator, the roll-off rates improve from 6 dB/octave to 12 dB/octave or more, since the LQG controller is strictly proper, as is the plant. As a consequence, we expect enhanced robustness to unmodeled dynamics at high frequencies.

Figure 8.3-1(a) shows a redrawing of the basic state estimate feedback plant-controller arrangement of Fig. 8.1-1, as a unity positive feedback system. The

(possibly time-varying) nonlinearity in the sector $(\frac{1}{2}, \infty)$. This can be tolerated in the otherwise optimal design without inducing instability. This property specializes to

(a)

(b)
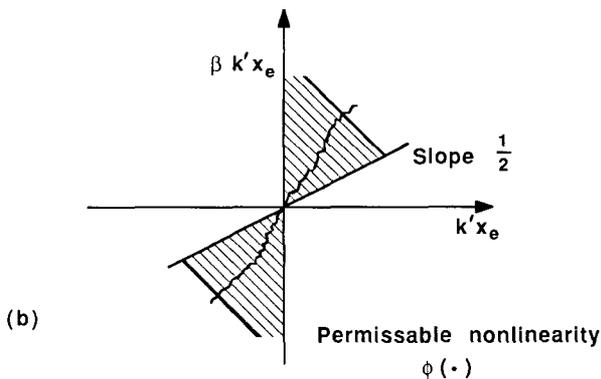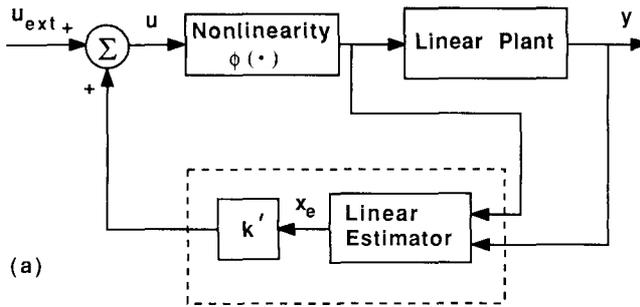
Permissable nonlinearity   **Figure 8.3-2**   Robustness to
$\phi (\cdot)$ insertion of nonlinearities.

saying that at the nonlinearity insertion point there is a gain margin $(\frac{1}{2}, \infty)$ and a phase margin of 60 deg.

Now it is important to understand that this guaranteed robustness occurs where it is not needed. There is not usually any unmodeled nonlinearity or uncertainty $\phi(\cdot)$ where indicated in the control scheme of Fig. 8.3-2. Rather, the more realistic situation is as in Fig. 8.3-3(a), where the control u is fed to the estimator prior to any plant input nonlinearities or uncertainty.

One way out of this dilemma is to introduce the same nonlinearity into the controller as in Fig. 8.3-3(b), so that the estimator once again becomes a true model of the plant. Correct estimation now takes place, and the guaranteed robustness property depicted in Fig. 8.3-2 is recovered.

However, there is a potential difficulty with this somewhat artificial arrangement. If the exact nature of the input nonlinearity to the plant is unknown, then it is impossible to construct the estimator. Of course, if the nature of the nonlinearity is roughly known, and the nonlinearity included in the estimator approximates that in the plant, then presumably performance will be satisfactory.

Another approach to exploiting the guaranteed passband robustness properties at the "wrong" place in the loop is to organize the design of estimator/controller gains K_e/K so that the pathway from the control u to the estimator is relatively inconsequential; that is, the estimator output is far more dependent on the plant
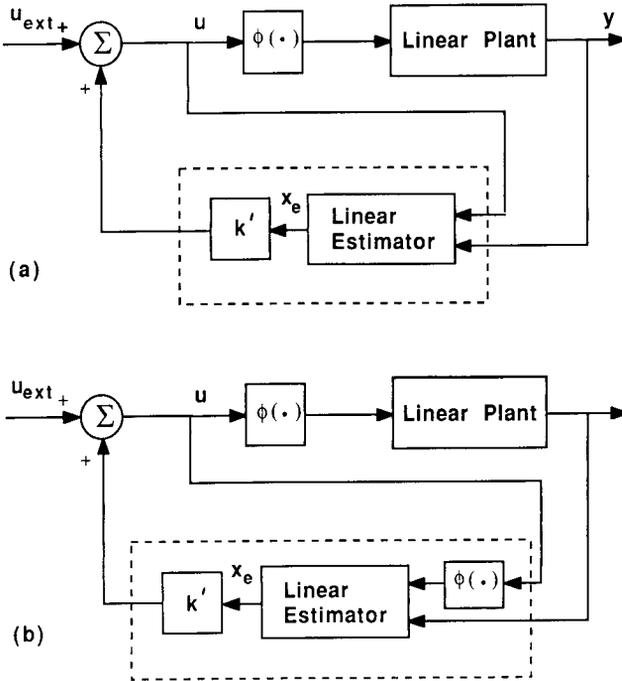
Figure 8.3-3 Introduction of nonlinearity.

output y than on the plant input u. In the extreme, if this pathway is effectively deleted, without upsetting optimality and the associated robustness properties, then the schemes of Figs. 8.3-2, 8.3-3 would be equivalent, and the guaranteed robustness properties would be just as useful as in the full-state feedback design.

A procedure that renders the path between the control u and the estimator less consequential as a parameter σ increases is the loop-recovery technique of the next section.

So far in this section, we have focused on robustness at the plant input. For the single-input, single-output case, of course, plant output robustness is identical to plant input robustness, since the Nyquist plots are identical whether the loop is opened at the plant input or output. Needless to say, for the mutivariable plant case, there is in general a different open-loop gain transfer function matrix for the loop opened at the plant output, being given from

$$V_{OL}(s) = [H'(sI - F)^{-1}G][K'(sI - F - GK' - K_eH')^{-1}(-K_e)] \qquad (8.3-2)$$

Properties of this are not studied further here.

One general reason for loss of robustness on insertion of an estimator is that when the plant differs from its model in the estimator, there is bound to be inaccurate state estimation. This means that even when the state feedback gains are appropriate for the actual plant, the inaccurate state estimates fed back can cause reduced performance or stability properties. A corollary of this remark is that this particular reason for poor robustness (viz., use of inaccurate state estimates) does

not arise when only output nondynamic feedback is used. Such an observation suggests that when one is designing for robustness as distinct from noise immunity, then it may be preferable to work with reduced order observers that have direct feedthrough of plant outputs. Preliminary studies not reported here support this approach.

Can there be any advantage of an LQG design over an LQ design in terms of robustness? We have not shown that there is always a loss of robustness, and indeed one could perhaps devise examples where there is an improvement in robustness properties due to the insertion of an estimator. Certainly, in the scalar input case, the introduction of an estimator will increase the roll-off rate of the open-loop gain matrix at high frequencies. This follows since both the plant and controller are strictly proper so that the open-loop gain will have a roll-off rate of at least 12 dB/octave as compared to 6 dB/octave for $K'(sI - F)^{-1}G$. As a consequence of the higher roll-off rate associated with an LQG design, we would expect improved robustness at frequencies well beyond the cut-off frequency of the LQG design. Of course, at such frequencies, robustness may or may not be a problem. With reduced order observers, the controllers are proper but not strictly proper, and there is no roll-off rate advantage.

For the remainder of this section, we present two examples that point up the possible poor passband input robustness properties of an optimal state-estimate feedback design. One example will be explored further in subsequent sections to demonstrate the benefits of the loop recovery technique and frequency shaping.

We consider now the first demonstration in the literature of a dramatic loss of robustness due to insertion of estimators into a linear quadratic design [5]. The system, with two associated eigenvalues at $+1$, is

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u + \begin{bmatrix} 1 \\ 1 \end{bmatrix} v$$

$$y = \begin{bmatrix} 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + w$$

(8.3-3)

in the notation of this text. Here the noise intensities for v, w are $\hat{Q} = \sigma > 0$, $\hat{R} = 1$, respectively. The performance integral has weights $Q = \rho[1 \quad 1]'[1 \quad 1]$, $R = 1$. [Note that the estimation and control problems have identical (dual) solutions when $\rho = \sigma$.]

Analytical solutions for the gain matrices are

$$K' = -\alpha[1 \quad 1], \qquad \alpha = 2 + \sqrt{4 + \rho}$$

$$K_e' = \beta[1 \quad 1], \qquad \beta = 2 + \sqrt{4 + \sigma}$$

(8.3-4)

The full system matrix is readily constructed as

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & -m\alpha & -m\alpha \\ \beta & 0 & 1-\beta & 1 \\ \beta & 0 & -\beta-\alpha & 1-\alpha \end{bmatrix}$$

with $m = 1$. The case when there is an actual plant input gain m, not necessarily unity, is also given by the above matrix. The characteristic polynomial has the form

$$s^4 + *s^3 + *s^2 + [\beta + \alpha - 4 + 2(m - 1)\alpha\beta]s + 1 + (1 - m)\alpha\beta = 0$$

where $*$ denotes complicated expressions not involving m. It is necessary for stability that the last two coefficients are positive. For sufficiently large α, β (that is, sufficiently large σ, ρ), then there is instability for arbitrarily small perturbations in m from unity in either direction. Thus *LQG designs exist with arbitrary small gain margins*.

As a second example drawn from [6], we use a relatively simple model representing a stable scalar plant that is disturbed by a colored noise process. It is assumed that the low-frequency dynamics are modeled with good precision. However, the plant contains observable and controllable lightly damped high-frequency modes that are not well defined. These are modeled by a reduced-order model. The objective is to design a controller that will reduce the disturbance response of the low-frequency mode.

The plant model and the process-noise model are combined to form the following stochastic model of the nominal plant, depicted in Fig. 8.3-4.

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -100 & -0.2 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 100 \\ 0 \end{bmatrix} u + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0.45 \end{bmatrix} v$$

$$y = \begin{bmatrix} 1 & 10 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + w$$

$$(8.3\text{-}5)$$
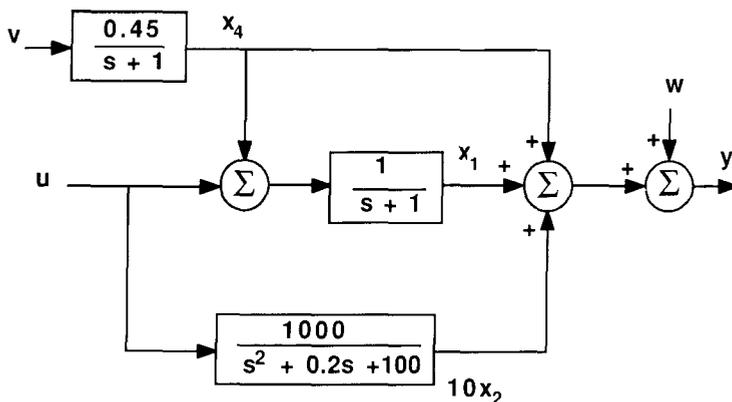
Figure 8.3-4 The stochastic signal model.

Also, here $[x_1 \quad x_2 \quad x_3 \quad x_4]'$ is the state vector, u is a single control, and y is a single measurement. Also v is white process noise and w is white measurement noise with the following properties:

$$E[w(t)] = E[v(t)] = 0, \qquad E[w(t)v(\tau)] = 0$$
$$E[v(t)v(\tau)] = \delta(t - \tau), \qquad E[w(t)w(\tau)] = 0.01\delta(t - \tau) \qquad (8.3\text{-}6)$$

First, a linear quadratic (LQ) regulator is designed based on the following cost function, recalling the formulations as in (8.2-5) and (8.2-6),

$$V = E[4x_1^2 + u^2] = \lim_{T \to \infty} \frac{1}{T} E \int_0^T (4x_1^2 + u^2)\, dt \qquad (8.3\text{-}7)$$

and subjected to the differential equation constraint of (8.3-5). Second, a nominal Kalman state estimator is designed based on the model. This state estimator is inserted into the control loop and forms with the full-state gain matrix a linear quadratic Gaussian (LQG) regulator which we term the *nominal LQG regulator*.

   With the system subjected to process noise and measurement noise with intensities specified by (8.3-6), the value of the cost function, V, (8.3-7) is calculated for three cases: (1) with the control loop open, in that u is set to $u = 0$, (2) with the control loop closed, using the full-state feedback LQ regulator, and (3) with the control loop closed, using the nominal LQG regulator. Table 8.3-1 shows that the regulator performance [as measured by the value of the cost function V, evaluated as in (8.2-16)] of the LQG regulator is within 3 percent of that of the LQ regulator. Both of these control laws provide a significant reduction in the response of the plant to stochastic disturbances. In this particular case, it turns out that most of the response is due to the process noise, and the contribution from the measurement noise is so small that it can be ignored. Thus by means of (8.2-16) and (8.2-17), the index (8.3-7) can be approximately written as

$$V \approx \frac{1}{2\pi} \int_{-\infty}^{\infty} [4|t_1(j\omega)|^2 + |t_2(j\omega)|^2]\, d\omega \qquad (8.3\text{-}8)$$

where $t_1(s)$, $t_2(s)$ are the transfer functions between v and x_1, u. The reason for the small loss in performance when the nominal Kalman filter is inserted into the control loop is illustrated in Fig. 8.3-5, which for the two control laws shows that $|t_1(j\omega)|$ and $|t_2(j\omega)|$ for the LQ and LQG designs are close to one another in the bandwidth of the system. This we would expect, since both regulators are optimal

TABLE 8.3-1   DISTURBANCE RESPONSE

	Performance cost
Open-loop plant	3.64
Closed-loop plant with full-order state feedback LQ regulator	0.77
Closed-loop plant with LQG regulator	0.79
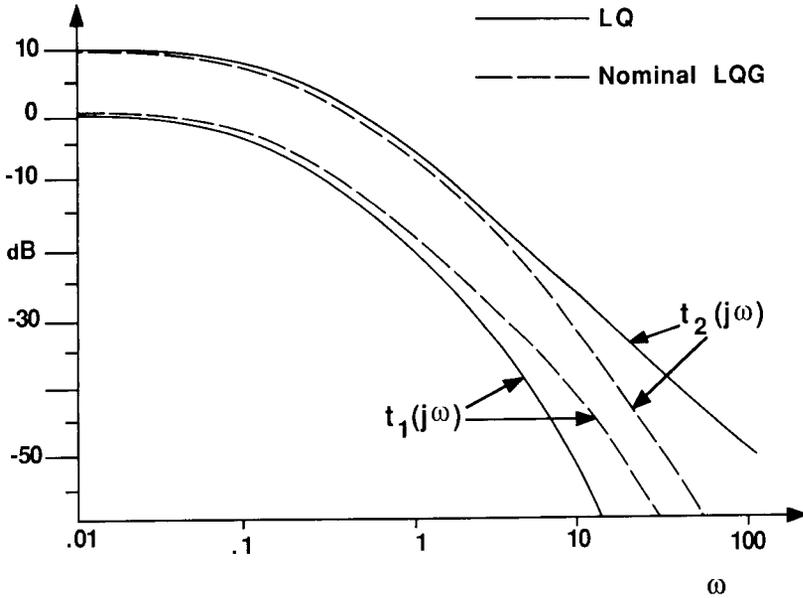
Figure 8.3-5   Magnitudes of $t_1(j\omega)$, $t_2(j\omega)$.

with respect to the same cost function, the effects of measurement noise are very small when compared to the effects of process noise, and state estimation error covariances for the nominal plant case are small.

In order to illustrate the passband robustness of the two control laws, consider Nyquist plots of the open-loop control-loop transfer functions $W_{OL}(j\omega)$ as shown in Fig. 8.3-6. Here $W_{OL}(s)$ is the appropriate specialization of (8.3-1). Applying the classical Nyquist criterion, we see that the LQ regulator has excellent gain and phase margins. However, these same margins are very small and totally inadequate for the nominal LQG regulator. In fact, *the phase margin is approximately 1 deg and the gain margin is for all practical purposes zero*. The LQ regulator has acceptable magnitude roll-off characteristics with increasing frequency. For example, the control-loop gain is reduced to approximately -18 dB at a frequency of 10 rad/s. Since we know that the uncertainty in the model increases with increasing frequency, it is necessary that at these higher frequencies the control-loop gain is as low as -18 dB so as not to destabilize any modeled or unmodeled modes. In fact, such a roll-off characteristic allows arbitrary phase at these higher frequencies. The LQG regulator has a much higher bandwidth than the LQ regulator with a crossover frequency at 10 rad/s. Certainly then, arbitrary phase changes cannot be tolerated at these higher frequencies for the LQG regulator, and there is the possibility that the mode at 10 rad/s and/or unmodeled modes at higher frequencies will be significantly, and unnecessarily, excited.

This example has illustrated the loss of robustness introduced in going from an LQ design to an LQG design.
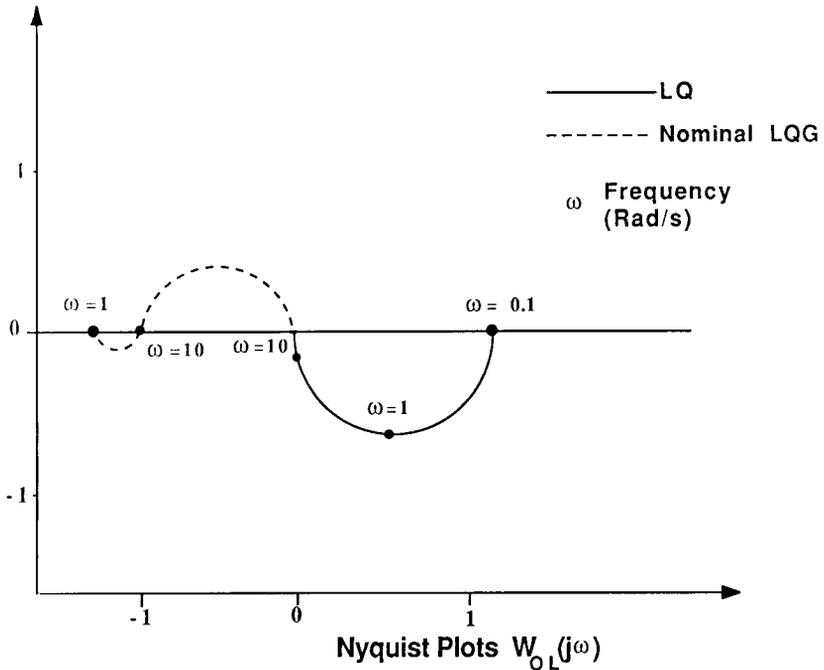
Figure 8.3-6   Control-loop frequency responses.

Main points of the section.   Passband robust properties of optimal state feedback designs can degenerate upon the introduction of a state estimator. On the other hand, high-frequency robustness can improve due to higher roll-off rates for the open-loop gain. It is suggested that poor passband robustness due to poor state estimation in a full-order state estimator can be to some extent ameliorated by using a reduced-order observer with direct feedthrough of plant outputs.

Problem 8.3-1.   Consider the first example. Evaluate the estimator transfer functions from y, u to $K'x_e$. Compare their relative significance as σ, ρ increase and robustness deteriorates.

8.4 LOOP RECOVERY

As noted in previous sections, the attractive passband robustness properties of full-state feedback optimal quadratic designs may disappear with the introduction of a state estimator. Can LQG designs be organized by appropriate weighting matrix selections so that they have the return difference inequality satisfied, and associated robustness properties guaranteed? In other words, can we recover the

loop properties of an LQ design by a suitable adjustment to the LQG design process?

In this section, we first focus on one relatively simple technique, based on work in [7], for state estimator design such that loop recovery takes place for minimum phase plant models. That is, the open-loop gain transfer functions for the LQG designs recover those of the LQ design, and there is recovery in associated robustness properties. We envisage a design situation where one starts with a nominal stochastic model and associated LQG design that lacks the robustness properties of the LQ design. By adding fictitious noise to the plant input model, representing in perhaps a loose way plant variations, uncertainty, or unmodeled dynamics, there is an adjustment to the estimator design. For the case of minimum phase plants there is loop recovery as the fictitious noise becomes larger and larger. In coping with the fictitious noise, the LQG controller becomes more robust to gain and phase changes at the plant input. Of course, there is no longer optimality for the original nominal stochastic model. This is a disadvantage when the response to disturbances as modeled in the nominal model is an important consideration. In this case there must be a trade-off in design between the performance loss for the nominal model and robustness gain. Otherwise, for a deterministic controller design environment when the noise intensities of an assumed original nominal stochastic model are merely convenient design parameters, there is no such trade-off required, although trade-offs between other performance and robustness measures associated with a nominal design may still be important.

A dual technique of loop recovery to recover output robustness properties associated with an estimator design will be discussed.

Loop recovery techniques can be applied to the nonminimum phase plant case with care. In the nonminimum phase plant case, classical control theory tells us that it may be difficult or impossible to design a robust controller. Such difficulties are not expected to be hurdled by any simple loop recovery technique. Some methods and results for this case are noted in this and the next section.

At this stage, we caution that loop recovery is one technique to achieve a restricted class of robustness properties. It may not achieve robustness to unmodeled dynamics or certain plant parameter variations. It is certainly not a universal panacea for solving robustness problems.

Asymptotic estimator properties.   Here we exploit some dual properties to the asymptotic regulator properties of Chapter 6. We consider time-invariant stochastic plant models, as usual stabilizable and detectable:

$$\dot{x} = Fx + Gu + v, \qquad y = H'x + w \tag{8.4-1}$$

The associated optimal estimator/regulator is

$$\dot{x}_e = (F + GK')x_e - K_e(y - H'x_e), \qquad u = K'x_e \tag{8.4-2}$$

and it is also time-invariant.

For the main results of this section, we assume that the plant $H'(sI - F)^{-1}G$ is square, and is such that

$$H'(sI - F)^{-1}G \text{ is nonsingular in Re } s \geq 0 \qquad (8.4\text{-}3)$$

For nonsquare plants that have full column rank or full row rank in Re $[s] \geq 0$, results can be achieved, but they are more complicated. They involve building a square minimum phase plant out of a nonsquare one [8].

Asymptotic properties of the estimator are now studied with the intensity of the process noise added to the plant input becoming infinite. Thus we assume

$$E[v(t)v'(\tau)] = \hat{Q}_\sigma \delta(t - \tau), \qquad \hat{Q}_\sigma = \sigma GG'$$
$$E[w(t)w'(\tau)] = \hat{R}\delta(t - \tau), \qquad \hat{R} > 0 \qquad (8.4\text{-}4)$$

and study the situation when the positive scalar σ becomes infinite. The first property is a direct dual of that in Sec. 6.2, namely that under the minimum phase assumption (8.4-3) and with $[F, H]$ not just detectable but in fact completely observable,

$$\lim_{\sigma \to \infty} \sigma^{-1/2} K_{e\sigma} \hat{R}^{1/2} = GV \qquad (8.4\text{-}5)$$

for some orthogonal V, where $K_{e\sigma}$ denotes K_e parametrized by σ. This property follows from the spectral factorization relation (7.3-33); details are omitted here.

From (8.4-5), we can argue that the estimator state x_e becomes much more influenced by y than by u; consider, for example, Fig. 7.2-1, and observe that the estimator inputs are coupled through gain matrices G and K_e (or here $K_{e\sigma}$). Now (8.4-5) shows that one of these, $K_{e\sigma}$, tends to infinity as $\sigma \to \infty$. This asymptotic property suggests that the schemes of Figures 8.3-3(a) and 8.3-3(b) may become equivalent as $\sigma \to \infty$, so that the effect of the nonlinearity $\phi(\cdot)$ becomes the same for both arrangements. In other words, the robustness properties of the full state feedback design may be recovered. We now move on to give greater precision to this robustness recovery, also termed loop recovery.

From (8.4-5), a further limiting property can be readily established; for all finite ω,

$$\lim_{\sigma \to \infty} [I - (j\omega I - F)^{-1}K_{e\sigma}H']^{-1}(j\omega I - F)^{-1}G = 0 \qquad (8.4\text{-}6)$$

To see this, note that asymptotically the left side is, when we recall (8.4-5),

$$[I - \sigma^{1/2}(j\omega I - F)^{-1}GV\hat{R}^{-1/2}H']^{-1}(j\omega I - F)^{-1}G$$
$$= (j\omega I - F)^{-1}G[I - \sigma^{1/2}V\hat{R}^{-1/2}H'(j\omega I - F)^{-1}G]^{-1}$$

which in turn approaches zero as $\sigma \to \infty$ when $H'(j\omega I - F)^{-1}G$ is nonsingular, and this is guaranteed by (8.4-3) for all finite ω.

State feedback regulator loop recovery.   Let us now apply the asymptotic property (8.4-6) to examine the behavior of the loop gain transfer function matrix

$$W_{OL}(s) = C_2(s)P(s) \tag{8.4-7a}$$

$$P(s) = H'(sI - F)^{-1}G \qquad \text{(plant)} \tag{8.4-7b}$$

$$C_2(s) = K'(sI - F - GK' - K_{e\sigma}H')^{-1}(-K_{e\sigma}) \quad \text{(compensator)} \tag{8.4-7c}$$

A left factorization of $C_2(s)$, easily checked, is

$$C_2(s) = X_L^{-1}(s)Y_L(s) \tag{8.4-8a}$$

$$X_L(s) = I - K'(sI - F - K_{e\sigma}H')^{-1}G \tag{8.4-8b}$$

$$= I - K'[I - (sI - F)^{-1}K_{e\sigma}H')^{-1}(sI - F)^{-1}G$$

$$Y_L(s) = K'(sI - F - K_{e\sigma}H')^{-1}(-K_{e\sigma}) \tag{8.4-8c}$$

[In the notation of the last section $Y_L(s) = C_3(s)$ and $X_L^{-1}(s) = C_1(s)$.] Application of (8.4-6) gives immediately that, for all finite ω

$$\lim_{\sigma \to \infty} X_L(j\omega) = I, \qquad \lim_{\sigma \to \infty} W_{OL}(j\omega) = Y_L(j\omega)P(j\omega) \tag{8.4-9}$$

Moreover, since

$$K'(sI - F)^{-1}G - Y_L(s)P(s) = K'[I + (sI - F - K_{e\sigma}H')^{-1}K_{e\sigma}H'](sI - F)^{-1}G$$

$$= K'[I - (sI - F)^{-1}K_{e\sigma}H']^{-1}(sI - F)^{-1}G$$

taking limits as $\sigma \to \infty$ for $s = j\omega$, ω finite, and applying (8.4-6) and (8.4-9) gives immediately an asymptotic property known as the *loop recovery property*:

$$\lim_{\sigma \to \infty} W_{OL}(j\omega) = K'(j\omega I - F)^{-1}G \qquad \text{for all finite } \omega \tag{8.4-10}$$

Summarizing then, we have established:

> **State feedback regulator loop recovery property.**   Consider a state-estimate feedback regulator for a stabilizable, completely observable, time-invariant, standard stochastic process model of this chapter. Suppose the plant is strictly minimum phase in that (8.4-3) holds. Consider the plant noise process (8.4-4) parametrized in terms of σ and let the filter gain K_e also parametrized in terms of σ be denoted by $K_{e\sigma}$. Then the associated control scheme loop gain transfer function matrix $W_{OL}(s)$ has the limiting property, (8.4-10).

Loop recovery and controller design.   When a plant has the same number of inputs or, as noted in [8], more outputs than inputs, then there is a chance of exploiting loop recovery as so far described. (Certain changes are needed for nonsquare plants.) When there are the same number of inputs or more inputs than outputs, the dual loop recovery technique described later in this section can be used (again with changes for nonsquare plants). The loop recovery property tells us that for stabilizable, observable minimum phase plants satisfying (8.4-3), the loop gain transfer function matrix in a full-state feedback design, namely

$K'(j\omega I - F)^{-1}G$, is recovered in a full-order state estimate feedback design under a certain limiting operation. As the signal model plant process noise injected at the plant input becomes infinitely intense, then loop recovery takes place. The consequence is that the return difference inequality is guaranteed to be satisfied in the limit, and this means that the passband input robustness properties guaranteed by this inequality are recovered. Thus in the scalar input case, phase margins of 60 deg are approached and gain margins of $(\frac{1}{2}, \infty)$ can be achieved, and so on. It is very important that for each injected noise intensity σ, the parameter of a class of designs, the controller designs are stabilizing for the nominal plant, so that σ can be selected to achieve an acceptable design *without* necessarily having to become very large.

Are there disadvantages to applying loop recovery with σ large to achieve robust designs? The poor roll-off rate of 6 dB/octave associated with a full-state design is also "recovered" at finite frequencies, and poor high-frequency robustness (to unstructured plant uncertainty) can accrue. Another disadvantage is that loop recovery relies on pole-zero cancellations in the open loop. Thus in the scalar variable case, plant zeros are cancelled and replaced by zeros of $K'(j\omega I - F)^{-1}G$. (This explains why the plant must be minimum phase with left half-plane zeros.) If the plant zeros are lightly damped, then any approximate cancellation due to plant variations from the nominal one could cause gross differences between the desired nominal open-loop transfer function and that actually achieved. Even with exact pole-zero cancellations of lightly damped zeros, there are concerns with oscillations associated with these modes during transients. More generally, designers are cautious of loop designs with high loop gains because of intolerance of unmodeled dynamics and plant parameter variations.

We have already foreshadowed that a controller exploiting loop recovery is no longer optimal for any original nominal stochastic plant, because the noise statistics adopted for loop recovery design differ from the actual noise statistics. This loss of optimality may be a severe disadvantage in terms of achieving good disturbance response characteristics.

How then can the loop recovery property be exploited in practice for controller design purposes? Let us consider a stabilizable, observable, minimum phase plant where we seek a controller and our use of the stochastic model only facilitates the design. Here then σ is merely a convenient design parameter, as may be Q, R, and $\hat{R}$. First a full-state feedback design is carried out with acceptable input passband robustness properties. Next a number of stochastic designs are tried with increasing values of σ until there is adequate loop recovery as measured by input passband robustness properties. The value of σ is not increased excessively to avoid problems associated with high gain loops. Some compromise between input robustness properties, and other robustness/performance measures may have to be reached. An example is studied later in the section which shows that there may be no reasonable compromise, and the more powerful frequency-shaped loop recovery techniques may be needed.

Frequency-shaped loop recovery.   In frequency-shaped loop recovery, the "fictitious" plant process noise is frequency-shaped in that its intensity is frequency dependent. We typically model the fictitious noise as arising from a linear system (shaping filter) driven by white noise. For the case when the fictitious noise is limited largely to a certain frequency band, then loop recovery will take place more so in this frequency band, and the system behavior will be relatively unaffected outside this band. In practice, if there are unacceptably poor phase margins in an LQG design, it is reasonable to inject fictitious plant noise in the vicinity of the cross-over frequency, and possibly in a frequency band where there is high plant input (actuator) uncertainty for a loop recovery LQG design. It is usual to use low-order high-pass, low-pass, or bandpass filters driven by white noise to generate the fictitious noise. These filters augment the plant model, as illustrated in an example later in the section; see also Problem 8.4-1. A rigorous asymptotic theory as the noise intensity increases can be derived by applying the known results for the standard (nonfrequency shaped) loop recovery to the augmented plant models; details are in [6].

If an LQG design is modified by the inclusion of frequency-shaped loop recovery, the closed-loop characteristics outside the frequency band of the loop recovery are not affected. If loop recovery takes place in a low-frequency band, for example, the roll-off rate characteristics of the original LQG design are preserved. A design example is given later in the section to show the advantages of, and perhaps the necessity of, frequency shaping when loop recovery is applied.

The notion of frequency-shaped LQG designs, including frequency-shaped loop recovery, is developed further in the next section and chapter.

The case of nonminimum phase plants.   What, then, about the nonminimum phase plant situation? A property of linear quadratic loop gain transfer functions $K'(sI - F)^{-1}G$ is that they are minimum phase. Thus to recover such open-loop transfer functions with a series compensator a plant must be minimum phase. Otherwise, there would have to be open-loop unstable pole-zero cancellations leading to instability. Even so, there are two possible roles for the loop recovery ideas so far described in designing state estimate feedback controllers for nonminimum phase plants. These are now studied in turn.

First trial approach.   One approach to applying loop recovery techniques to nonminimum phase plants is to proceed tentatively as if the plant were minimum phase. Just increase noise variances until there is a maximum degree of robustness enhancement—never expecting a full recovery of state feedback open-loop properties, nor expecting continual improvement in robustness as the noise variance continues to increase. In this approach frequency shaping could well permit adequate loop recovery in the frequency band of interest, as long as the nonminimum phase zeros are outside this band. This general approach is studied in more detail in [9] when there is present one right half-plane zero. As expected, the

approach is shown to work reasonably well when this zero is outside the frequency band of the feedback loop.

A factored plant model approach. In exploring further the application of loop recovery ideas when there are nonminimum phase plants, it proves convenient to work with a minimum-phase, all-pass factorization of the plant $P(s) = P_{MP}(s)P_{AP}(s)$, as in Fig. 8.4-1; see also [10]. The stable all-pass factor has its transfer function $P_{AP}(s) = J_{AP} + H'_{AP}(sI - F_{AP})^{-1}G_{AP}$ satisfying $P'_{AP}(-s)P_{AP}(s) = I$, so that it introduces only phase changes. Its zeros are the right half-plane zeros of the plant. Spectral factorization techniques lead to achieving a minimum phase factor $P_{MP}(s) = H'_{MP}(sI - F_{MP})^{-1}G_{MP}$ associated with a plant $P(s)$, having the same poles as $P(s)$. The factorization is straightforward in the scalar plant case, and more complex in the nonscalar case, as studied in [11]. Details for this latter case are not studied here. The factored plant is assumed to have an all-pass factor state x^{AP} and a minimum phase factor state x^{MP}. The plant has a nonminimal state vector $x' = [x^{AP'} \ x^{MP'}]$, and the nonminimal modes are all unobservable and stable, being poles of $P_{AP}(s)$ if x^{AP} and x^{MP} are separately minimal.

A factored plant model design approach is now described.

Step 1. First design a stabilizing state feedback law for the nonminimal model of Figure 8.4-1 as

$$u_1 = K'_{AP}x^{AP} + K'_{MP}x^{MP} + u_{\text{ext}} \tag{8.4-11}$$

Nonminimal plant representation P(s)
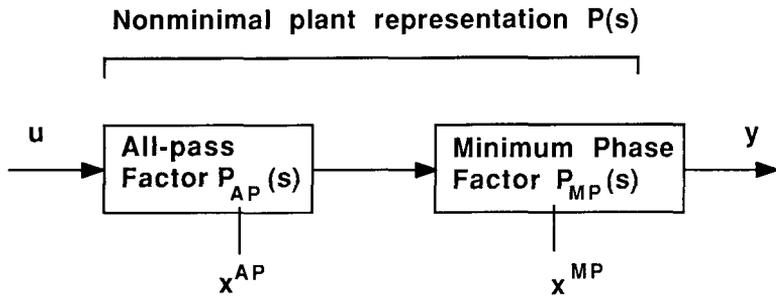
Figure 8.4-1   Plant factorization.

Such a design can be achieved by using the linear quadratic approach to achieve a robust design.

Step 2. Achieve a partial state estimate feedback design as follows. Consider a state estimate of x^{AP}, denoted x_e^{AP}, constructed from a parallel model driven only from u, and not also from y, as follows: $x_e^{AP}(s) = (sI - F_{AP})^{-1}G_{AP}u(s)$. Consider also

$$u_2 = K'_{AP}x_e^{AP} + K'_{MP}x^{MP} + u_{\text{ext}} \tag{8.4-12}$$

Notice that since $P_{AP}(s)$ is stable, $x_e^{AP}(t) \to x^{AP}(t)$ as $t \to \infty$ for arbitrary plant inputs u (including u_2, of course), so that $u_2(t) \to u_1(t)$ as $t \to \infty$, and u_2 is stabilizing as is u_1.

Step 3.   Organize the control u_2 as dynamic feedback of x^{MP} as follows:

$$u_2(s) = K'_{AP}(sI - F_{AP})^{-1}G_{AP}u_2(s) + K'_{MP}x^{MP}$$
$$= K'_{MP}(s)x^{MP}(s) \tag{8.4-13a}$$

$$K'_{MP}(s) = [I - K'_{AP}(sI - F_{AP})^{-1}G_{AP}]^{-1}K'_{MP} \tag{8.4-13b}$$

This is a stabilizing controller leading to closed-loop modes being those of the state original feedback design and those of $P_{AP}(s)$.

Step 4.   Implement the state estimate feedback law

$$u_3(s) = K'_{MP}(s)x_{e\sigma}^{MP}(s) + u_{\text{ext}}(s) \tag{8.4-14}$$

where $x_{e\sigma}^{MP}$ is a state estimate of x^{MP} with fictitious noise variance σ, as for the minimum phase plant case, save that now the nominal plant model is used with input u, output y, and states $x = [x^{AP'}x^{MP'}]'$. Thus $x_{e\sigma}^{MP}$ is in fact a subvector of the full state estimate $x_{e\sigma}$. The associated open-loop transfer function of this design is denoted $W_{OL}(s)$. It is easily checked that it has the structure $X(s)P_{AP}(s)$ for some $X(s)$.

We use the same theoretical techniques as those for establishing the standard loop recovery property; the relevant generalization is as follows:

Loop recovery property.   In the notation of this section,

$$\lim_{\sigma \to \infty} W_{OL}(j\omega) = K'_{MP}(s)(sI - F_{MP})^{-1}G_{MP}P_{AP}(s) \tag{8.4-15}$$

Proof.   We present here a simple heuristic proof outline. The state estimation of the states x^{AP}, x^{MP} for the model of Figure 8.4-1 leads to estimates via

$$\dot{x}_e^{MP}(t) = (F_{MP} + K_{e\sigma}H'_{MP})x_e^{MP}(t) - K_{e\sigma}y(t)$$
$$+ G_{MP}[H'_{AP}x_e^{AP}(t) + J_{AP}u(t)]$$

In fact, $K_{e\sigma}$ is identical to what is obtained for the case when $P_{AP}(s)$ is set to I (not proved here). Thus $K_{e\sigma}$ satisfies (8.4-5) with G_{MP} replacing G as

$$\lim_{\sigma \to \infty} \sigma^{-1/2}K_{e\sigma}\hat{R}^{1/2} = G_{MP}V$$

for some orthogonal V. As a consequence, as $\sigma \to \infty$, then in calculating x_e^{MP} the term $K_{e\sigma}y(t) = \sigma^{1/2}G_{MP}Vy(t)$ dominates the term $G_{MP}[H'_{AP}x_e^{AP}(t) + J_{AP}u(t)]$. This domination is a mild generalization of that for the minimum phase case, and leads likewise to the loop recovery property. Further details are omitted; see [10].

Clearly, this result specializes to the standard one given earlier in this section when the plant is minimum phase so that $P_{AP}(s) = I$, $P_{MP}(s) = P(s)$, and $K'_{MP}(s) = K'$.

We stress again that since the designs do not achieve minimum phase open-loop transfer functions, a property of linear quadratic designs, they will not achieve the input robustness properties in terms of the return difference inequality of the linear quadratic designs. Even so, if a design with acceptable performance and robustness is achieved with the control law (8.4-13), then loop recovery can be achieved in a state estimate feedback design.

For some situations, it makes sense to replace the plant $P(s)$ for design purposes by its minimum phase factor $P_{MP}(s)$. Thus a state estimate feedback design is implemented on the assumption that the plant has $P_{AP}(s) \equiv I$. The resulting design must be robust to the inclusion of $P_{AP}(s)$ into the loop at the nominal plant input. For example, let us consider a scalar input plant case such that in the frequency band where loop gains exceed unity, the all-pass factor introduces phase changes $\Delta\varphi(j\omega)$ small relative to $\pi/3$. Then it is reasonable to set $P_{AP}(s) = I$ for LQG designs with loop recovery, since such designs achieve phase margins of near $\pi/3$ degrees, and thus could tolerate the introduction of the all-pass factor with its small phase changes in the region of the cross-over frequency.

It now makes sense to ask for what class of plants would the all-pass factor introduce relatively small phase changes in the vicinity of the cross-over frequency. A loose answer is when the original plant has right half-plane zeros near the $j\omega$ axis and outside the cross-over frequency region.

If the all-pass factor in the above discussion has the property that $|\Delta\varphi(j\omega) + \pi|$ is small relative to $\pi/3$ where the loop gain exceeds one, and in the cross-over frequency band, then the all-pass factor can be approximated by a simple gain of -1. This situation arises, in loose terms, when the original plant has (an odd number of) zeros with large positive real parts as compared with the bandwidth where the loop gain exceeds 1.

Dual asymptotic regulator property.

So far, in discussing loop recovery, we have focused on recovering input robustness properties of a full-state feedback design. There are dual results for plant output robustness, relying on the regulator asymptotic properties of Chapter 6. In these, the estimator is a standard full-state estimator as in Sec. 7.3, but the regulator design is modified. Thus the controller Q matrix is selected as

$$Q = \rho HH' \qquad (8.4\text{-}16)$$

and the following asymptotic regulator property of Chapter 6 is exploited. Under the plant minimum phase assumption (8.4-3), and with $[F, G]$ not just stabilizable but completely controllable, then

$$\lim_{\rho \to \infty} \rho^{-1/2} R^{1/2} K'_\rho = WH' \qquad (8.4\text{-}17)$$

for some orthogonal W. Here K_ρ denotes the state feedback gain K parametrized by ρ. This result leads to duals of the estimator asymptotic properties including the loop recovery property given earlier. In particular, the dual of (8.4-6) is, for all finite ω,

$$\lim_{\rho \to \infty} H'(j\omega I - F)^{-1}[I - GK'_\rho(j\omega I - F)^{-1}] = 0 \qquad (8.4\text{-}18)$$

This result now leads to the dual loop recovery properties explored in the next subsection.

Dual state estimator loop recovery. Now we define a loop gain transfer matrix with the loop opened at the plant output, as

$$V_{0L}(s) = P(s)C_2(s) = P(s)Y_R(s)X_R^{-1}(s) \qquad (8.4\text{-}19)$$

where

$$P(s) = H'(sI - F)^{-1}G \qquad \text{(plant)}$$

$$C_2(s) = K'_\rho(sI - F - GK'_\rho - K_e H')^{-1}(-K_e) \quad \text{(compensator)} \qquad (8.4\text{-}20)$$

and

$$X_R(s) = I - H'(sI - F - GK'_\rho)^{-1}K_e$$

$$Y_R(s) = K'_\rho(sI - F - GK'_\rho)^{-1}(-K_e) \qquad (8.4\text{-}21)$$

Dualizing the loop recovery properties then gives, for all finite ω, under (8.4-4)

$$\lim_{\rho \to \infty} X_R(j\omega) = I, \qquad \lim_{\rho \to \infty} V_{0L}(j\omega) = P(j\omega)Y_R(j\omega)$$

$$\lim_{\rho \to \infty} V_{0L}(j\omega) = H'(j\omega I - F)^{-1}K_e \qquad (8.4\text{-}22)$$

Now as $\rho \to \infty$, the control scheme of Fig. 8.4-2(a) behaves as that of Fig. 8.4-2(b). This tells us that the estimator design is crucial to the closed-loop system behavior. Now consider the return difference inequality associated with Fig. 8.4-2(b), namely

$$[I - H'(j\omega I - F)^{-1}K_e]\hat{R}[I - K'_e(-j\omega I - F')^{-1}H] \geq \hat{R} \qquad (8.4\text{-}23)$$

for all ω, where $\hat{R}$ is the measurement noise covariance matrix. This inequality guarantees robustness properties at the plant output. For example, the inequality (8.4-23) guarantees in the scalar case phase margins of 60 deg and gain margins of $(\frac{1}{2}, \infty)$. Indeed there can be tolerated output nonlinearities (possibly time-varying) in the sector $(\frac{1}{2}, \infty)$, without disturbing stability.

We conclude that in this dual asymptotic result there is a form of estimator loop recovery and output passband robustness recovery. There is also, unfortunately, a recovery of the poor roll-off rates of 6 dB/octave (20 dB/decade). Also with high loop gains, there is again the possibility of poor robustness to unmodeled dynamics and plant parameter variations.

Design using dual loop recovery. When there is concern about output robustness, or when a plant has more inputs than outputs, or when simply there is a desire to achieve a certain loop gain, then it makes sense to consider this dual loop recovery technique (with adjustment in the nonsquare plant

$$-V_{OL}(s)$$

(a)

(b)
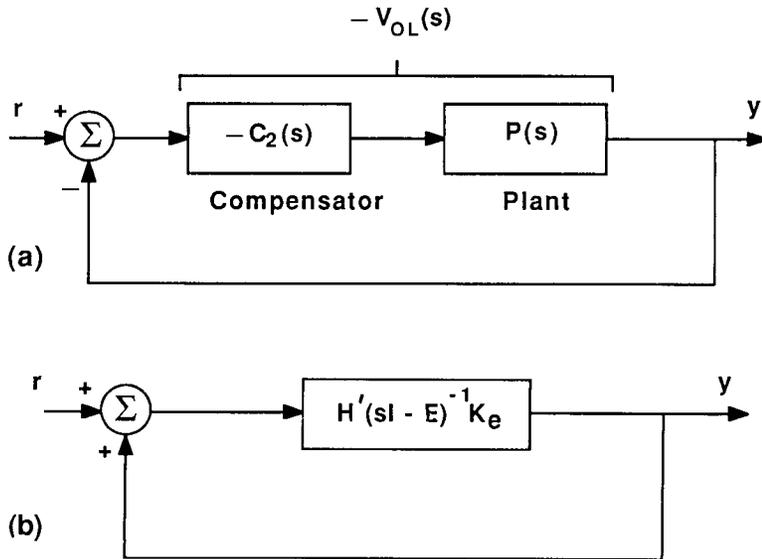
Figure 8.4-2 Two control schemes.

case). The estimator design for K_e remains unchanged, while the state feedback control law design K_ρ, parametrized in terms of ρ, is adjusted with trial values. The value of ρ is increased until the desired loop gain is adequately approximated, or there is acceptable output robustness, without too much loss in noise rejection capability, and other performance robustness properties associated with an original design with $\rho = 0$. Notice that the unity negative feedback version of the controller design as depicted in Fig. 8.4-2(a) approaches that of the estimator loop of Fig. 8.4-2(b) as $\rho \to \infty$. Ironically then, the estimator design dominates the controller behavior as ρ becomes large.

For the nonminimum phase plant situation, then the case $\rho \to \infty$ will not lead to loop recovery—the dual of the earlier situation. Frequency shaping has a role in this dual design procedure—the model of the plant is augmented at the output by a band limiting filter, giving an augmented output y^f which is then penalized in the regulation performance index in lieu of y. Such designs are the subject of the next chapter. There is also a dual of the factored plant method for achieving loop recovery for nonminimum phase plants.

Design example.   Consider the plant model (8.3-5)–(8.3-6) of the previous section, but with fictitious noise v_f, representing plant input uncertainty, injected at the plant input. The equations are

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x} \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -100 & -0.2 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} u + \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 100 \\ 0.45 & 0 \end{bmatrix} \begin{bmatrix} v \\ v_f \end{bmatrix} \qquad (8.4\text{-}24)$$

$$y = x_1 + 10x_2 + x_4 + w \qquad (8.4\text{-}25)$$

We consider first the cases when v_f is white noise. The cost index is again (8.3-7), repeated as

$$V = E[4x_1^2 + u^2] \qquad (8.4\text{-}26)$$

The plant here, even apart from v_f, is a stochastic one, and noise performance is important. A full loop recovery by using white noise v_f of high intensity turns out to achieve satisfaction of the return difference inequality and associated input robustness, as expected by the theory, but very poor rejection of the noise terms v, w, as subsequently depicted.

Let us look then more closely at our design objectives. We require robustness to unmodeled lightly damped dynamics at high frequency. Thus we require that the LQG regulator exhibit the same or better magnitude roll-off as the LQ design beyond a frequency of 5 rad/s. More specifically, as discussed in the last section, we require the control-loop gain to be equal to or less than −18 dB at a frequency of 10 rad/s. Recall from the last section that the LQ design certainly achieves this objective, but not the original LQG design. Since, in this case, the model is fairly accurate at the lower frequencies, it is not necessary to modify the optimum LQG controller characteristics to match the (excellent) gain and phase characteristics of the LQ design below 5 rad/s at a performance cost at low frequencies. Instead, we impose the requirements that the cross-over frequency be at or below 1 rad/s and that the minimum gain margin and phase margin are equal to or better than 2 dB and 20 deg, respectively. This approach is typical for many practical situations with imprecisely modeled high-frequency dynamics.

With the above objectives in mind, it is proposed here to employ frequency-shaped loop recovery. The fictitious noise, representing plant uncertainty, is emphasized in the frequency band above 5 rad/s.

Thus we consider for this second case the plant (8.4-24) augmented with a filter driven by white noise η as

$$\dot{x}_5 = -10x_5 + \eta, \qquad v_f = -10x_5 + \eta \qquad (8.4\text{-}27)$$

The transfer function from η to v_f is $s(s + 10)^{-1}$, which is a simple high-pass filter. More sophisticated band-limited filter designs turn out to give no significant further benefit so are not discussed further. We expect to achieve a greater degree of loop recovery (input robustness) at frequencies above 5 rad/s than below with a consequent loss of performance significant only in the handling of signals above 5 rad/s, where the presence of the fictitious noise renders the modified LQG design nonoptimal. With the filter (8.4-27) specified, then applying loop recovery techniques to the augmented plant of (8.4-24) and (8.4-27) can be interpreted as frequency-shaped loop recovery for the original plant.

Several modified LQG regulators are designed for the two cases (standard loop recovery and frequency-shaped loop recovery) using various intensities for the fictitious noise. The trade-off between performance and robustness as a function of fictitious noise intensity is shown implicitly in Fig. 8.4-3, where the vertical axis is
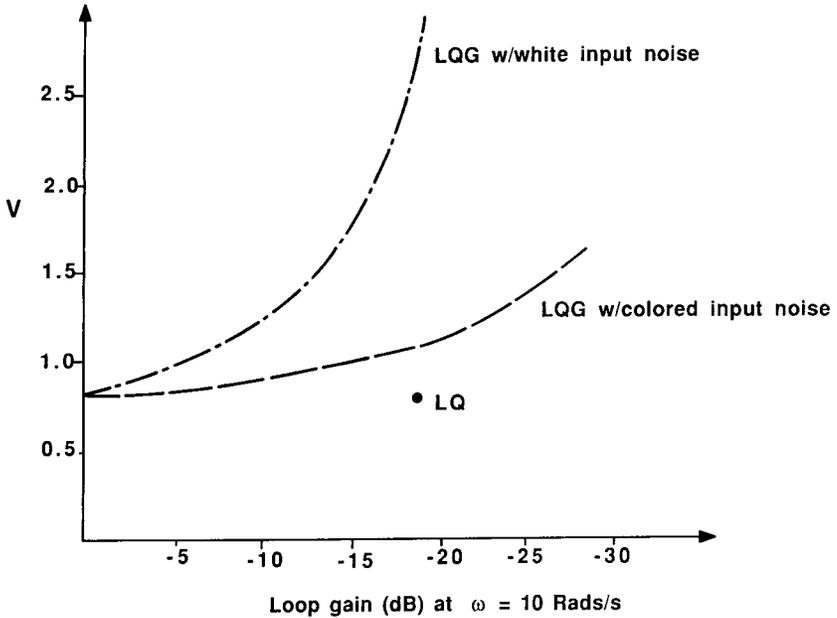
Figure 8.4-3   Performance versus control-loop gain roll-off.

the value of the cost index V, and the horizontal axis shows an appropriate robustness measure for this example, namely, the loop gain at the critical frequency of 10 rad/s. (Note that the fictitious noise is used to design the controller, but not to compute the performance V, precisely because it is fictitious.) We see that if we attempt to achieve a small loop gain at 10 rad/s for robustness purposes, for example −18 dB, then the cost index increases dramatically in the white noise v_f case, but remains close to the LQ cost with the frequency-shaped noise injection. The performance index V takes values that are large when the effects of the disturbance $v(\cdot)$ are not well countered by $u(\cdot)$; in frequency domain terms, V becomes large when the transfer functions $t_1(j\omega)$, $t_2(j\omega)$ from v to x_1 and v to u respectively have large mean square values; see the approximate expression for V of (8.3-8). The amplitude Bode diagram for $t_1(j\omega)$, $t_2(j\omega)$ thus gives insight into the values of V obtained. Looking at Figs. 8.4-4, 8.4-5, we see that using the colored noise rather than white noise has less impact on the frequency response between the process noise and the variables in the cost function. Fig. 8.4-6 (which should be compared with Fig. 8.3-6), shows that when fictitious white input noise is used, the control loop frequency response approaches that of the LQ regulator, in particular with regard to the phase characteristics. On the other hand, when colored fictitious input noise is used, the control-loop frequency response is modified only enough to meet our design requirements for gain roll-off at high frequency as well as gain and phase margins within the control bandwidth.

Many examples illustrating the features of loop recovery-based design are studied in the literature—see, for example, [7, 12].
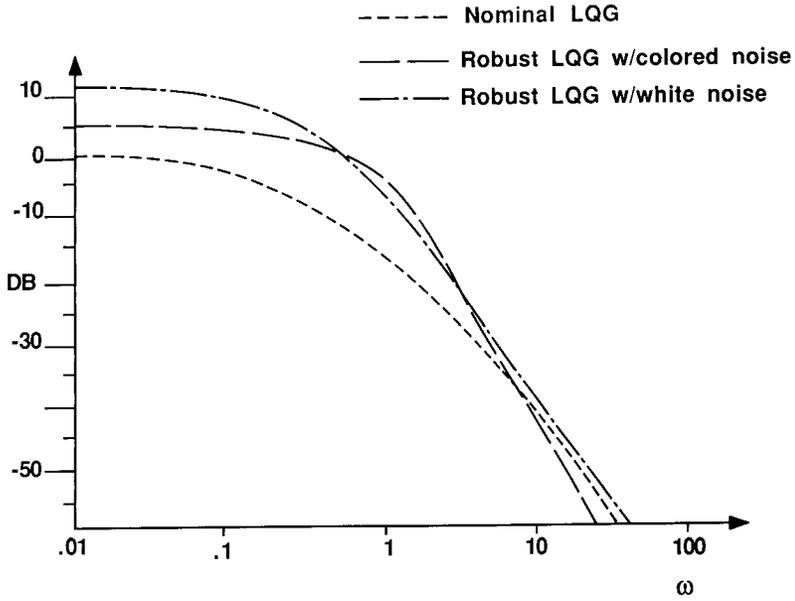
Figure 8.4-4   Magnitude of $t_1(j\omega)$.
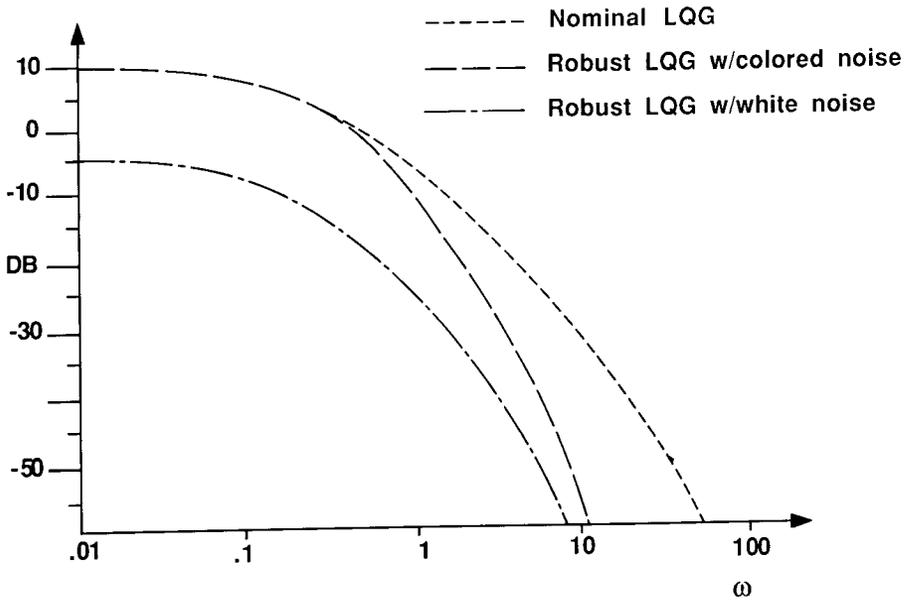
Figure 8.4-5   Magnitude of $t_2(j\omega)$.
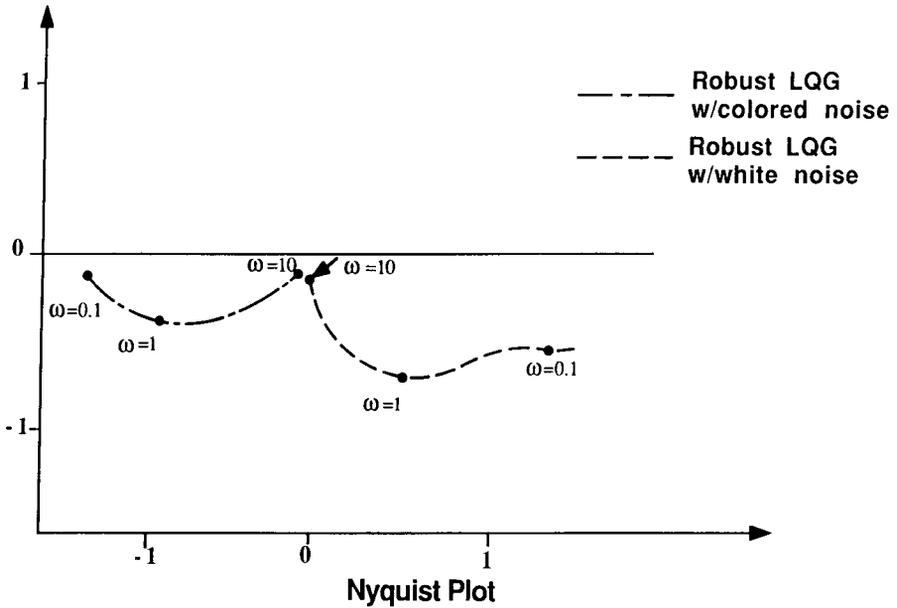
Figure 8.4-6   Control-loop frequency responses (compare with Fig. 8.3-6).

Discrete-time loop recovery.   The continuous-time techniques for loop recovery apply also to the discrete-time case. Two cases are of interest. The first is when calculation time is negligible compared to the discrete-time sampling interval, and the other is when there must be a unit lag in the feedback due to processing time. Asymptotic recovery properties for square minimum phase plants are explored in [13 and 14].

Main points of the section.   For minimum phase plants, loop recovery methods can be applied to achieve trade-offs between performance for a nominal plant and noise environment and input (or output) robustness in LQG designs. There is a scalar parameter that can be varied between zero and infinity to achieve a range of stabilizing controllers with attractive performance properties for a nominal plant when zero, and attractive input or output robustness properties associated with return difference inequalities of full state designs when infinity.

In the case of nonminimum phase plants, two approaches have been presented to achieve loop recovery results.

Working with filters that augment the plant, frequency-shaped designs allow different performance/robustness trade-offs in different frequency bands.

Problem 8.4-1.   Show a loop-recovery property in the presence of plant process noise invariant with respect to σ. In particular, show that (8.4-5) holds with $\hat{Q}_\sigma = \sigma GG' + \hat{Q}$ for any $\hat{Q} \geq 0$.

Problem 8.4-2.   Consider frequency-shaped plant process noise produced by white noise driving a (minimum phase) filter. Show that the full-order state estimator in this case can be viewed as that for the white noise case but with the filter gain K_e replaced by a filter $K_e(s)$, which reflects the frequency shaping of the process noise. Conclude heuristically that if $K_e(s)$ is thought of as a frequency-shaped gain, then loop recovery is frequency-shaped. [*Hint:* Work with an augmented plant consisting of the original plant augmented with the noise filter—see also [6] for loop recovery properties in this case.]

Problem 8.4-3.   (Requiring computer solution). Introduce various all-pass factors in series with the plant of the design example of this section, to see the effect of zeros on the Nyquist plots.

8.5 ROBUSTNESS IMPROVEMENT VIA RESIDUAL FEEDBACK

The loop recovery technique limits itself to working with state estimate feedback designs. Can we achieve robustness enhancement of an LQG design by relaxing the constraint of working only with state estimate feedback? In Sec. 8.1, it is pointed out that a key property of a state estimate feedback design as depicted in Fig. 8.1-1 is that the transfer function matrix from u_{ext} to plant output y is invariant with respect to aspects of the estimator design, being identical to that of a full-state regulator design. Can we modify the state estimate feedback controller in such a way as to preserve the transfer function matrix from u_{ext} to y and at the same time improve robustness properties?

To lead into answers to these questions we now discuss the class of all stabilizing one-degree-of-freedom controllers. This section is perhaps a bridge between the now "classical" LQ theory and current and possibly future trends in linear control theory. Consider a plant $P(s)$ and negative feedback controller $C(s)$. The control loop is well-posed and the controller $C(s)$ is said to stabilize $P(s)$ if and only if

$$\begin{bmatrix} I & C(s) \\ -P(s) & I \end{bmatrix}^{-1} \text{ exists and is stable and proper}$$

Equivalently, the four possible closed-loop transfer function matrices $(I + CP)^{-1}$, $P(I + CP)^{-1}$, $(I + CP)^{-1}C$, $P(I + CP)^{-1}C$ exist and are stable and proper. These conditions, of course, exclude open-loop unstable pole-zero cancellations. We seek the class of all proper $C(s)$ satisfying this stability property, that is, the class of all one-degree-of-freedom stabilizing controllers. To this end, let us begin by considering a particular basic stabilizing state estimate feedback controller of Sec. 8.1 (see Fig. 8.1-1), but modified to include a filter Q_f with transfer function matrix $Q_f(s)$ driven from the residuals ξ of the estimator, defined as $(y + r_2 - H'x_e)$ in the usual

notation, and with output q adding to the state estimate feedback $K'x_e$. The situation is depicted in Fig. 8.5-1(a). Of course, when the estimator is an optimal one for the noise processes r_1, r_2 then the residuals are precisely the innovations. So as to emphasize that the controller is driven from the plant output y and feeds to the plant input u, it is redepicted in Fig. 8.5-1(b) for the case $u_{ext} = 0$. [Fig. 8.5-1(b) also implicitly defines the transfer function matrix $J(s)$ used below]. The figure highlights the fact that the controller transfer function matrix is dependent on the filter transfer function matrix $Q_f(s)$.

Class of all stabilizing one-degree-of-freedom controllers. Consider the plant-controller arrangement of Fig. 8.5-1(a) and (b) with the estimator stable and $K'x$ a stabilizing state feedback law. Then the controller class characterized in terms of $Q_f(s)$, with $Q_f(s)$ stable and proper but otherwise arbitrary, is the entire class of stabilizing proper one-degree-of-freedom controllers for the plant. Moreover, transfer function matrices between external additive plant inputs r_1, r_2 and responses u, y are affine in $Q_f(s)$, and from u_{ext} to u and y are independent of $Q_f(s)$.

Close examination will show that the order of $J(s)$ together with $Q_f(s)$ is generically higher than the order of the plant, so that low order controllers are achieved by appropriate cancellations within the $J(s)$, $Q_f(s)$ arrangement.
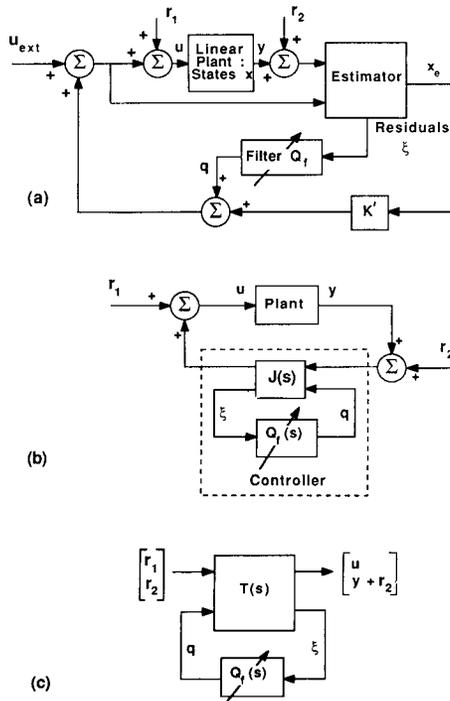
Figure 8.5-1   Class of all stabilizing controllers.

A proof of the above result is now given with certain details requested in Problem 8.5-1. This proof may be omitted by readers with no interest or background in matrix fraction descriptions. Let us denote the plant transfer function matrix as $P(s)$ and the negative feedback controller transfer function matrix in the case $Q_f(s) = 0$ as $C(s)$. Then there exist factorizations as

$$P(s) = B_R(s)A_R^{-1}(s) = A_L^{-1}(s)B_L(s) \tag{8.5-1a}$$

$$C(s) = Y_R(s)X_R^{-1}(s) = X_L^{-1}(s)Y_L(s) \tag{8.5-1b}$$

with stable proper factors

$$\begin{bmatrix} A_R(s) & -Y_R(s) \\ B_R(s) & X_R(s) \end{bmatrix} = \begin{bmatrix} K' \\ H' \end{bmatrix} (sI - F - GK')^{-1}[G \quad -K_e] + \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \tag{8.5-2a}$$

$$\begin{bmatrix} X_L(s) & Y_L(s) \\ -B_L(s) & A_L(s) \end{bmatrix} = \begin{bmatrix} K' \\ H' \end{bmatrix} (sI - F - K_eH')^{-1}[-G \quad K_e] + \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \tag{8.5-2b}$$

This may be checked by back substitution of (8.5-2) into (8.5-1); see also Appendix B. Now straightforward manipulations give the "double Bezout" equations

$$\begin{aligned} &\begin{bmatrix} A_R(s) & -Y_R(s) \\ B_R(s) & X_R(s) \end{bmatrix} \begin{bmatrix} X_L(s) & Y_L(s) \\ -B_L(s) & A_L(s) \end{bmatrix} \\ &= \begin{bmatrix} X_L(s) & Y_L(s) \\ -B_L(s) & A_L(s) \end{bmatrix} \begin{bmatrix} A_R(s) & -Y_R(s) \\ B_R(s) & X_R(s) \end{bmatrix} \\ &= \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \end{aligned} \tag{8.5-3}$$

which tells us that the factorizations (8.5-2) are coprime, [15]. This means that there are no unstable common zeros in the factors B_R, A_R, and so on of (8.5-1). Also, with reference to Fig. 8.5-1(b) and 8.5-1(c), straightforward manipulations not spelled out in detail here give

$$J(s) = \begin{bmatrix} J_{11}(s) & J_{12}(s) \\ J_{21}(s) & J_{22}(s) \end{bmatrix} = \begin{bmatrix} -C(s) & X_L^{-1}(s) \\ X_R^{-1}(s) & -X_R^{-1}(s)B_R(s) \end{bmatrix} \tag{8.5-4}$$

$$T(s) = \begin{bmatrix} T_{11}(s) & T_{12}(s) \\ T_{21}(s) & T_{22}(s) \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} I & C(s) \\ -P(s) & I \end{bmatrix}^{-1} \begin{bmatrix} A_R(s) \\ B_R(s) \end{bmatrix} \\ [B_L(s) \quad A_L(s)] \qquad \qquad 0 \end{bmatrix} \tag{8.5-5}$$

Now let us denote the transfer function matrix of the controllers of Fig. 8.5-1(b) as a function of Q_f as $C(Q_f, s)$ (note that $J_{11}(s) = -C(s) = -C(Q_f, s)|Q_f = 0$). Then manipulations involving the Bezout relations (8.5-3) give the Youla parameterizations [16]

$$\begin{aligned} C(Q_f, s) &= C(s) - X_L^{-1}(s)Q_f(s)[I - J_{22}(s)Q_f(s)]^{-1}X_R^{-1}(s) \\ &= Y_R(Q_f, s)X_R^{-1}(Q_f, s) \\ &= X_L^{-1}(Q_f, s)Y_L(Q_f, s) \end{aligned} \tag{8.5-6}$$

where

$$Y_R(Q_f, s) = Y_R(s) - A_R(s)Q_f(s) \tag{8.5-7a}$$

$$X_R(Q_f, s) = X_R(s) + B_R(s)Q_f(s) \tag{8.5-7b}$$

$$Y_L(Q_f, s) = Y_L(s) - Q_f(s)A_L(s) \tag{8.5-7c}$$

$$X_L(Q_f, s) = X_L(s) + Q_f(s)B_L(s) \tag{8.5-7d}$$

are stable and proper. It is readily shown that under (8.5-7), then the Bezout identity (8.5-3) holds with $Y_R(s)$ replaced by $Y_R(Q_f, s)$, and so on, so that the various inverses in (8.5-6) exist and the factorizations (8.5-6) are coprime. The mapping (8.5-6) is called a linear fractional map and it yields a unique $C(Q_f, s)$ for each Q_f; conversely, for each proper stabilizing controller $\overline{C}(s)$ for $P(s)$ there is a unique proper stabilizing $Q_f(s)$ such that $\overline{C}(s) = C(Q_f, s)$. A simple derivation follows below.

Also with reference to Fig. 8.5-1(c) and (8.5-5), manipulations give the transfer function from $[r_1' \quad r_2']'$ to $[u' \quad y' + r_2']$, as

$$\begin{bmatrix} I & C(Q_f, s) \\ -P(s) & I \end{bmatrix}^{-1} = T_{11}(s) + T_{12}(s)Q_f(s)T_{21}(s) \tag{8.5-8}$$

$$= \begin{bmatrix} I & C(s) \\ -P(s) & I \end{bmatrix}^{-1} + \begin{bmatrix} B_R(s) \\ A_R(s) \end{bmatrix} Q_f(s)[A_L(s) \quad B_L(s)]$$

Pre- and postmultiplication by $[Y_L(s) \quad X_L(s)]$ and $[X_R'(s) \quad Y_R'(s)]'$, respectively, and applying the Bezout identity (8.5-3) gives a unique solution for $Q_f(s)$ in terms of any proper stabilizing controller $\overline{C}(s) = C(Q_f, s)$, as

$$Q_f(s) = [Y_L(s) \quad X_L(s)]\left\{ \begin{bmatrix} I & \overline{C}(s) \\ -P(s) & I \end{bmatrix}^{-1} - \begin{bmatrix} I & C(s) \\ -P(s) & I \end{bmatrix}^{-1} \right\} \begin{bmatrix} X_R(s) \\ Y_R(s) \end{bmatrix} \tag{8.5-9}$$

Now from (8.5-8), $C(Q_f, s)$ is a stabilizing proper negative feedback controller for $P(s)$ when a proper $C(s)$ is a stabilizing negative feedback controller for $P(s)$, and when $Q_f(s)$ is proper and stable. Also from (8.5-9), if $\overline{C}(s)$ and $C(s)$ are proper stabilizing negative feedback controllers for $P(s)$, then $Q_f(s)$ such that $\overline{C}(s) = C(Q_f, s)$ is stable and proper. This completes the proof of the first claim.

Now from (8.5-8), it is clear that the transfer functions from r_1, r_2 to $u, y + r_2$ are affine in $Q_f(s)$. Also observe that since the state estimation error $(x - x_e)$ is known to be independent of u_{ext}, it follows immediately that the transfer function from u_{ext} to the residuals ξ is zero, and consequently the transfer functions from u_{ext} to u, y, x_e, and so on in the closed-loop system are independent of $Q_f(s)$.

This completes the proof of our claims concerning the class of all stabilizing controllers. Incidently, this constitutes a proof of the Appendix B Section 11, 12 results, viz. that the class of all stabilizing controllers is given by (8.5-6) and (8.5-7) in terms of any coprime factorizations satisfying (8.5-1) and (8.5-3). Here, the particular factorizations (8.5-2) are the relevant ones satisfying (8.5-1) and (8.5-3). Some remarks are now in order.

1. That $T_{22}(s)$ is zero can be seen directly from the fact that the transfer function matrix from u_{ext} to ξ is zero, independently of $Q_f(s)$. For $T_{22}(s)$ is the transfer function matrix from q to ξ; see Fig. 8.5-1(c). Referring to Fig. 8.5-1(a), we see this is identical with the transfer function matrix from u_{ext} to ξ. This property is crucial to achieving the linearity with respect to $Q_f(s)$ in the closed-loop transfer functions (8.5-8). This linearity assists any optimization of performance or loop robustness with respect to $Q_f(s)$ selections. Some details are given subsequently. The fact that $T_{22}(s)$ is zero also makes clear an eigenvalue separation property, viz. that with Q_f realized as a separate subsystem, the closed-loop modes consist of those of the state feedback regulator, the estimator, and those of Q_f.

2. We stress that a stabilizing regulator designed by any technique whatsoever can be viewed as a state estimate feedback controller designed with particular $Q, R, \hat{Q}, \hat{R}$ and with insertion of some filter $Q_f(s)$ between residuals ξ and the control u.

3. The introduction of the filter $Q_f(s)$ does not change the transfer functions from u_{ext} to u, y, these being identical to the optimal linear quadratic design. But the introduction of Q_f does change the loop gain properties and response to disturbances r_1, r_2. This allows us in principle to select $Q_f(s)$ to optimize such disturbance responses or loop gain properties without affecting the tracking capabilities of a linear quadratic design. (Recall that u_{ext} could be optimally designed to achieve a tracking objective.) Two approaches to Q_f selection are now studied. Also the frequency-shaping LQG design techniques of the next chapter can be viewed as methods to select Q_f.

4. The mapping from $Q_f(s)$ to $C(Q_f(s), s)$ is bijective as can be seen from manipulations on (8.5-6); see Problem 8.5-4.

5. A dual result can of course be obtained concerning the class of all proper rational plants stabilized by a single fixed negative feedback controller $C(s)$. Such a class can be parametrized in terms of one such plant $P(s)$ and an arbitrary proper stable $Q_p(s)$ as

$$P(Q_p(s), s) = P(s) + A_L^{-1}(s)Q_p(s)[I - A_L^{-1}(s)Y_R(s)Q_p(s)]^{-1}A_R^{-1}(s)$$

being a dual of the first equality in (8.5-6). The other duals of (8.5-6) are easily written down. Of course $P(0, s) = P(s)$, and the mapping from $Q_p(s)$ to $P(Q_p(s), s)$ is bijective, there being an obvious dual of the result of Problem 8.5-4.

Let us now think of $P(s)$ as a nominal plant for the design of a nominal negative feedback controller $C(s)$. Then any actual plant $P_a(s)$ can be organized in the form $P(Q_p(s), s)$ for some $Q_p(s)$, not necessarily stable [although if $C(s)$ were stabilizing for $P(Q_p(s), s)$, then $Q_p(s)$ would be stable]. Let us now apply a controller $C(Q_f(s), s)$, derived as described earlier, as a negative feedback controller to $P_a(s) = P(Q_p(s), s)$. It is interesting to ask when this control loop is stable. It turns

out (see Problem 8.5-5) that there is stability if and only if $Q_f(s)$ stabilizes $Q_p(s)$. Thus the problem of tuning a controller for stabilization is equivalent to the problem of tuning $Q_f(s)$ to stabilize $Q_p(s)$. Notice from Problem 8.5-4 and its dual that $Q_f(s)$ and $Q_p(s)$ are frequency-shaped versions of $[C(s) - C(Q_f(s), s)]$ and $[P(s) - P(Q_p(s), s)]$, respectively.

Sensitivity/loop recovery by Q_f selection.

As noted in the previous section, one technique for passband robustness improvement of an LQG design is loop recovery. Such must also be achievable by a suitable Q_f selection in view of the second remark above. Indeed, one can seek $Q_f(s)$ to minimize some measure of the difference between the open-loop gain matrices $K'(sI - F)^{-1}G$ obtained with full-state feedback and $-C(Q_f, s)P(s)$. The procedure we develop is termed *sensitivity recovery* rather than loop recovery for reasons which will become clear as we proceed, following [15].

Let us denote the maximum singular value of the sensitivity function difference over all frequency ω for $s = j\omega$ as

$$\Delta S_1 = \left\| [I - K'(j\omega I - F)^{-1}G]^{-1} - [I + C(Q_f, j\omega)P(j\omega)]^{-1} \right\|_\infty \quad (8.5\text{-}10)$$

We choose the index (8.5-10) for two reasons. First, the sensitivity function is affine in $Q_f(s)$ as

$$[I + C(Q_f, j\omega)P(j\omega)]^{-1} = A_R(j\omega)[X_L(s) + Q_f(s)B_L(s)]$$

Thus the difference term is affine in $Q_f(s)$ as follows:

$$\Delta S_1 = \left\| A_R(j\omega)[I - X_L(j\omega)] - A_R(j\omega)Q_f(j\omega)B_L(j\omega) \right\|_\infty$$

Minimizing this index is really a sensitivity recovery process. A second reason to work with this index in that it is an appropriate frequency-shaped loop recovery index as can be seen by its reorganization as

$$\Delta S_1 = \left\| A_R(j\omega)[K'(j\omega I - F)^{-1}G + C(Q_f, j\omega)P(j\omega)][I + C(Q_f, j\omega)P(j\omega)]^{-1} \right\|_\infty$$

The inner term $K'(j\omega I - F)^{-1}G + C(Q_f, j\omega)P(j\omega)$ is simply the error between the desired loop gain and the achieved loop gain. The first and third factors serve to frequency-weight this error. Since $A_R(j\omega) = [I - K'(j\omega I - F)^{-1}G]^{-1}$, this weight will be small when the loop gain is large, and thus tends to emphasize the unity gain cross-over frequency in the expression for ΔS_1. The same holds true of $[I + P(j\omega)C(Q_f, j\omega)]^{-1}$.

A dual index to (8.5-10) is

$$\Delta S_2 = \left\| [I + P(j\omega)C(j\omega)]^{-1} - [I + P(j\omega)C(Q_f, j\omega)]^{-1} \right\|_\infty$$
$$= \left\| [I - X_R(j\omega)]A_L(j\omega) - B_L(j\omega)Q_f(j\omega)A_L(j\omega) \right\|_\infty \quad (8.5\text{-}11)$$

Now when $P(s)$ is minimum phase, with full column rank in Re $s > 0$, then a left inverse $B_L^{-L}(s)$ exists and is stable, although not proper if $P(s)$ is strictly proper. In this case ΔS_1 is optimized with a selection

$$Q_f(s) = [I - X_L(s)]B_L^{-L}(s) \quad (8.5\text{-}12)$$

which is stable but not necessarily proper, (The issue of nonproperness is discussed subsequently). The optimum index is $\Delta S_1^* = 0$. Likewise in the dual situation when $B_R^{-R}(s)$ exists and is stable, then ΔS_2 is optimized with

$$Q_f(s) = B_R^{-R}(s)[I - X_R(s)] \qquad (8.5\text{-}13)$$

It is readily shown that $Q_f(s)$ in (8.5-12) or (8.5-13) has poles which are the set (or a subset) of the zeros of B_L^{-L} or B_R^{-R}, being the zeros of the plant.

When the plant $P(s)$ is nonminimum phase, an optimization procedure known as H^∞ optimization can be used to find the optimum Q_f [17]. Details are omitted here.

The optimum stable Q_f will not be proper for plants of relative degree $l = 2$ or more, although $s^{-(l-1)}Q_f(s)$ will be proper. In practice then the optimum $Q_f(s)$, if improper, must be approximated by a proper transfer function. This can be achieved by approximating s by $\alpha s / (\alpha + s)$ for suitably large $\alpha > 0$.

When we introduce nonzero Q_f to an LQG optimally designed regulator, we can worsen the closed-loop performance, while improving robustness. Trade-off between performance and robustness in a sensitivity recovery operation can be achieved by using a filter Q_f from (8.5-12) or (8.5-13) with a gain m that can range from zero to unity. The controllers will be stabilizing for all $m \geq 0$ since $[mQ_f(s)]$ is stable. With $m = 0$ the original design can be recovered, and with $m = 1$ full sensitivity and loop recovery can be achieved, so that a suitable trade-off can be made with a selection of $0 \leq m \leq 1$. Design examples showing the strength of this sensitivity recovery approach are given in [17], one of which is included below. The relative attractiveness of the sensitivity recovery approach over the recovery approach of the previous sections is not surprising, since more parameters are used in the design and there is a natural in-built frequency-shaping to the loop recovery in sensitivity recovery.

Frequency-shaped sensitivity recovery methods can be readily devised. In such cases, as indeed for the Q_f selections above, the controller order may be excessive, so that controller reduction via the techniques in Chapter 10 could well be in order for a final design.

A design example.   Consider the nonminimum phase plant

$$P(s) = \frac{s-1}{s^2 - 3s + 3} = \begin{bmatrix} 1 & -1 \end{bmatrix} \left(sI - \begin{bmatrix} 3 & -3 \\ 1 & 0 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

with nonminimum phase zero at 1 and unstable poles at $(1.5 \pm j\sqrt{3}/2)$. Here $R = 1$, $Q = [1 \quad 1]'[1 \quad 1]$, $\hat{R} = 1$, $\hat{Q} = [1 \quad 0]'[1 \quad 0]$, and the controller/estimator LQG gains

$$K' = \begin{bmatrix} -6.21 & -0.16 \end{bmatrix} \quad \text{and} \quad K_e' = \begin{bmatrix} 0.16 & 6.53 \end{bmatrix}$$

Consider now minimization of the sensitivity index ΔS_1 over stable proper Q_f. The H^∞ optimization algorithm of [17] gives an optimal nonproper Q_f as $Q_f = 6.05(s + 3.5)$, which we approximate as

$$Q_a = \frac{6.05(s + 3.53)}{(0.05s + 1)}$$

In this nonminimum phase plant case, only partial loop and sensitivity recovery can be achieved. Figure 8.5-2 shows the Nyquist plots for the various open loop transfer functions. Curve 1 shows the plot of the open-loop transfer function for the nominal LQG implementation with K, K_e given as above. When the loop recovery technique of Sec. 8.4 is used, with fictitious noise intensity 10,000, Curve 2 is obtained. There is a marginal improvement over the nominal case. With the fictitious noise intensity increased further, no improvement is recorded. In fact, there is a degradation over certain frequency ranges.

Using the sensitivity recovery technique with Q_a as above, we obtain Curve 3 of Fig. 8.5-2. Clearly there is an improvement over the nonfrequency-shaped loop technique as far as robustness in the critical frequency band is concerned. The improvement in the critical frequency band is obtained at the expense of some degradation in the noncritical frequency band. This demonstrates clearly the effectiveness of the in-built frequency weightings of the sensitivity recovery to achieve the desired loop properties.

Other robustness optimizations.   We mention in passing that $Q_f(s)$ can be selected with robustness measures other than loop recovery in mind. For example, one can select stable proper $Q_f(s)$ to minimize the sensitivity function-based index

$$\left\| [I + C(Q_f, s)P(s)]^{-1} \right\|_\infty = \left\| A_R(j\omega)[X_L(j\omega) + Q_f(j\omega)B_R(s)] \right\|_\infty$$
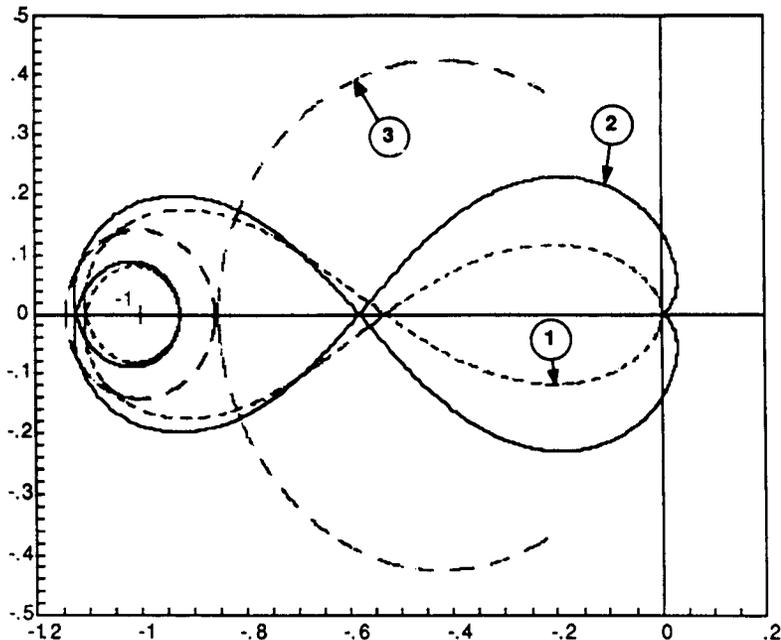
Figure 8.5-2   Nyquist plots for sensitivity/loop recovery.

which has $Q_f(s)$ entering in an affine manner. The optimization of $Q_f(s)$ is again an H^∞ optimization task [18]. In a like manner duals or complementary sensitivity functions can be optimized. Further details are omitted here.

All stabilizing two-degrees-of-freedom controllers.

In dealing with external inputs, as in tracking, controllers are driven by both the plant output and a reference signal giving rise then to two-degrees-of-freedom controllers. This motivates us to conveniently characterize the class of all stabilizing proper two-degrees-of-freedom controllers. A key nontrivial observation is that the class of all two-degrees-of-freedom controllers for a plant $P(s)$ can be viewed as the class of all one-degree-of-freedom controllers for the augmented plant $[0 \quad P'(s)]'$. Some derivations for this are requested in Problem 8.5-2—see also [19]. A special subset of the class of all two-degrees-of-freedom controllers is the class of model matching controllers that achieves a prescribed input-output transfer function, perhaps that of a nominal optimal design. Results for this are studied in [20]. Again the classes of controllers of interest can be constructed by means of LQG designs with added arbitrary stable proper filters $Q_f(s)$.

Main points of the section.
State estimate feedback arrangements with additional arbitrary stable proper residual feedback filters $Q_f(s)$ can be used to generate the class of all stabilizing proper controllers. With differing $Q_f(s)$ selections, differing robustness properties can be achieved, including sensitivity and loop recovery, while maintaining closed-loop transfer functions.

Problem 8.5-1.   Verify (8.5-1) to (8.5-8) by back substitution.

Problem 8.5-2.   Consider the factorizations for the augmented plant $[0 \quad P'(s)]'$ and stabilizing controller $C(s) = [C_1(s) \quad C_2(s)]$, in the notation of Sec. 8.1.

$$\bar{P}(s) = \begin{bmatrix} 0 \\ P(s) \end{bmatrix} = \bar{B}_R(s)\bar{A}_R^{-1}(s) = \bar{A}_L^{-1}(s)\bar{B}_L(s)$$

$$C(s) = -\bar{Y}_R(s)\bar{X}_R^{-1}(s) = -\bar{X}_L^{-1}(s)\bar{Y}_L(s)$$

with

$$\bar{B}_R(s) = \begin{bmatrix} 0 \\ B_R(s) \end{bmatrix}, \qquad \bar{B}_L(s) = \begin{bmatrix} 0 \\ B_L(s) \end{bmatrix}, \qquad \bar{A}_R(s) = A_R(s), \qquad \bar{A}_L(s) = \begin{bmatrix} I & 0 \\ 0 & A_L(s) \end{bmatrix}$$

$$\bar{Y}_R(s) = [-A_R(s)X_L(s)C_1(s) \quad Y_R(s)]$$

$$\bar{X}_R(s) = \begin{bmatrix} I & 0 \\ B_R(s)X_L(s)C_1(s) & X_R(s) \end{bmatrix}$$

$$\bar{Y}_L(s) = [-X_L(s)C_1(s) \quad Y_L(s)], \qquad \bar{X}_L(s) = X_L(s)$$

Write down the class of all one-degree-of-freedom controllers for $\bar{P}(s)$ using (8.5-6) and (8.5-7). Then reinterpret to give formulations for the class of all two-degrees-of-freedom stabilizing controllers for $P(s)$, using these factorizations.

Problem 8.5-3

(a) Consider a plant $P_r(s) = (sI - F)^{-1}G$ with a stabilizing (positive feedback) controller K'; then show that factorizations exist as

$$\begin{bmatrix} A_{rR}(s) \\ B_{rR}(s) \end{bmatrix} = \begin{bmatrix} K' \\ I \end{bmatrix}(sI - F - GK')^{-1}G + \begin{bmatrix} I \\ 0 \end{bmatrix}$$

$$[B_{rL}(s) \quad A_{rL}(s)] = (sI - F - GK')^{-1}[G \quad GK'] + [0 \quad I]$$

$$X_{rL}(s) = X_{rR}(s) = I, \qquad Y_{rL}(s) = Y_{rR}(s) = -K'$$

which satisfy the double Bezout equation

$$\begin{bmatrix} X_{rL}(s) & Y_{rL}(s) \\ -B_{rL}(s) & A_{rL}(s) \end{bmatrix}\begin{bmatrix} A_{rR}(s) & -Y_{rR}(s) \\ B_{rR}(s) & X_{rR}(s) \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$$

(b) Propose duals for K_e stabilizing $P_e(s) = H'(sI - F)^{-1}$.

(c) Generate the class of all stabilizing controllers for $P_r(s) = (sI - F)^{-1}G$ in terms of arbitrary stable proper $Q_f(s)$ using the coprime factorizations of (a) and (8.5-6), (8.5-7).

Problem 8.5-4.   Derive the relationship from (8.5-6)

$$Q_f = X_L(Q_f)[C - C(Q_f)]X_R$$
$$= X_L[C - C(Q_f)]X_R(Q)$$

Suppose C is strictly proper. Then show that $C(Q_f)$ is strictly proper if and only if Q_f is strictly proper.

Problem 8.5-5.   Derive the expression

$$\begin{bmatrix} I & C(Q_f) \\ -P(Q_p) & I \end{bmatrix}^{-1} = \begin{bmatrix} I & C \\ -P & I \end{bmatrix}^{-1} + \begin{bmatrix} A_R & Y_R \\ B_R & X_R \end{bmatrix}\left\{\begin{bmatrix} I & -Q_f \\ -Q_p & I \end{bmatrix}^{-1} - I\right\}\begin{bmatrix} X_L & Y_L \\ B_L & A_L \end{bmatrix}$$

and conclude that $C(Q_f)$ as a negative feedback controller stabilizes the plant $P(Q_p)$ if and only if Q_f stabilizes Q_p.

REFERENCES

[1] J. B. Moore, "A Note on Feedback Compensators in Optimal Linear Systems," *IEEE Trans. Auto. Control*, Vol. AC-15, No. 4 (August 1970), pp. 494–495.

[2] J. B. Moore and G. Ledwich, "Minimal Order Observers for Estimating Linear Func-

tions of a State Vector," *IEEE Trans. Auto. Control*, Vol. AC-20, No. 5 (October 1975), pp. 623–631.

[3] D. C. Youla, J. J. Bongiorno, Jr. and C. N. Lu, "Single-loop Feedback Stabilization of Linear Multivariable Plants," *Automatica*, Vol. 10 (1974), pp. 159–173.

[4] M. H. A. Davis, *Linear Estimation and Stochastic Control*. London: Chapman and Hall, 1977.

[5] J. C. Doyle, "Guaranteed Margins in LQG Regulators," *IEEE Trans. Auto. Control*, Vol. AC-23, No. 4 (August 1978), pp. 664–665.

[6] J. B. Moore, D. Gangsaas, and J. Blight, "Performance and Robustness Trades in LQG Regulator Designs," *Proc. 20th IEEE Conf. on Dec. and Contr.*, San Diego, Cal., December 1981, pp. 1191–1199.

[7] J. C. Doyle and G. Stein, "Robustness with Observers," *IEEE Trans. Auto. Control*, Vol. AC-24, No. 4 (August 1979), pp. 607–611.

[8] J. C. Doyle and G. Stein, "Multivariable Feedback Design: Concepts for a Classical/ Modern Synthesis," *IEEE Trans. Auto. Control*, Vol. AC-26, No. 1 (February 1981), pp. 4–16.

[9] Z. Zhang and J. S. Freudenberg, "Loop Transfer Recovery with Nonminimum Phase Zeros," *Proc. 26th IEEE Conf. on Dec. and Contr.*, Los Angeles, December 1987, pp. 956–957.

[10] J. B. Moore and Lige Xia, "Loop Recovery and Robust State Estimate Feedback Designs," *IEEE Trans. Auto. Control*, Vol. AC-36, No. 6 (June 1987), pp. 512–517.

[11] C. Chu and J. Doyle, "On Inner-Outer Spectral Factorizations," *Proc. 23rd IEEE Conference Decision and Control*, Las Vegas, Nev., December 1984, pp. 1764–1765.

[12] G. Stein and M. Athans, "The LQG/LTR Procedure for Multivariable Feedback Control Design," *IEEE Trans. Auto. Control*, Vol. AC-32, No. 2 (February 1987), pp. 105–114.

[13] J. M. Maciejowski, "Asymptotic Recovery for Discrete-Time Systems," *IEEE Trans. on Auto. Control*, Vol. AC-30, No. 6 (June 1985), pp. 602–605.

[14] T. Ishihara and H. Takeda, "Loop Transfer Recovery Techniques for Discrete-Time Optimal Regulators Using Prediction Estimators," *IEEE Trans. on Auto. Control*, Vol. AC-31 (December 1986), pp. 1149–1151.

[15] C. N. Nett, C. A. Jacobson, and M. J. Balas, "A Connection between State Space and Doubly Coprime Fractional Representations," *IEEE Trans. Auto. Control*, Vol. AC-29, No. 9 (September 1984), pp. 831–832.

[16] D. C. Youla, J. J. Bongiorno, Jr. and H. A. Jabr, "Modern Wiener-Hopf Design of Optimal Controllers Parts I, II," *IEEE Trans. on Auto. Control*, Vol. AC-21, No. 1 (1976), pp. 3–14, and Vol. AC-21, No. 6 (June 1976), pp. 319–330.

[17] J. B. Moore and T. T. Tay, "Loop Recovery via H^∞ Sensitivity Recovery," *Int. J. Control*, to appear.

[18] B. A. Francis, *A Course in H^∞ Control Theory*. Berlin: Springer Verlag, 1987.

[19] T. T. Tay and J. B. Moore, "Performance Enhancement of Two-degree-of-freedom Controllers via Adaptive Techniques," *Int. J. Adap. Control and Signal Processing*, to appear.

[20] J. B. Moore, Lige Xia, and K. Glover, "On Improving Control-loop Robustness of Model Matching Controllers," *Systems and Control Letters*, Vol. 7 (1986), pp. 83–87.

9

Frequency Shaping

9.1 BLENDING CLASSICAL AND LINEAR QUADRATIC METHODS

In the development of linear quadratic methods for control system design so far, we have been able to assess the quality of the designs from a robustness point of view in terms of classical frequency domain insights. We have focused on their strength in achieving attractive gain margins of $(\frac{1}{2}, \infty)$, guaranteed 60-deg phase margins, or multivariable equivalents. We have suggested that in the adjustment of design parameters, bandwidth considerations should be taken into account, as in classical designs.

In this chapter we go further in blending classical control design insights and techniques with linear quadratic methods so that each is used at its point of strength. The objective is to achieve in a systematic manner designs that are as good as, or better, than those which can be achieved by purely classical techniques where these apply, and yet extend to situations intractable to classical design methods; we seek also to improve on the standard linear quadratic-based designs. The new designs will be termed *frequency-shaped designs*. Basic results for such designs are developed in [1–4].

Linear quadratic methods are in essence time-domain methods, and yet they result in designs with attractive frequency domain properties associated with the return difference inequality. There are also useful interpretations of asymptotic

properties as weight gains increase with transmission zeros playing their expected role.

Classical control methods are executed in the frequency domain, with compensators designed to appropriately frequency-shape the return difference, or equivalently the open-loop gain. There is an emphasis on setting bandwidths for effective control action. There is also the concept of closed-loop system poles approaching open-loop zeros as the loop gain increases, suggesting that the designer may introduce suitably located zeros via series compensators. Such concepts may be implemented by using proportional plus integral PI compensators or proportional plus integral plus differential PID compensators. These are almost universal in classical designs, and allow good disturbance rejection, including asymptotic rejection of constant disturbances.

How then can linear quadratic methods with their computational elegance and power harness more fully the classical notions of frequency shaping, transmission zero adjustment, and PID compensation?

In the previous chapter, Sec. 8.2, the notion of studying stochastic system performance of an LQG design in the frequency domain is introduced. Also, the use of such insights to modify the quadratic performance index to yield an improved design is illustrated by a design example. The resulting controllers are primitive frequency-shaped designs. The frequency-shaping approach can be carried a step further by augmenting the plant with frequency-shaped filters so as to penalize their outputs in addition to other cost terms in the performance index. For example, if the performance response spectrum shows too much control (or output) energy in a certain frequency band, then the plant can be augmented by a filter driven from the control (or output) with a response only in this passband. Penalizing the output of such a filter in the performance index associated with the augmented plant may well improve the performance response—or at least allow trade-offs to be made. To illustrate this simple approach to achieving a frequency-shaped design, we return to the design example of Sec. 8.2.

Yaw damper frequency-shaped design.
Recall from Sec. 8.2 that the open-loop spectral response of a low-order model of an aircraft subject to wind gusts is as in Fig. 8.2-2, showing resonances at 1.5 and 22 rad/s. A simple LQG design with the index $E[y_f^2 + y_a^2 + 0.2u^2]$ leads to the closed-loop response of Fig. 8.2-3. There is a considerable reduction in the resonances (factors of two or three). To achieve further improvement in performance, it makes sense to work with a modified index, further penalizing the aft response y_a at less than 3 rad/s. Thus, let us filter y_a to achieve a filtered response y_a^f as

$$y_a^f(s) = W^f(s)\, y_a(s), \qquad W^f(s) = \frac{9}{s^2 + 2s + 9} \tag{9.1-1}$$

and work with the new performance index

$$V = E[y_a^2 + y_f^2 + 8(y_a^f)^2 + 0.1u^2] \tag{9.1-2}$$

This index is, of course, a standard quadratic index for the plant augmented with the frequency-shaping filter at the plant output. This augmented plant has a state-variable representation of the form

$$\begin{bmatrix} \dot{x} \\ \dot{x}^f \end{bmatrix} = \begin{bmatrix} F & 0 \\ G^f H' & F^f \end{bmatrix} \begin{bmatrix} x \\ x^f \end{bmatrix} + \begin{bmatrix} G \\ 0 \end{bmatrix} u \qquad (9.1\text{-}3)$$

where the original plant is described by $\dot{x} = Fx + Gu$, $y_a = H'x$, and the filter is described by $\dot{x}^f = F^f x^f + G^f u^f$, $y_a^f = H^{f'} x^f$. Of course, $u^f = y_a$, and $H^{f'}(sI - F^f)^{-1} Gf = W^f(s)$.

The augmented plant state vector entries consist of those of the original plant x, together with those of the filter x^f. The state feedback law for the augmented plant is obtained by standard methods to yield a feedback law $u = K'x + K_f'x^f$. It is easily shown that the augmented plant feedback controller is equivalent to the dynamic state feedback law for the original plant as

$$u(s) = K'(s)x(s), \qquad K'(s) = K' + K_f'(sI - F^f)^{-1}G^f H' \qquad (9.1\text{-}4)$$

In an LQG design, there is a corresponding dynamic state estimate feedback $u(s) = K'(s)x_e(s)$.

For the design problem of Sec. 8.2, it turns out that the frequency-shaped design gives a significant improvement over the nonfrequency-shaped design. This is depicted in Fig. 9.1-1. Notice also the improvement at less than 1 rad/s and mild improvement elsewhere over the primitive frequency-shaped design of Fig. 8.2-4.
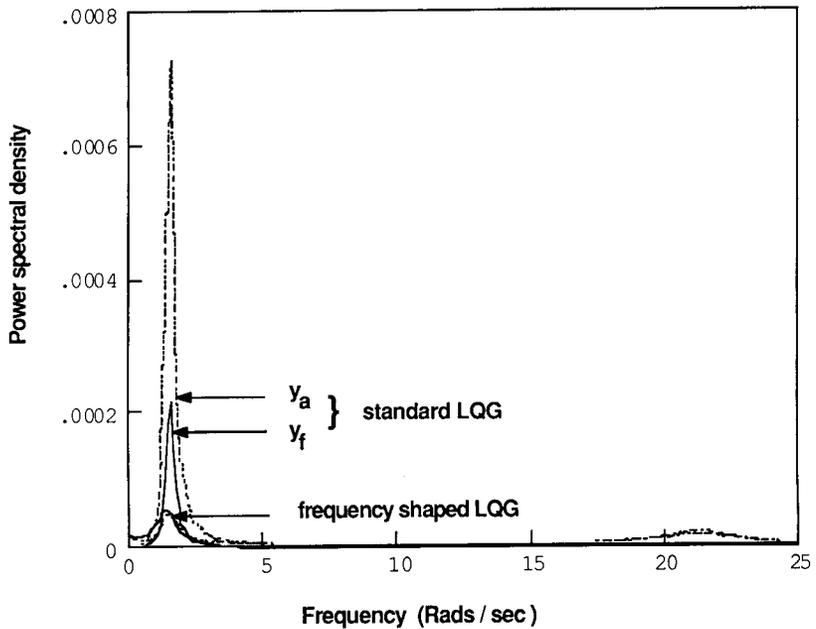
Figure 9.1-1 Frequency weighted LQG design.

Of course, the improvement is gained at the cost of controller complexity increase (by an order increase from 6 to 8).

Notice that were we dealing with a deterministic index and had closed-loop stability, the cost term penalizing y_a^f could be written in the frequency domain as

$$\int_0^\infty (y_a^f)^2 \, dt = \frac{1}{2\pi} \int_{-\infty}^\infty y_a^f(-j\omega) \, y_a^f(j\omega) \, d\omega$$

$$= \frac{1}{2\pi} \int_{-\infty}^\infty y_a(-j\omega) \, W^f(-j\omega) \, W^f(j\omega) \, y_a \, (j\omega) \, d\omega \qquad (9.1\text{-}5)$$

Clearly, the constant weight on y_a^f can be interpreted as a frequency-shaped weighting on y_a.

The concept of augmenting a plant with filters to lead to frequency-shaped designs applies to designs with only deterministic objectives as well as those with stochastic objectives, such as studied in the above design. A number of possible augmentation arrangements are depicted in Fig. 9.1-2.

In the remainder of this section, we further discuss the augmented plant approach to frequency-shaping, appealing at times to examples studied in earlier chapters, and then summarize the work of the next sections.

Figure 9.1-2(a) shows the addition of a filter on the control signal leading to an augmented plant having the original plant input u and output y, but with state
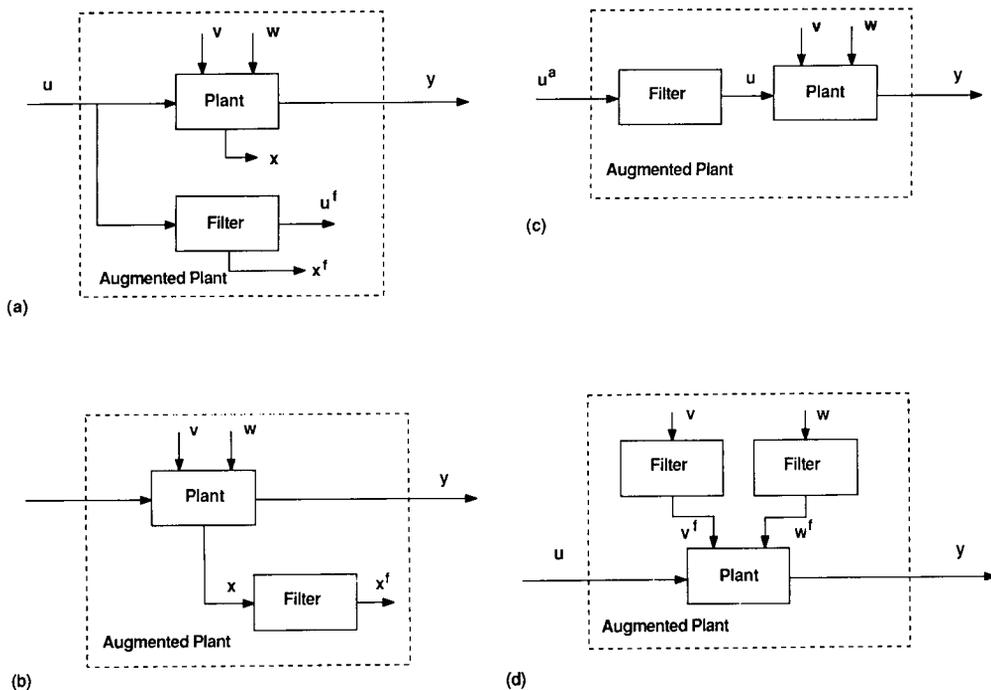
Figure 9.1-2 Augmented plants.

variables consisting of those of the filter, denoted x^f, as well as those of the original plant. Thus a performance index of the form

$$V = E[x'Qx + u'Ru + x^{f'}Q^f x^f] \qquad (9.1\text{-}6)$$

can be considered, for example. Of course, it may well be that $u^f = H^{f'}x^f$ for some H^f and $Q^f = H^f H^{f'}$, so that (9.1-6) in effect is penalizing u^f. Figure 9.1-2(b) shows the addition of a filter, with states x^f, driven from the plant states.

Augmentations that lead to a plant with a modified input are shown in Fig. 9.1-2(c), and those that modify process and/or sensor noise are illustrated in Fig. 9.1-2(d). Of course, combinations of filters can be employed if desired.

Design examples using this augmented plant approach have been studied earlier in the text. For example, in the previous chapter, Sec. 8.4, there is introduced fictitious frequency-shaped plant process noise v^f to achieve a frequency-shaped loop recovery in a design example. There results a state estimator for the original plant with a frequency-shaped filter gain $K_e(s)$ instead of the usual constant gain K_e. Another case studied earlier is in Chapter 4, Problem 4.3-5, where a plant is augmented with an integrator at its input. This means that the augmented plant input u^a is $\dot{u}$, the derivative of the original plant input. Thus in penalizing the augmented plant input by the quadratic term $u^{a'}R_a u^a$, there is a penalty on the rate of change of u, namely $\dot{u}'R_a\dot{u}$. There results a dynamic state feedback controller $K'(s)$ for the original plant rather than a constant state feedback gain K' as in nonfrequency-shaped designs. The disadvantage of this particular augmentation is that the return difference inequality is satisfied at the augmented plant input, and not the original plant input. Thus input robustness properties usually associated with an LQ design could well be lost.

The use of a frequency-dependent weighting that emphasizes high frequencies in the penalty on control effort is studied in [5]. The weighting is motivated by the assumption that the plant model may not be accurate at high frequencies, so that control activity should be attenuated at high frequency. The conclusion from analytical studies in [5] is that such frequency weighting will improve robustness of a state feedback design to unstructured uncertainty outside the passband, relative to the nonfrequency-dependent weighting situations, but in the process some robustness may be lost in the passband, with gain and phase margins being reduced. We do not explore input augmentations for frequency-shaped designs further in this chapter.

The frequency shaping based on augmented plants discussed above can actually be interpreted as working with the original plant and a frequency-shaped performance index. Thus the cost term $u^{a'}R_a u^a = \dot{u}'R_a\dot{u}$ is really a time domain version of $u'(-j\omega)(\omega^2 R_a)u(j\omega)$. What has happened is that instead of working with a constant penalty matrix R, there is a frequency-shaped penalty matrix $R(j\omega) = \omega^2 R_a$. All the frequency-shaping concepts based on plant augmentations of Fig. 9.1-2 can be likewise interpreted as generalizing the original performance index matrices for regulation and state estimation, as

$$R(s), \, Q(s), \, \hat{R}(s), \, \hat{Q}(s) \quad (\text{with } s = j\omega)$$

Moreover, the consequence of such generalizations is to replace the constant gains K', K_e by frequency-shaped gains

$$K'(s), K_e(s)$$

Problem 8.4-2 develops this interpretation for the case of state estimation when the constant matrices $\hat{R}$, $\hat{Q}$, K_e for the constant R case generalize to $\hat{R}(s)$, $\hat{Q}(s)$, (with $s = j\omega$) and $K_e(s)$. Also recall that Problem 4.3-5 in effect illustrates the same concept for the state feedback regulator case when R, Q, K' generalize to $\omega^2 R$, Q, $K'(s)$.

In Section 2 of this chapter, we view the class of all stabilizing strictly proper controllers for a nominal plant as frequency-shaped state estimate feedback arrangements with the frequency shaping in the state estimator gain K_e, or the state feedback gain K, or both. That is, we permit a generalization of the gain, K, to a transfer function $K(s)$ or the gain K_e to $K_e(s)$, or both K, K_e to $K(s)$, $K_e(s)$. As pointed out, such designs with dynamic gains $K(s)$ and/or $K_e(s)$ can be achieved by using the linear quadratic approach where there is a generalization of the performance index weights Q, R and noise intensities $\hat{Q}$, $\hat{R}$ to transfer functions $Q(s)$, $R(s)$ and $\hat{Q}(s)$, $\hat{R}(s)$. Thus frequency-shaped penalties are applied to control signal and state costs, and in the state estimation.

In Sections 3 and 4, we show how to achieve in a systematic manner practical linear-quadratic-based design exploiting classical ideas, focusing on the case when K, K_e become proportional plus integral filters. The properties of such designs are studied, including their rejection of constant disturbances in both state estimation and control. The approach can be generalized for rejection of ramp or periodic disturbance inputs.

Of course, the frequency-shaped LQG designs have an increased complexity over standard designs, but controller reduction techniques set out in Chapter 10 offer the possibility of taking care of this problem in a systematic manner. There is also increased designer effort in the approach of this chapter. With the increase in flexibility and performance/robustness trade-offs there is now not only the task of selecting performance index weights, but the frequency shaping. The challenge addressed in this chapter is to present one or two possible paths for the designer to follow in making such decisions—with no superiority claims over other approaches.

Main points of the section.   Although the linear quadratic (LQ) design approach is in the time domain, aspects of performance can be viewed in the frequency domain. Moreover, cost terms in a performance index can be penalized in the frequency domain by appropriate augmentation of the plant with frequency-shaping filters. Standard LQ designs for the augmented plants can be interpreted in terms of a frequency-shaped design for the original plant.

Problem 9.1-1.   (Requiring computer). For the yaw damper design of this section, try first-order frequency-shaping filters $W^f(s)$ to improve resonance suppression.

9.2 STATE ESTIMATE FEEDBACK WITH FREQUENCY SHAPING

In designing a controller associated with a nominal plant, the first requirement is that the controller stabilize the nominal plant and thus belongs to the class of all stabilizing controllers for the nominal plant. Of course, state estimate feedback LQG designs belong to a rather restricted subset of this class, and it is not surprising that improved robustness properties can be achieved by designs outside the set of standard LQG designs.

In this section, we reexamine the class of all stabilizing strictly proper controllers for a nominal plant as studied in Chapter 8, Section 5. We show that one can view this class as feedback of state estimates from a fixed but arbitrary standard state estimator via dynamic filters $K'(s)$, rather than via constant gains K'. The class of filters $K'(s)$ is the class of all stabilizing proper state feedback filters for the plant. Alternatively, the class of all proper stabilizing controllers may be viewed as coming via a fixed but arbitrary stabilizing constant gain state estimate feedback, where the state estimator has a stabilizing proper dynamic gain $K_e(s)$ rather than a constant one. These results motivate and indeed justify our study in the next sections of linear quadratic methods to design suitable $K(s)$ and/or $K_e(s)$. We shall also examine benefits that can accrue from such designs. In the remainder of this section, we make precise and prove the above statements.

The results we develop are specializations of those presented in [6]. We suggest that the reader having difficulty with factorization theory move on to the next section.

Let us work with the notation of previous chapters and assume a plant with transfer function matrix

$$P(s) = H'(sI - F)^{-1}G \qquad (9.2\text{-}1)$$

For regulator and estimator designs, we work also with components of this plant, namely,

$$P_r(s) = (sI - F)^{-1}G, \qquad P_e(s) = H'(sI - F)^{-1} \qquad (9.2\text{-}2)$$

Observe that stabilizing negative feedback controllers for $P_r(s)$, $P_e(s)$ are provided by stabilizing state feedback regulator gains $(-K')$, and stabilizing filter gains $(-K_e)$, respectively. Now, as foreshadowed in Problem 8.5-3, the class of all stabilizing controllers for $P_r(s)$, $P_e(s)$ can be conveniently parametrized in terms of arbitrary stable proper $Q_r(s)$, $Q_e(s)$, respectively, as follows.

Define factorizations as

$$P_r(s) = B_{rR}(s) A_{rR}^{-1}(s) = A_{rL}^{-1}(s) B_{rL}(s) \qquad (9.2\text{-}3a)$$

$$(-K') = Y_{rR}(s) X_{rR}^{-1}(s) = X_{rL}^{-1}(s) Y_{rL}(s) \qquad (9.2\text{-}3b)$$

with

$$\begin{bmatrix} A_{rR}(s) \\ B_{rR}(s) \end{bmatrix} = \begin{bmatrix} K' \\ I \end{bmatrix} (sI - F - GK')^{-1}G + \begin{bmatrix} I \\ 0 \end{bmatrix} \qquad (9.2\text{-}4a)$$

$$[B_{rL}(s) \quad A_{rL}(s)] = (sI - F - GK')^{-1}[G \quad GK'] + [0 \quad I] \qquad (9.2\text{-}4b)$$

$$X_{rL}(s) = X_{rR}(s) = I, \qquad Y_{rL}(s) = Y_{rR}(s) = (-K') \qquad (9.2\text{-}4c)$$

These are stable and proper, and indeed coprime, since they are readily seen to satisfy the double Bezout identity

$$\begin{bmatrix} X_{rL}(s) & Y_{rL}(s) \\ -B_{rL}(s) & A_{rL}(s) \end{bmatrix} \begin{bmatrix} A_{rR}(s) & -Y_{rR}(s) \\ B_{rR}(s) & X_{rR}(s) \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \qquad (9.2\text{-}5)$$

We now appeal to Section 8.5. The class of all stabilizing controllers for $P_r(s)$ is depicted in Fig. 9.2-1 in terms of an arbitrary stable proper $Q_r(s)$ and

$$J_r(s) = \begin{bmatrix} K' & I \\ I & -B_{rR}(s) \end{bmatrix} \qquad (9.2\text{-}6)$$

(a) $Q_r(s)$ arbitrary proper stable

(b) $Q_e(s)$ arbitrary proper stable
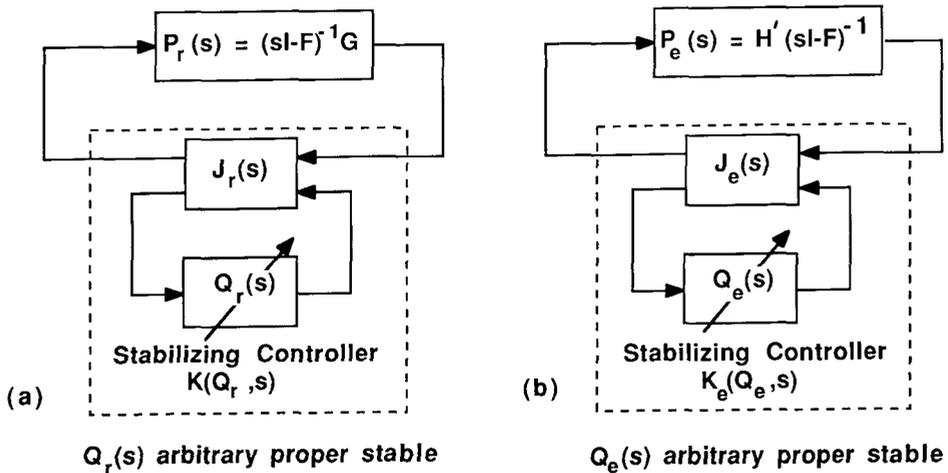
Figure 9.2-1   Stabilizing controllers $K(Q_r, s)$, $K_e(Q_e, s)$.

The dual for $P_e(s)$ is also depicted in terms of arbitrary stable proper $Q_e(s)$ and

$$J_e(s) = \begin{bmatrix} K_e & I \\ I & -B_{eR}(s) \end{bmatrix} \qquad (9.2\text{-}7a)$$

$$B_{eR}(s) = H'(sI - F - K_eH')^{-1} \qquad (9.2\text{-}7b)$$

The verification of the above results follows by direct manipulation as for the results in Chapter 8, Section 5; see also Problem 8.5-3.

Let us now consider state estimate feedback schemes with the usual state feedback gain K' replaced by a frequency-shaped gain $K'(Q_r, s)$, and the usual state estimation gain K_e replaced by a frequency-shaped gain $K_e(Q_e, s)$; see Fig. 9.2-2. Of course, $K(Q_r, s)$ and $K_e(Q_e, s)$ are constructed as explained above. The main result of this section is now stated.
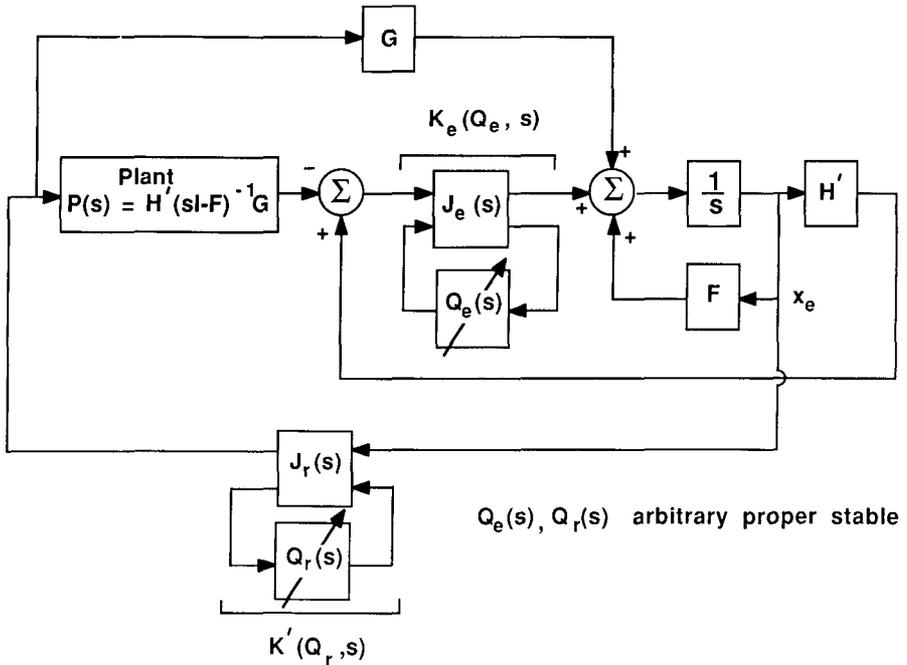
Figure 9.2-2 Frequency-shaped state estimate feedback.

Class of all Strictly Proper Stabilizing Controllers. Consider the plant $P(s)$, and a stabilizing state estimate feedback controller with constant state estimate feedback gain K', and constant state estimator gain K_e. Consider also the transfer function matrices $K(Q_r, s)$, $K_e(Q_e, s)$ parametrized in terms of arbitrary stable proper $Q_r(s)$, $Q_e(s)$ with $K = K(Q_r, s)|Q_r = 0$, $K_e = K_e(Q_e, s)|Q_e = 0$. Then the class of all stabilizing *strictly* proper controllers for $P(s)$ can be generated by the frequency-shaped state estimate feedback scheme with K replaced by $K(Q_r, s)$ and K_e replaced by $K_e(Q_e, s)$, as depicted in Fig. 9.2-2. Moreover, if K has full column rank, then the class can be generated with the gain pairs K, $K_e(Q_e, s)$ in terms of arbitrary stable proper $Q_e(s)$, and if K_e has full column rank, then the class can be generated with the gain pairs $K(Q_r, s)$, K_e.

The proof that the classes described above are stabilizing follows since $K'(Q_r, s)$, $K_e(Q_e, s)$ stabilize $P_r(s)$, and $P_e(s)$ respectively, for all stable proper $Q_r(s)$, $Q_e(s)$, and the modes of the closed-loop system are the union of the modes of the two closed-loop systems of Fig. 9.2-1; see Problem 9.2-1. For full details, see [6].

That the classes described generate the *entire* class of stabilizing strictly proper controllers is more difficult to prove. A key intermediate result is that the classes in this result are equivalent to the class of strictly proper controllers of Sec. 8.5

obtained by constraining $Q_f(s)$ to be strictly proper. Of particular interest are the following two cases.

$$\text{If } Q_r(s) = 0, \quad \text{then} \quad Q_f(s) = -K'(sI - F - K_eH')^{-1}Q_e(s), \quad (9.2\text{-}8a)$$

$$\text{If } Q_e(s) = 0, \quad \text{then} \quad Q_f(s) = -Q_r(s)(sI - F - GK')^{-1}K_e \quad (9.2\text{-}8b)$$

Details to establish the claims of (9.2-8) are in [5] and are not repeated here. Notice that under the full column rank conditions on K and K_e, then (9.2-8) suggests how Q_e or Q_r can be defined from Q_f.

$$\text{If } Q_r(s) = 0, \quad \text{then} \quad Q_e(s) = -(sI - F - K_eH')K(K'K)^{-1}Q_f(s) \quad (9.2\text{-}9a)$$

$$\text{If } Q_e(s) = 0, \quad \text{then} \quad Q_r(s) = -Q_f(s)(K_e'K_e)^{-1}K_e'(sI - F - GK') \quad (9.2\text{-}9b)$$

Hence if $Q_r(s) = 0$ an arbitrary proper stable $Q_e(s)$ maps via (9.2-8a) into a strictly proper $Q_f(s)$, and consequently leads to a stabilizing controller via the theory of Sec. 8.5. Likewise, for any stabilizing strictly proper controller there is an associated strictly proper $Q_f(s)$ and via (9.2-9a) an associated proper $Q_e(s)$. Corresponding results hold for the case $Q_e(s) = 0$. These special cases give insight into the more general results claimed above.

The relevance of the above result for us here is that with frequency shaping allowed in either the state feedback gain or state estimator gain, the class of all stabilizing strictly proper controllers can be generated, given constant K_e or K, assuming the full rank conditions on K_e or K. With K, K_e resulting from an LQG design then the full rank conditions are satisfied with G, H full rank and the Riccati solutions P and P_e positive definite. Notice that with the frequency shaping in the state feedback gain, then the state estimates x_e are unaffected by the frequency shaping, whereas when the frequency shaping is in the estimator gain, then x_e is no longer a standard state estimate but should be thought of as a frequency-shaped estimate.

It is important to realize that even low-order stabilizing strictly proper controllers with order much less than that of the plant can be viewed as belonging to the class of frequency-shaped state estimate feedback schemes—even though generically these are actually of higher order than that of the plant. In such cases, the frequency shaping to some extent cancels some dynamics of the state estimation step in a controller design.

An extension of the result of this section, established in [6], is that for the scheme of Fig. 9.2-2 with an additional arbitrary stable proper filter $Q_f(s)$ driven from the residuals $\xi = (y - H'x_e)$ and with output adding to the state estimate feedback $K'(Q_r, s)x_e$, there is still stability of the closed-loop system. Moreover, with $Q_e(s)$, $Q_r(s)$ fixed and chosen such that either $K(Q_r, s)$ or $K_e(Q_e, s)$ is minimum phase (full rank in Re $s \geq 0$), then the entire class of stabilizing proper controllers is characterized in terms of $Q_f(s)$. This is a generalization of the results in Sec. 8.5 and this section. This observation leads to results, proved in [6], which allow the strict properness controller constraint on the results of this section to be relaxed to just a

properness constraint. Thus, with $Q_r(s)$ fixed such that $K(Q_r, s)$ is minimum phase, then the class of all stabilizing proper controllers is parametrized in terms of arbitrary stable proper $Q_e(s)$ and arbitrary constant Q_f. The dual also holds, in that with $Q_e(s)$ fixed such that $K_e(Q_e, s)$ is minimum phase, then the same class is generated in terms of arbitrary stable proper $Q_r(s)$, and arbitrary constant Q_f.

In a practical design situation, it may well be better to include some frequency shaping in the estimator and some in the state estimate feedback law, rather than trying to achieve all the effects in the one step of the design. This is suggested by an example in the next section; however, theory provides at this stage no overall guidance as to the best allocation of the frequency shaping.

In the next section, we explore the case when $K(Q_r, s)$, $K_e(Q_e, s)$ are proportional plus integral gains rather than merely constant gains as treated in earlier chapters.

Main points of the section.   All stabilizing strictly proper controllers for a nominal plant can be viewed as state estimate feedback arrangements with frequency shaping in either the state feedback gain, or the gain in the state estimator, or both. The addition of an arbitrary constant gain between the estimator residuals and feedback control allows generation of all proper stabilizing controllers. The message of this section is, then, that frequency-shaped linear quadratic designs have the potential to improve on nonfrequency-shaped designs, provided that systematic procedures for introducing the frequency shaping can be advanced.

Problem 9.2-1

(a) For frequency-shaped state estimate feedback controlled as in Fig. 9.2-2, give a proof that the controllers are stabilizing for arbitrary stable proper $Q_e(s)$, $Q_r(s)$. [*Hint:* Generalize the proof for the situation of the previous chapter when $Q_e(s) = 0$, $Q_r(s) = 0$ and $K_e(Q_e, s) = K_e$, $K'(Q_r, s) = K'$.]

(b) Consider the scheme of Fig. 9.2-2, with the introduction of arbitrary stable $Q_f(s)$ driven from the residuals $\xi = (y - H'x_e)$ and adding to the state estimate feedback control. Establish that this is stabilizing. [*Hint:* Build on the proof approach of part (a)].

9.3 PROPORTIONAL PLUS INTEGRAL STATE FEEDBACK

In classical control design, it is known that proportional plus integral control can yield attractive controller performance and robustness properties. Also, such controllers reject (asymptotically) constant external disturbances, and have attractive low-frequency disturbance response properties. They are used for set point regulation. In this section we study proportional plus integral state feedback regulation in a linear quadratic optimization context, and give frequency domain interpreta-

tions of the frequency-shaped designs achieved. The work of this section is dualized in the following section to achieve proportional plus integral estimators which are then used in conjunction with the proportional plus integral state feedback regulators to achieve state estimate feedback designs with frequency shaping in both the estimator gains and state feedback gains, thus making connection with the results of the previous section.

The emphasis of this section is on exploiting properties and results of earlier chapters for design purposes rather than developing new theory. The design process here is a series of steps now presented with some rationalization at each step. We focus first on augmentations of the plant states with frequency-shaped filters.

Our starting point is the plant state space description, assumed minimal,

$$\dot{x} = Fx + Gu \qquad (9.3\text{-}1)$$

$$y = H'x \qquad (9.3\text{-}2)$$

We have in mind the situation when the plant input is subject to constant disturbances u_{ext} which are unknown, and we seek to regulate the plant states to zero in the presence of such disturbances. For the set point regulation case, there is an external reference r, assumed constant, and we seek a controller such that the plant output y tracks r in the presence of constant unknown disturbances u_{ext}.

A first trial approach.   A simple approach to tackling the above regulation problem is to assume in the design stage that $r = 0$ and $u_{ext} = 0$, and work with a performance index that penalizes the integral of y, or rather its square, as well as the usual terms $(x'Qx + u'Ru)$. Thus we work with the original plant augmented with integrators at the output. We shall apply standard LQ theory to the augmented plant, recognizing that the integral of y, denoted y^f, is now a linear combination of the states of the augmented plant, and so can be penalized in the performance index for the augmented plant. The augmented plant is described by

$$\begin{bmatrix} \dot{x} \\ \dot{y}^f \end{bmatrix} = \begin{bmatrix} F & 0 \\ H' & 0 \end{bmatrix} \begin{bmatrix} x \\ y^f \end{bmatrix} + \begin{bmatrix} G \\ 0 \end{bmatrix} u \qquad (9.3\text{-}3)$$

Here y^f, the integral of y, is penalized in a performance index associated with the augmented plant as follows:

$$V = \int_0^\infty (x'Qx + u'Ru + y^{f\prime}Q^f y^f)\, dt$$

$$= \int_0^\infty \left(\begin{bmatrix} x' & y^{f\prime} \end{bmatrix} \begin{bmatrix} Q & 0 \\ 0 & Q^f \end{bmatrix} \begin{bmatrix} x \\ y^f \end{bmatrix} + u'Ru \right) dt \qquad (9.3\text{-}4)$$

The optimal linear quadratic control has the form $K'x(t) + K_f' y^f(t)$ for some optimal gains K', K_f'. The situation is depicted in Fig. 9.3-1. Here u_{ext} is an external (constant) disturbance and r is a (constant) reference signal, both ignored in the initial design stage, but now recognized to be present. The control in the arrangement of Fig. 9.3-1 is now

$$H'(sI - F - GK')^{-1}G$$
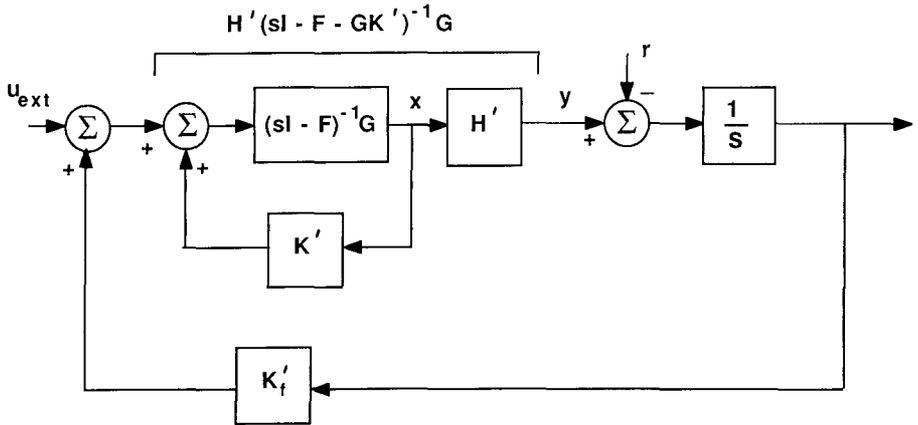
Figure 9.3-1   Proportional plus integral feedback.

$$u(t) = K'x(t) + K_f' \int_0^t [y(\tau) - r]\, d\tau + u_{\text{ext}} \qquad (9.3\text{-}5)$$

This is clearly a form of proportional plus integral feedback. Instead of proportional state feedback gains K' we now have

$$K'(s) = (K' + K_f' H' s^{-1}) \qquad (9.3\text{-}6)$$

Classical controllers frequently employ integral feedback. A well-known by-product is rejection of constant disturbances and/or achieving of set point regulation properties. The idea is that the input to the integrator can be asymptotically zero, yet the output asymptotically constant so as to allow cancellation of the disturbance u_{ext} and/or asymptotic set point reference tracking of the reference r.

For set point regulation, which includes standard regulation with a reference signal $r = 0$, in the absence of u_{ext} the requirement is that the closed-loop transfer function matrix from r to $(y - r)$ be zero at $s = 0$. Equivalently, as it turns out,

$$H'F^{-1}GK_f' \text{ is nonsingular} \qquad (9.3\text{-}7)$$

That is, the poles of the integrator (at the origin) must not be cancelled by zeros in the transfer function matrix $H'(sI - F)^{-1}GK_f'$. A necessary condition is that the dimension of the plant input u be no less than that of the plant output y, to achieve $H'F^{-1}G$ full row rank.

For rejection of disturbances u_{ext} having arbitrary constant values, clearly the dimension of y^f must be no less than that of u_{ext}. This and the previous dimension constraint suggest that we can achieve fully the goals of arbitrary set point regulation only in the presence of arbitrary constant disturbance inputs, when the plant is square and with no zero at the origin.

Further incorporation of classical control ideas.   The first trial approach just described may not always lead to suitably high performance

robust designs. We offer the following design steps, each one based on insights from classical control.

Step 1. Squaring the plant.   Output variables (termed regulated variables) y_1 are now defined, being linear combinations of the states as

$$y_1 = D'x \qquad (9.3\text{-}8)$$

The elements of y_1 will usually consist of actual plant measurements or combinations of such. The intention is to regulate the variables y_1 to zero, or have them track a constant reference r of arbitrary magnitude in the presence of arbitrary constant plant input disturbances. From our discussion above, it makes sense to seek such properties only when the dimension of y_1 is the same as that of u—thus the phrase "squaring the plant." Our approach is to penalize y_1, or filtered y_1, in the performance index. Clearly, it is preferable if the elements of y_1 have physical significance. This will usually be the case in a set point regulation situation. If y_1 is not measured, then it will be subsequently estimated, and this estimate set point regulated.

In squaring the plant we ensure that

$P_1(s) = D'(sI - F)^{-1}G$ is square with $[F, D]$ completely observable, and in addition when F is nonsingular there are no zeros at the origin ($D'F^{-1}G$ is nonsingular)   (9.3-9a)

and where possible achieve the following desirable property (discussed further below).

$P_1(s)$ is minimum phase within the bandwidth of significant control energy and has a smooth gain characteristic "close" to that of an easy-to-control low-order plant [any zeros of $P_1(s)$ in the pass band should be well damped].   (9.3-9b)

The nonsingularity of $D'F^{-1}G$ permits infinite loop gains at the origin with use of integral feedback. (When F is singular, the plant itself contains a pure integration, and it may not be necessary to use full integral feedback.) The minimum phase property for $P_1(s)$ is desirable if measurements y_1 are used to recover the state x through an estimator and the robustness associated with exact state feedback is not to be lost in using an estimator. Note also that the intention is to penalize y_1 in the performance index, and consequently the zeros of $P_1(s)$ will be attractors of closed-loop poles in the frequency band of significant loop gain. It may be for all reasonable sensor selections and combinations of measurements in squaring the plant that $P_1(s)$ is nonminimum phase in the bandwidth of the control action. In this case, the requirement here of "squaring the plant" could reasonably be relaxed to "squaring the augmented plant." Such an approach is discussed subsequently.

The idea of ignoring the phase characteristic of $P_1(s)$ is that linear quadratic

state feedback designs give automatic attention to plant input phase margins via the return difference inequality. The desire to have $P_1(s)$ minimum phase is motivated by the fact that a state estimate feedback design can inherit the phase margins of the state feedback design. The reason for seeking smooth gain characteristics of $P_1(s)$ close to those of a low-order system with well-damped zeros is that experience tells us that LQG designs give robust high performance controllers for such plants, whereas for arbitrary high-order plants there can be real robustness problems in an LQG design.

In order to apply proportional plus integral plus derivative (PID) classical control concepts, as distinct from proportional plus integral (PI) concepts, there is a requirement to achieve a proper controller that

$$D'G = 0 \tag{9.3-10}$$

If no suitable D can be found satisfying this constraint, then either the PI approach should be used and in the following development the differentiator gains K_D, k_D set to zero, or a filter with a proper transfer function employed, perhaps as an approximation to a PID filter. The constraint that $D'G = 0$ allows the construction of the derivative $\dot{y}_1$ as

$$\dot{y}_1 = D'\dot{x} = D'Fx + D'Gu = D'Fx$$

Step 2. Assigning zeros—Augmented plant construction.

The next step is to append proportional plus integral plus derivative (PID) filters to each element of y_1. Now a PID filter $(k_p + k_I s^{-1} + k_D s)$ has a pair of zeros and a pole at the origin. The PID gains are selected to achieve minimum phase target zeros intended to attract the closed-loop poles in a closed-loop design. This means that the zero assignments are approximate closed-loop pole assignments, so should be suitably damped and in appropriate frequency ranges. Again, classical control experience comes to bear in the assignment of PID coefficients. Notice, however, that the PID filters are not controllers so that their design is much more straightforward than for a classical controller design.

This step in our design process is where frequency shaping takes place. We have now constructed an augmented plant as in Fig. 9.3-2 with output

$$y_2 = K_P y_1 + K_I \int y_1 dt + K_D \dot{y}_1 \tag{9.3-11}$$
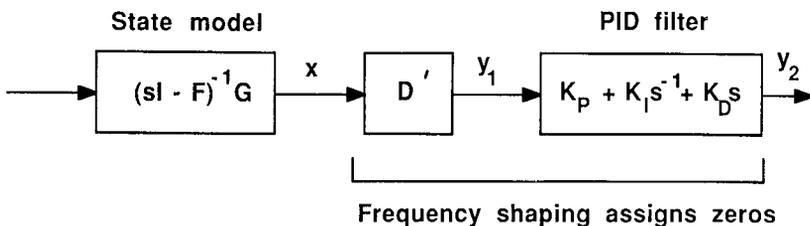
Figure 9.3-2 Augmented plant.

and transfer function matrix, easily verified to be

$$P_2(s) = [(K_P D' + K_D D' F) + K_I D' s^{-1}](sI - F)^{-1}G \qquad (9.3\text{-}12)$$

where K_P, K_D, K_I are diagonal matrices with elements $K_P^{(i)}$, $K_D^{(i)}$, $K_I^{(i)}$ being the individual PID filter coefficients. Its states consist of x and $x_1 = \int y_1 dt$, and its state equations are

$$\begin{bmatrix} \dot{x} \\ \dot{x}_1 \end{bmatrix} = \begin{bmatrix} F & 0 \\ D' & 0 \end{bmatrix} \begin{bmatrix} x \\ x_1 \end{bmatrix} + \begin{bmatrix} G \\ 0 \end{bmatrix} u \qquad (9.3\text{-}13)$$

$$y_2 = [(K_P D' + K_D D' F) \quad K_I] \begin{bmatrix} x \\ x_1 \end{bmatrix} \qquad (9.3\text{-}14)$$

Linear quadratic state feedback control laws for this augmented plant which penalize the augmented plant output y_2 have the structure

$$u = K'x + K_1'x_1$$

$$= K'x + K_1' \int (D'x)\, dt \qquad (9.3\text{-}15)$$

being proportional plus integral state feedback as in Fig. 9.3-2. Thus instead of proportional state feedback K', we have dynamic state feedback $K'(s) = (K' + K_1'D_s'^{-1})$. The situation is depicted in Fig. 9.3-3 where in the figure, we have additionally introduced r_1, an arbitrary constant set point reference signal, and u_{ext}, an external constant disturbance of arbitrary magnitude. For set point regulation (including the case $r = 0$), in the presence of a constant external disturbance u_{ext} of arbitrary magnitude, theory developed earlier in the section tells us that K_1 must be
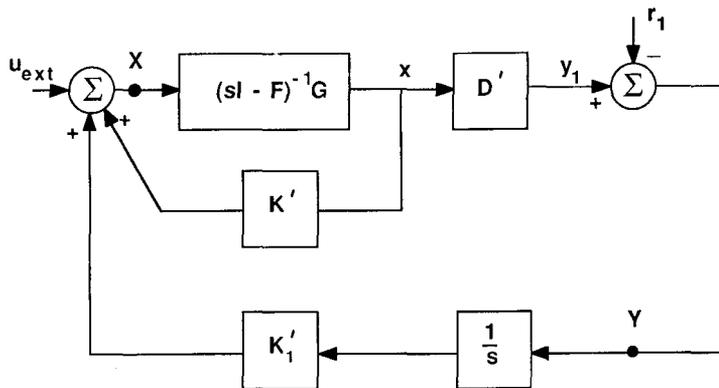
X   Control loop breaking point

Y   Reference loop breaking point

Figure 9.3-3   Proportional plus integral state feedback.

nonsingular; see also Problem 9.3-1. The (strict) minimum phase assumption on the PID augmentations ensures that no unstable pole-zero cancellations can take place in forming the augmented plant, thereby ensuring stabilizability and detectability of the augmented plant. As a consequence, the linear quadratic controller exists and is guaranteed to be stabilizing.

Of course, a designer can choose more complex frequency-shaping filters than the simple decoupled PID filters. The filters used to replace the PID filter can reasonably be proper, obviating the need of (9.3-10), and have poles and zeros such that the augmented plant has a boosted response in any frequency band of interest. For example, it may be that there is a desire to suppress resonances at a particular frequency. In this case it makes sense to have the frequency-shaping filter empha- size the resonance frequency band so that the performance index more heavily penalizes response in this frequency band. Such a situation arises in a design example studied subsequently. Clearly, when PID filters are used, the integrators emphasize low frequencies and the differentiators emphasize high frequencies. Alternative frequency shaping may well be more appropriate. The appropriate selection of more general frequency-shaping filters appears a more difficult task than for the PID filters studied here.

Should the squared plant $P_1(s)$ have nonminimum phase zeros, then $P_1(s)$ augmented with filters, as above, will also have the same zeros. To avoid such a situation in the case when there are more plant sensors than controls, it could well be possible to achieve a squared augmented minimum phase plant via a non square $P_1(s)$. In this case $P_1(s)$ is augmented with PI/PID filters that are coupled.

Step 3. Performance index formulation.   Consider now a
performance index associated with the augmented plant with input u, output y_2, and transfer function matrix $P_2(s)$ of (9.3-12), as

$$V = \int_0^\infty (y_2' Q_2' y_2 + u'Ru) \, dt \qquad (9.3\text{-}16)$$

This index is seen to penalize the augmented plant outputs y_2 which are frequency-shaped versions of the regulated variables y_1. Optimizing this index is a means to achieving appropriate control signal bandwidths and reference signal response bandwidths in the system, constrained by the fact that each actuator is known (or assumed) to be effective only over a certain range of frequencies. Thus any index optimization, irrespective of the physical meaning of the index, must achieve realistic bandwidth goals consistent with actuator/plant constraints. Like- wise, any regulated response with physical interpretations is known to be effectively controlled over a certain frequency band, associated with the modes that influence strongly the variables. Again, any index optimization must achieve results consistent with this knowledge.

Here we propose to start with a crude trial selection $Q_2 = I$, $R = I$, and pro- ceed in a trial and error design. Other starting designs based on Chapter 5 results could also be used.

Step 4. Trial and error weight selection.   With a trial design, examine the open-loop frequency response and adjust weights as follows. Open the control loop at point X in Fig. 9.3-3 at the first input and examine the cross-over frequency, which we recall approximates the closed-loop control system bandwidth. If this bandwidth is too high (low) to achieve a practical design, then increase (decrease) the first component of the diagonal matrix R. Repeat for all inputs. Likewise, when considering response to reference signals r_i, the reference loop is opened at point Y in Fig. 9.3-3, and the following rule applied. Increase (decrease) the ith diagonal element of Q_2 if the reference-loop bandwidth is too low (high) for a practical design. Repeat for all i.

One can repeat the above trial adjustments a number of times until the resulting design achieves reasonable control-loop and reference-loop bandwidths. These bandwidths may differ depending on the bandwidth of reference signal and desired system response. For this stage of the design, it is of course important to introduce engineering insight associated with the plant to be controlled. It must be known what are the frequency ranges for actuators and sensors to be effective. In this design process the weights Q_2, R are merely design parameters.

The rationale for the weights adjustment/bandwidth trade-off follows closely that of Chapter 6, and is not repeated here. Experience supports the approach, but we offer no guarantees. We do not suggest that trial and error adjustments to the frequency shaping is an essential part of the design approach, although in some cases such may be necessary to refine a controller design.

Design example.   To illustrate aspects of the PID design approach of this section, and also what can go wrong in such an approach, let us return to the yaw damper problem of Secs. 8.2 and 9.1. Here there is a single-input, two-output plant. Our first step is to select a scalar regulated variable $y_1 = D'x$ to achieve, if possible, a minimum phase plant with a Bode gain plot that is like that of a second-order system. In addition, we require that its response closely reflect the response for the original plant outputs y_a, y_f. Since selecting y_1 as y_a, or y_f, or a linear combination of y_a and y_f leads to a nonminimum phase $P_1(s)$, we first discuss the simpler case when y_1 is selected so that $P_1(s)$ is minimum phase. Thus we select

$$D' = [0 \quad 0 \quad 3 \quad -2 \quad 0 \quad 0]$$

There result minimum phase zeros at

$$\{-3.3, \ -0.88 \ \pm \ j0.93, \ -0.007, \ -4.7 \times 10^{-8}\}$$

Three designs are studied for illustration purposes. The objective we now examine is to minimize the response of $y_1 = D'x$ to the noise disturbance— irrespective of the response to the plant outputs $y' = [y_a, y_f]'$.

Case 1.   Here we just penalize u and y_1 in a performance index and over a few trials "minimize" the peaks of the spectral response. We select the response achieved with the index $E[y_1^2 + 0.005u^2]$.

Case 2.   Here the integral of y_1, denoted y_2, is penalized in the index $[y_2^2 + 10u^2]$ and the weight on u^2 is selected to get the "best" spectral response of y_1.

Case 3.   Here y_1 is passed through an approximation to a PI filter to yield

$$y_2(s) = \frac{s^2 + 5s + 100}{s^2 + 5s} y_1(s)$$

The approximation avoids differentiation. Again, the response y_1 is "optimized," penalizing y_2 in the index as $\int (y_2^2 + 0.01u^2) \, dt$

The spectral responses for the three cases are presented in Fig. 9.3-4 indicating that a much more dramatic reduction is achieved by the PI augmentation.
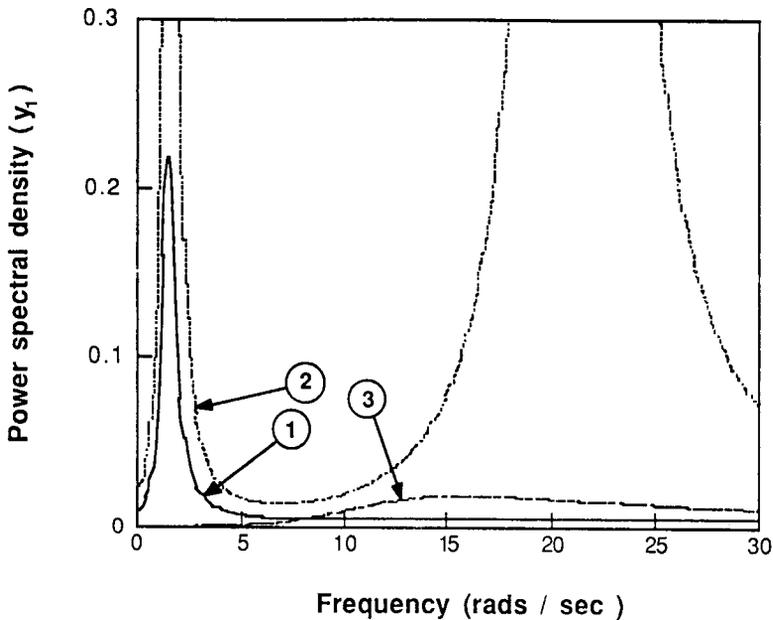
Figure 9.3-4   Improvement with PID frequency shaping.

So far, in this design example we have illustrated the power of the PID augmentation approach when $P_1(s)$ is so constructed as to be minimum phase. We also wish to point out the limitation of the approach when $P_1(s)$ is not minimum phase, as when y_1 is set as y_a or y_f or a linear combination. In this case, this PID approach, without modification, is unattractive in the state estimate feedback situation. In fact, the more sophisticated frequency shaping of the design of Sec. 9.1 turns out to lead to a better state estimate feedback design. To improve this PID approach, the augmented plant output should be selected as a combination of

filtered versions of both y_a and y_f so that the augmented plant output is minimum phase, as suggested in Step 2 above. Of course, a special case of this approach is the design of Sec. 9.1, so further designs are not repeated here.

The notion of proportional plus integral feedback as developed in this section can be extended to proportional plus integral plus double integral and so on, although details are not spelled out here.

Problem 9.3-3 studies an alternative approach to achieving proportional plus integral state feedback designs. It works with the augmentation of an integrator at the plant input for an initial design, being in the first instance a special case of the approach of [5].

The reader is referred to [7] for certain quite general frequency-shaped linear quadratic regulator results in the case of dynamic output feedback. Spectral factorization techniques are used rather than the augmentations of this section. For minimum phase plants the input robust properties associated with standard linear quadratic design are guaranteed.

Main points of the section.   Proportional plus integral state feedback designs can be achieved by applying linear quadratic methods to plants augmented with integrators or proportional plus integral plus derivative (PID) filters or approximations to these. The PID filters can be designed to give specified zeros between the controls and frequency-shaped regulated variables and to give open-loop gain responses that are like those of a low-order easy-to-control system. The zeros then attract closed-loop system poles in a state feedback design for the augmented plant and so shape the closed-loop response. The weights of the quadratic index can be systematically selected to achieve reasonable bandwidths associated with the control variables. The Riccati theory takes care of plant input phase margins. As in the case of classical controllers designed with proportional plus integral controllers, there is asymptotic cancellation of constant disturbances at the plant input for state proportional plus integral feedback designs (with K_1 nonsingular). Also, there is asymptotic tracking of constant reference signals even in the presence of unknown constant plant input disturbances. The more realistic situation of state estimate feedback is considered in the next section.

Problem 9.3-1.   Consider a linear quadratic design associated with the augmented plant of Fig. 9.3-1 and index (9.3-16) resulting in the controller arrangement of Fig. 9.3-3. Show that K_1 is nonsingular if $Q_2 > 0$, and K_I has full rank. [*Hint:* Examine the 22-block of the associated Riccati equation.]

Problem 9.3-2.   Show how to obtain a pair of complex zeros with poles at the origin by means of a proportional plus integral (PI) filter having two inputs and two outputs. Using such an arrangement in multivariable designs obviates the need for the differentiator and thus the constraint $D'G = 0$ assumed in this section, and its dual $H'D_e = 0$ of the next section.

Problem 9.3-3

(a) Set up a frequency-shaped regulator design by augmenting the plant $\dot{x} = Fx + Gu$ at the input with integrators, so that the augmented plant has an input $\dot{u}$.

(b) Show how state feedback for the augmented plant results in a dynamic feedback controller for the original plant.

(c) Give a method to reorganize this as proportional plus integral state feedback. [*Hint:* In achieving proportional plus integral state feedback, use the plant state equations as well as those of the controller.]

(d) Show that constant input disturbances are asymptotically rejected by such designs.

(e) When applying linear quadratic methods for the augmented plant state feedback design, interpret the index in terms of the original plant variables.

(f) What can be said about input robustness properties of this form of frequency-shaped linear quadratic design?

9.4 PROPORTIONAL PLUS INTEGRAL STATE ESTIMATE FEEDBACK

In this section we first briefly present procedures dual to those in the previous section for state estimator design. These result in state estimators with proportional plus integral gains, and asymptotic rejection in the state estimation of certain constant plant disturbances entering with the process noise. Such estimators are then combined with the proportional plus integral regulator designs of the previous section to achieve frequency-shaped state estimate feedback.

Estimator with proportional plus integral gain.   The full-order state estimators of Chapter 7 have the property that if constant disturbances enter where the process noise normally enters, but otherwise the process noise and the measurement noise are zero, state estimates x_e do not approach the states x. Of course, by increasing the estimator loop gains there can well be a decrease in the asymptotic errors $x - x_e$, but then sensor noise may not receive adequate filtering. Here we propose a frequency-shaped estimator having proportional plus integral gains rather than just a proportional gain as in the standard case. There is then, under reasonable conditions, asymptotic rejection of constant disturbances appearing with the process noise in that $x_e \to x$, assuming zero process and measurement noise. This property suggests strongly a more robust design than for the standard estimator. The design stages presented are duals of those for the proportional plus integral state regulator, and are therefore only briefly outlined.

Step 1. Square the signal process model.

Our starting point is a stochastic signal model based on the deterministic model (9.3-1) and (9.3-2).

$$\dot{x} = Fx + Gu + D_e v_1 \tag{9.4-1}$$

$$y = H'x + w \tag{9.4-2}$$

Here we have included a fictitious colored process disturbance $D_e v_1$ and white sensor noise w, where D_e is selected so that the transfer function $P_{e1}(s) = H'(sI - F)^{-1}D_e$ is square and desirably minimum phase, and in the case that F is nonsingular with $H'F^{-1}D_e$ nonsingular. Further, any constant disturbance later postulated to be acting on the plant must be of the form $D_e b$, for some constant b. (It is such a disturbance for which we shall demonstrate rejection). This is a further restriction on D_e. In the case that PID concepts are employed rather than just PI concepts, there is the additional constraint $H'D_e = 0$. Again, it is preferable that $P_{e1}(s)$ either have no zeros, or have heavily damped zeros, and again relaxation of these requirements can be achieved as discussed for the dual results of the previous section.

Step 2. Zero assignment.

Now append PI filters (or approximate PID filters) at the process noise input point as in Fig. 9.4-1. (For simplicity, we consider only PI implementations, and v_2 is a zero mean, white noise having covariance $\hat{Q}_2\delta(t - s)$. Later, we shall allow the possibility of constant disturbances being introduced at v_1, but this will occur after the estimator design). In Fig. 9.4-1 K_{eP} and K_{el} can be block diagonal, allowing zero assignment with the view to attracting closed-loop filter poles in a state estimator design for the augmented model. The augmented signal model has a transfer function

$$P_{e2}(s) = H'(sI - F)^{-1}(D_e K_{eP} + D_e K_{el}s^{-1})$$

The noise $D_e K_{el} \int v_2 \, dt$ represents for design purposes either plant or disturbance uncertainty at low frequencies, and is actually a Wiener process.
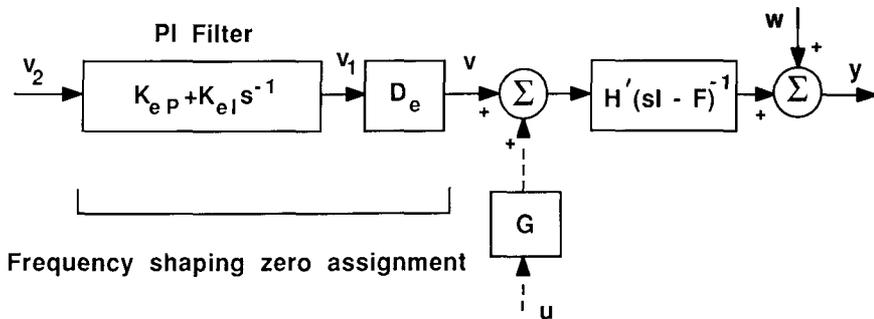
Figure 9.4-1 Augmented signal model.

Now applying a full-order state estimator design for this model leads to the frequency-shaped estimator design structure of Fig. 9.4-2.

Step 3. Setting bandwidths.

The quadratic weights are now the intensities $\hat{R}$ of the sensor noise w, and intensities $\hat{Q}_2$ of the process noise v_2. Here these parameters are adjusted to set estimator loop bandwidths, rather than being known a priori quantities. Initially set $\hat{R} = I$, $\hat{Q}_2 = I$ in the absence of further information. Evaluate open-loop bandwidths by breaking the resulting estimator design at the points X_e, Y_e of Fig. 9.4-2. The loops are broken one at a time for each line in the vector of lines at the points X_e, Y_e. The loops are termed sensor (estimator) and uncertainty loops, respectively, as indicated in Fig. 9.4-2. The term *uncertainty loop* is used because the loop exists to take care (in part) of plant uncertainty at low frequencies. Associated with the break point X_e at the ith sensor there is a sensor (estimator) loop cross-over frequency. If this coincides with the bandwidth over which the ith sensor gives reliable information then accept the ith diagonal element of $\hat{R}$. If the cross-over frequency is too low (high), then decrease (increase) the ith diagonal element of $\hat{R}$. Likewise for the break point Y_e giving rise to uncertainty loops, increase (decrease) the ith diagonal element of $\hat{Q}_2$ to increase (decrease) the cross-over frequency associated with the ith uncertainty loop, which copes with the ith element of the disturbance (uncertainty) v_1; see discussion below. Repeat the process of adjustments four times or so until there is a reasonable compromise reached in terms of loop bandwidths.

The above procedure exploits the rough equality between open-loop cross-over frequency and closed-loop bandwidth mentioned in Chapter 6. For the

Proportional plus integral gain

X_e Sensor estimator loop breaking point

Y_e Uncertainty loop breaking point
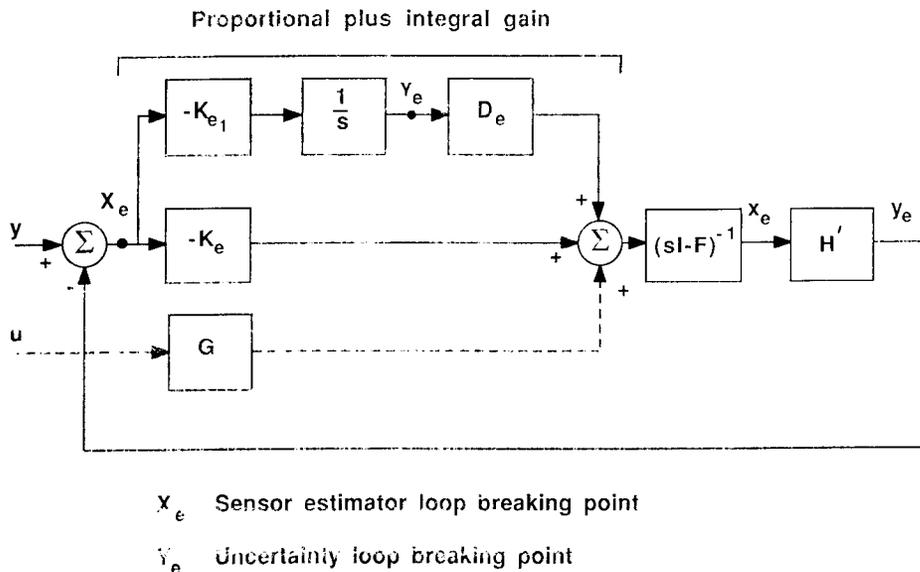
Figure 9.4-2   Frequency-shaped estimator.

ith sensor, the bandwidth of significance is that of the transfer function linking $y^{(i)}$ to $y_e^{(i)}$, the estimate of $y^{(i)}$. Typically, sensors tend to be fairly accurate over some low-frequency band and become unreliably noisy at high frequency. Knowledge of such sensor characteristics sets the estimator bandwidth in the above design procedure. Now the disturbance v_1 represents plant uncertainty together with process noise, and is frequency shaped. We expect a bandpass characteristic between v_1 and a state estimation error $(x - x_e)$, with the low-frequency corner frequency associated with the dominant time constant of this state estimation. It is this frequency (time constant) which is adjustable via the diagonal entries of $\hat{Q}_2$.

Properties of proportional plus integral estimator.

Some of these have been foreshadowed. The estimator loop gain is high at low frequencies by virtue of the integrator in the loop and the nonsingularity conditions. It is optimum for a plant process noise which is high in intensity at low frequencies, relative to the plant sensor noise. We conclude that the estimator relies on sensor data more at low frequencies and the model more at higher frequencies. This is a desirable property for an estimator in practice, since frequently sensor information is unreliable at high frequencies. Of course, one could attempt to achieve similar estimator properties with a signal model which involves a differentiation in a model for the sensor noise, and a nonfrequency-shaped plant process noise; however, such an approach with differentiation of white noise in the model is not a well-defined filtering task.

A crucial property for robustness to low-frequency disturbances or errors in modeling at low frequencies is the steady-state response properties of the estimator. These are studied in Problem 9.4-1. The key property is that with K_{e1} and $H'F^{-1}D_e$ nonsingular then $x_e \to x$, irrespective of the presence of an extra constant disturbance entering the plant through a gain D_e. This disturbance introduces a constant error $-H'F^{-1}D_eb$ into y and $-F^{-1}D_e$ into x; the stability of the estimation loop ensures that there will be a corresponding $-H'F^{-1}D_eb$ part of y_e. As a result the input to the integrator in the estimator will not depend on b, while the integrator output will have a DC value b in order to cause $x_e = F^{-1}D_eb$. This is an attractive robustness property not achieved in a standard estimator.

A design example.   Let us consider the first-order signal model

$$\dot{x} = -x + 2u + v, \qquad y = x + w$$

with noise intensities $\hat{Q} = 1$, $\hat{R} = 1$. The estimator gain is then $K_e = -0.414$, and the estimated closed-loop pole is at -1.414. The unit step response from plant (noise) input v to filtered estimate gives an unacceptable steady state filtered output error $y(\infty) - \hat{y}(\infty) = 0.7$, in contrast to a zero steady state error from a step input applied to the controls u. In a robustness test, changing the plant to

$$\dot{x} = -1.5x + 4u + 2v \qquad y = x + w$$

and preserving the estimator design leads to a steady state filtered output error of

0.44 from a plant input step response and 0.96 from the (noise) input v. These are now both unacceptable.

The steady state errors can be reduced by increasing the filter loop gain. Thus consider that $\hat{Q} = 1$, $\hat{R} = 0.01$ leading to an estimator gain $K_e = -9.05$ and closed-loop pole at -10.05. For the nominal plant, the steady state filtered error step response is still zero for a step input at u, and is reduced to 0.1 for a step from v, but the filter bandwidth is now too high and there is insufficient filtering of measurement noise.

Now consider a proportional-plus-integral-estimator designed using the nominal signal model augmented at the process noise input term with a proportional plus integral filter $(1 + 2s^{-1})$. The associated filter gain for the case $\hat{Q} = \hat{R} = 1$ is $K_e = (1.5 + 2s^{-1})$ and the closed-loop poles are at a frequency and damping $\omega = 1.414$, $\zeta = 0.86$. The step responses from u and v now both lead to zero steady state error in the filtered estimate for the nominal plant model and also for the second model used for a robustness test. Moreover, the filter bandwidths can be set as described earlier for good noise filtering and transient responses.

State estimate feedback.   The frequency-shaped design approach of this chapter is to use proportional plus integral estimators in conjunction with proportional plus integral state feedback regulator designs, replacing states in the control law with state estimates as in Fig. 9.4-3. There results what we term proportional plus integral LQG controllers. Clearly they are a subset of the class of all stabilizing controllers studied in Sec. 9.2, with the frequency-shaped gains being proportional plus integral gains. In Fig. 9.4-3, as sketched, the regulated variable y_1 is estimated as y_{1e} and is fed back via the integral feedback gain. Of course, if

Proportional plus integral state estimate feedback
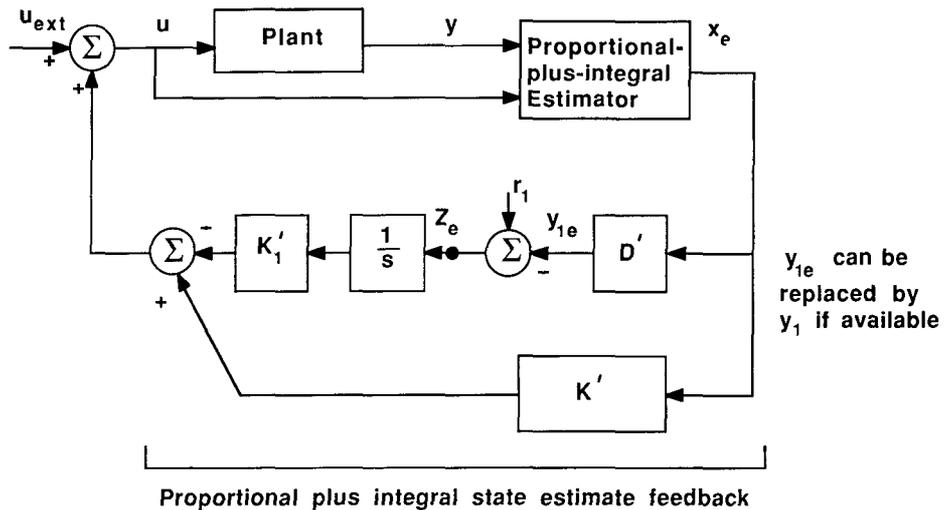
Figure 9.4-3   Proportional plus integral state estimate feedback.

$y_1 = Ly$ for some known L, then y_1 could be used in lieu of y_{1e} in the feedback arrangement depicted in Fig. 9.4-3. They have the important properties that under appropriate nonsingularity conditions, y_{1e} tracks the arbitrary constant reference signals r_1 in the presence of arbitrary constant disturbances u_{ext}. Moreover, this applies when there are constant plant input disturbances.

The eigenvalue separation property holds for the proportional plus integral LQG designs of this section, being a specialization of the eigenvalue separation property of more general frequency-shaped designs. By assigning the frequency-shaping minimum phase zeros, there is a crude assignment of those closed-loop poles attracted to these zeros in a closed-loop design.

By adjustment of the weights in the individual state feedback regulator and estimator designs, there is achieved a compromise in terms of control loop, and estimator loop bandwidths. In this regard, a significant property is that these sensor, reference, and uncertainty loop gains associated with the state estimator and state feedback regulator remain invariant upon switching to state estimate feedback. Derivations of this property are requested in the problems.

It is not surprising then that in the hands of an experienced designer, with classical and linear quadratic design insights, the frequency-shaped LQG design approach of this chapter can yield enhanced designs compared to the more standard ones of Chapter 8. The relative attractiveness of the frequency-shaped LQG design approach has been verified by a number of high-order designs for realistic aircraft models—see, for example, [8]. The ideas also build on other experience to achieve practical controllers as studied in [9].

Main points of the section.   Just as proportional plus integral plus differential (PID) controllers give robust classical designs, rejecting constant disturbances asymptotically, so frequency shaping with PI or PID filters in applying linear quadratic methods gives robust "optimal" designs involving proportional plus integral state estimate feedback and asymptotic suppression of certain constant disturbances. The last two sections show how to achieve proportional plus integral state estimate feedback designs in a systematic manner by using insights from classical control theory and the theory of earlier chapters. There is no claim that the approach is the best or universal, but rather that it is a systematic one that can lead to enhanced designs.

Problem 9.4-1   Consider the proportional plus integral estimator of this section. It is easily shown that the dual result of Problem 9.3-1 is that the integral gain K_{eI} is nonsingular, if K_{eI} is full rank, and $\hat{Q}_2 > 0$.

(a) Referring to Figure 9.4-2 for the case $u = 0$ with K_{eI} nonsingular, show that the steady state response of the proportional plus integral estimator to an asymptotically constant input y is $y_e(\infty) = y(\infty)$. [*Hint:* First show that the input to this integrator block in Fig. 9.4-2 is zero in steady state.]

(b) Suppose that the proportional plus integral estimator is driven by the plant with state x, constant input v_1 in place of process noise and having a transfer function

$H'(sI - F)^{-1}D_e$. Suppose also that K_{e1} is nonsingular as is the zero frequency gain $H'F^{-1}D_e$. Then show the steady-state state estimator error $[x(\infty) - x_e(\infty)]$ is zero.

Problem 9.4-2. Consider the frequency-shaped LQG design of Fig. 9.4-3 with estimator of Fig. 9.4-2, applied to the augmented plant model of Fig. 9.4-1. Verify that the sensor, uncertainty and reference loop gains are those of the open-loop estimator, and state feedback regulator as appropriate. It is assumed that the loops are opened at points X_e, Y_e indicated in Fig. 9.4-2 and the point Z_e in Fig. 9.4-3. [*Hint:* Referring to Fig. 9.4-2, first note that the effect of u cancels that of y as far as opening the loop at X_e, Y_e. When opening at Z_e, first see that $x_e \rightarrow x$ asymptotically.]

REFERENCES

[1] U. Shaked, "A General Transfer Function Approach to the Steady-State Linear-Quadratic-Gaussian Stochastic Control Problem," *Int. J. of Control,* Vol. 24, No. 6 (1976), pp. 77–80.

[2] M. G. Safonov, A. J. Lamb, and G. L. Hartmann, "Feedback Properties of Multi-variable Systems: The Role of Use of the Return Difference Matrix," *IEEE Trans. Auto. Control,* Vol. AC-16, No. 1 (Feb. 1981), pp. 47–65.

[3] N. K. Gupta, "Frequency-Shaped Loop Functionals: Extensions of Linear-Quadratic-Gaussian Design Methods," *J. of Guidance and Control,* Vol. 3 (Nov.–Dec. 1980), pp. 529–535.

[4] N. K. Gupta, M. G. Lyons, J. N. Auburn, and G. Mergulies, "Frequency Shaping Methods in Large Space Structures Control," *Proc. of AIAA Guidance and Control Conf.,* AIAA, New York, 1981.

[5] B. D. O. Anderson and D. L. Mingori, "Use of Frequency Dependence in Linear Quadratic Control Problems to Frequency Shape Robustness," *J. of Guidance and Control,* Vol. 8, No. 3 (May–June, 1985), pp. 397–401.

[6] J. B. Moore, K. Glover, and A. Telford, "All Stabilizing Controllers as Frequency Shaped State Estimate Feedback," *IEEE Trans. Auto. Control,* (1989) to appear.

[7] J. B. Moore and D. L. Mingori, "Robust Frequency-Shaped LQ Control," *Automatica,* Vol. 23, No. 5 (1987), pp. 641–646.

[8] C. M. Thompson, E. E. Coleman, and J. D. Blight, "Integral LQG Controller Design for a Fighter Aircraft," A.I.A.A. Conference Paper 87-2452, August 1987.

[9] D. Gangsaas, K. R. Bruce, D. J. Blight, and U.-L. Ly, "Application of Modern Synthesis to Aircraft Control: Three Case Studies," *IEEE Trans. Auto. Control,* Vol. AC-31, No. 11 (November 1986), pp. 995–1014.

10

Controller Reduction

10.1 INTRODUCTION: SELECTION OF FREQUENCY WEIGHTING

The techniques presented in this book to this point generally lead to controllers with order roughly equal to the plant order. In case the plant order is above four or five, it is natural to consider whether there might be a simpler (i.e., lower order) controller that will perform almost as well as the full-order controller resulting from a linear-quadratic design. Low-order controllers are normally to be preferred to high-order controllers, given comparable performance: there are fewer things to go wrong in the hardware, or bugs to fix in the software; their operation is easier to grasp at the conceptual level; that is, one is more likely to be able to identify parts of the controller as achieving certain subgoals of control, such as cancelling a pole, or injecting a phase compensation; and in a discrete-time implementation the computational requirements (operations per unit time and, probably, word length) are less.

These considerations motivate us to ask how a low-order design might be achieved. One approach is to seek to obtain a low-order controller directly; that is, one formulates the controller design problem ab initio with an order constraint in it. Call this the direct approach. Examples of the direct approach include the work of [1, 2], and we offer some discussion on this in Section 10.4. Generally, a quadratic optimization problem is posed with an order constraint and, naturally, a closed-loop stability constraint. Then there are two main issues to be considered. The first is

289

that of providing a satisfactory numerical procedure for executing the optimization. This is far from being a trivial task, and at the time of writing, no procedures are yet available in commercial control system software design packages, although it appears that future provision would not be out of the question. Moreover, even in-house packages that do exist require considerable experience and expertise for effective use. This leads us to the second issue, which relates to *utility of the whole* approach for achieving closed-loop design goals *apart from* minimization of a performance index. Much of our discussion has implicitly, and sometimes explicitly, focused on the fact that a quadratic index may possess little intrinsic significance, but rather serves as a vehicle for securing (via full-order LQG design) a number of closed-loop properties—relating to bandwidth, robustness, modal eigenvectors, and so on. The extent to which such closed-loop properties could be reflected in a quadratic index for a constrained order controller problem is simply not clear.

In contrast to the direct approach to low-order controller design, we can conceive of an indirect approach in which we use LQG methods to design a full-order controller, and then perform a further step, approximation of that full-order controller by a low-order controller. In this chapter, we shall study methods for carrying out this approximation.

There is a third possible approach to achieving low-order controller design which deserves mention. One begins the whole design procedure by approximating the plant with a low-order model. One designs (by LQG methods) a controller for this low-order model. The controller is, of course, attached to the original plant. There are two objections to this approach. First, in any design method involving approximation, it is logical to postpone approximation where possible until the later steps of the process. This is particularly so when it is not straightforward to keep track of the effect in later stages of a design process of an approximation made in an early stage of that design process. Performance of a low-order controller must be assessed for a model that approximates the plant as closely as possible, so that there should be no incentive to only define a low-order plant model and work with this exclusively. A second and more specific objection is that what constitutes satisfactory approximation of the plant necessarily involves the controller: it is the closed-loop behavior that one is ultimately interested in, and it is clear that a controller design could yield a situation in which very big variations in the open-loop plant in a limited frequency range had little effect on the closed-loop performance, while rather small variations in another frequency range could dramatically affect the closed-loop behavior. Now since the definition of a good plant approximation involves the controller, and since the controller is not known at the time of approximation, one is caught in a logical loop. Iteration might provide a way out, but the situation is by no means clear.

So let us return to the second idea, that of approximating a high-order controller by a low-order controller. For most designs, it is crucial to accept that the problem of controller reduction is distinct from the problem of (open-loop) model reduction, because of the presence of the plant, and because of the desire to have
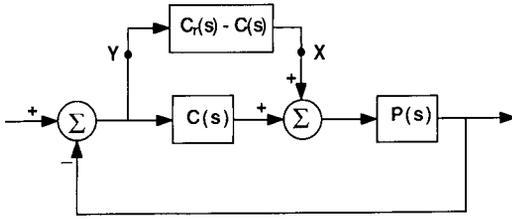
Figure 10.1-1   Redrawing of closed loop with compensation C_r.

good approximation in the closed-loop performance, which depends not just on the controller transfer function, but necessarily involves the plant.

Frequency weighting.   First and foremost, controller reduction has to preserve closed-loop stability. We want it also (as far as possible) to preserve other closed-loop properties, such as the closed-loop transfer function. Also, we may wish to maintain robustness properties in a controller reduction process. To preserve phase margins, for example, we may seek to preserve open-loop transfer functions in the vicinity of the cross-over frequency. Turning such goals into a quantitative statement generates a frequency-weighted approximation problem, as we shall now see. We focus first on closed-loop stability.

Let $P(s)$ be the transfer function matrix of a given plant, and let $C(s)$ be a stabilizing high-order series compensator with unity negative feedback. Let $C_r(s)$ be a low-order compensator, which we are seeking. Regard the system with compensator $C_r(s)$ replacing $C(s)$ as being equivalent to that of Fig. 10.1-1. Suppose that $C(s)$ and $C_r(s)$ have the same number of poles in Re $(s) \geq 0$, and in addition that, with $\bar{\sigma}\{\cdot\}$ denoting maximum singular value,

$$J_s = \max_\omega \bar{\sigma}\{[C(j\omega) - C_r(j\omega)]P(j\omega)[I + C(j\omega)P(j\omega)]^{-1}\} < 1 \qquad (10.1\text{-}1)$$

Then it is possible to show that $C_r(s)$ is also stabilizing. [The idea is very similar to one used in Chapter 5 in discussing the robustness of a design to plant variation, the difference here being that it is the controller rather than the plant which is varying. The result is suggested, but not proved, in the more restricted situation where $C - C_r$ is stable; now $C - C_r$ is stabilized by $P(I + CP)^{-1}$ as in Fig. 10.1-2. This is a redrawing of Fig. 10.1-1 to include a single transfer function matrix from X to Y, which is necessarily stable in view of the stabilizing property of $C(s)$. In Fig. 10.1-2, the two blocks are stable, and (10.1-1) states that the loop gain is smaller than 1. Hence stability of the closed loop in Fig. 10.1-2 follows, and thus also stability of the closed loop formed by C_r and P.]
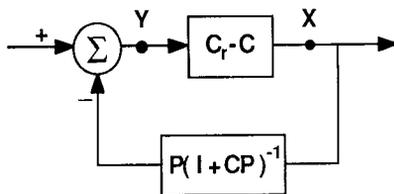
Figure 10.1-2   Redrawing of loop in Fig. 10.1-1.

Note that the requirements of $C(s)$ and $C_r(s)$ to have the same number of poles in Re $[s] \geq 0$ and for (10.1-1) to hold are sufficient conditions, not necessary conditions, for stability with the reduced-order controller. Again, in place of (10.1-1), one can have the condition

$$J_s' = \max_\omega \bar{\sigma}\{[I + P(j\omega)C(j\omega)]^{-1}P(j\omega)[C(j\omega) - C_r(j\omega)]\} < 1 \quad (10.1\text{-}2)$$

as an alternative sufficiency condition, when taken with equality of the pole count in Re $[s] \geq 0$ of C and C_r. This is obtained by virtually the same argument that led to (10.1-1), but is a different condition as soon as P, C are not scalar transfer functions.

Condition (10.1-1) and the condition that C and C_r have the same number of unstable poles suggest the following procedure for constructing a reduced-order controller. Write

$$C(s) = C_+(s) + C_-(s) \tag{10.1-3}$$

where $C_+(s)$ is strictly proper; that is, $C_+(\infty) = 0$, with all poles in Re $[s] \geq 0$, and $C_-[s]$ has all poles in Re $[s] < 0$. Copy the unstable part of $C_+(s)$ into $C_r(s)$; thus

$$C_r(s) = C_+(s) + C_{-r}(s) \tag{10.1-4}$$

where $C_{-r}(s)$ has all poles in Re $[s] < 0$. Choose $C_{-r}(s)$ so that J_s is minimized over all $C_{-r}(s)$ of prescribed degree. (How this last step can be performed will be considered later.) Should the value J_s turn out to be less than 1, stability with $C_r(s)$ is assured. Should J_s exceed 1, then instability may be associated with the minimizing $C_r(s)$, or it may not. Because (10.1-1) is part of a sufficient, rather than necessary, condition, one cannot be sure.

It is interesting to consider the significance of the weighting $P(I + PC)^{-1}$. Were this weighting replaced by the identity matrix, and were we working with scalar C and C_r, this minimization problem would require us to obtain frequency response (Nyquist) diagrams as close as possible for C, C_r subject to a constraint on the number of unstable poles, and total number of poles of C_r. The weighting term implies that it is more important to have C and C_r close in some frequency ranges than others. Notice that $\bar{\sigma}\{P[I + CP]^{-1}\}$ will be small when either the maximum singular value $\bar{\sigma}(P)$ is small or the minimum singular value $\underline{\sigma}(C)$ is large. Thus this weighting will be small in the stopband, or in the passband, if the latter corresponds to use of a C to produce high-loop gain. On the other hand, it is more likely to be large near the unity gain cross-over frequencies for the loop gain C. This means that the weighting matrix tends to require greater accuracy in the controller approximation near the cross-over frequency, an idea which should be familiar from classical control with its concern for phase margins.

A further issue deserving of comment is that of scaling. Consider Fig. 10.1-1. Suppose that $P(s)$ is replaced by $P(s)\Lambda$ and $C(s)$ and $C_r(s)$ are replaced by $\Lambda^{-1}C(s)$, $\Lambda^{-1}C_r(s)$, for some constant nonsingular Λ. This introduces an apparently inessential coordinate basis change at the plant input; if Λ is diagonal, it amounts to a scaling of the different inputs. Now observe that J_s in (10.1-1) is changed to

$$J_s = \max_\omega \bar{\sigma}\{\Lambda^{-1}[C(j\omega) - C_r(j\omega)]P(j\omega)[I + C(j\omega)P(j\omega)]^{-1}\Lambda\} \quad (10.1\text{-}5)$$

which is certainly not the same, except in certain trivial cases, for example, scalar $P(s)$ or $\Lambda = \lambda I$ for some scalar λ. With a different criterion, a different optimal C_r is to be expected.

How should Λ be chosen? In the event that process and measurement noise signals are present, $u(\cdot)$ will be a stationary random process with spectrum $\Phi_{uu}(j\omega)$, and it might make sense to scale the zero-lag covariance matrix of u, viz.

$$E[u(t)u'(t)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi_{uu}(j\omega)\, d\omega$$

so that the diagonal entries are all unity. More generally, the notion of equalizing signal levels on different inputs seems sensible. But it is no guarantee of appropriateness of scaling constants.

The frequency weighting $P(I + CP)^{-1}$ is derived above by focusing exclusively on stability considerations. Let us note two other ways whereby a weighting could be advanced. First, suppose the original closed-loop system with high-order controller $C(s)$ operates in the presence of stationary process and measurement noise. Let $q(t)$ be the input signal to $C(s)$. In the absence of driving signals other than the noises, $q(t)$ will be a stationary random process with spectrum computable from C, P, and the driving noise spectrum. Let the spectrum of $q(t)$ be $\Phi_{qq}(j\omega)$. Now one could argue that it is most important for $C(j\omega)$ to be accurately approximated by $C_r(j\omega)$ in those frequency bands where most signal energy is present in actual operation. So if $V(j\omega)$ is a spectral factor of $\Phi_{qq}(j\omega)$ in the sense that $V(s)$ and $V^{-1}(s)$ are stable, and

$$V(j\omega)V'(-j\omega) = \Phi_{qq}(j\omega) \tag{10.1-6}$$

one could seek C_r such that C and C_r have the same number of unstable poles and

$$J_n = \max_{\omega} \overline{\sigma}\{[C(j\omega) - C_r(j\omega)]V(j\omega)\} \tag{10.1-7}$$

is minimized. [*Note:* Given $\Phi_{qq}(j\omega)$, it is possible to compute $V(j\omega)$; see [3], [4]].

A further alternative to the selection of a weighting is as follows. One could seek to choose C_r so that the closed-loop transfer function $PC[I + PC]^{-1}$ is closely approximated. Now

$$PC[I + PC] - PC_r[I + PC_r]^{-1} \simeq [I + PC]^{-1}P[C_r - C][I + PC]^{-1} \tag{10.1-8}$$

This suggests use of a two-sided weighting for $C_r - C$, that is, the minimization of

$$J_c = \max_{\omega}\{\overline{\sigma}[I + PC]^{-1}P[C_r - C][I + PC]^{-1}\} \tag{10.1-9}$$

The weightings in J_c should be compared with those in J_s' (see (10.1-2)). Evidently, errors between C_r and C receive less weighting in the high loop gain bands with (10.1-9) than with (10.1-2).

For a further variant on (10.1-9), one could incorporate a multiplicative factor to reflect knowledge about the spectrum of external inputs (if it were available). Thus one might seek to minimize, over C_r of prescribed order with the same unstable pole count as C, an index

$$J_c' = \max_\omega \{\overline{\sigma}[I + PC]^{-1}P[C_r - C][I + PC]^{-1}\hat{V}(j\omega)\} \qquad (10.1\text{-}10)$$

where the input spectrum is $\hat{V}(j\omega)\hat{V}'(-j\omega)$.

None of the indices advanced are free from some criticism. Thus J_n, J_c, J_c' take no direct account of stability, which is a disadvantage. On the other hand, J_s, J_s' take no account of performance issues apart from stability, and address the stability issue only via a sufficiency condition. For no index can it be said that it will assure that PC and PC_r have the same roll-off rate as $\omega \to \infty$. Some of these disadvantages will be addressed in Section 10.3 when we describe the use of fractional representation of the controllers; by using such representations, alternative indices can be found.

What of the problem of actually achieving the minimization of a particular index, with the constraints on C_r? It turns out that there is a straightforward approach to an approximate minimization. We discuss this in the next section.

Main points of the section.   Practical considerations suggest the desirability of low-order controllers. Reduction of the order of an LQG controller should usually be contemplated. Intelligent inclusion of the plant in setting up a formulation of the reduction problem leads to a frequency-weighted approximation problem. The weight may reflect concern with stability, spectra, or closed-loop transfer functions, that is, performance. Scaling of plant inputs affects the approximation problem.

Problem 10.1-1.   Suppose that $C_r(s) = C(s)[I + L(s)]$ and that $C(s)$ is stabilizing. Suggest a criterion involving $L(s)$ that would be pertinent for controller approximation, basing the criterion on stability issues. [*Hint:* Recall Section 5.3 results.]

10.2 FREQUENCY-WEIGHTED BALANCED TRUNCATION

In this section, we focus on the following problem. Given transfer function matrices $C(s)$ of order n and $W(s)$ with all poles in $\mathrm{Re}[s] < 0$, find $C_r(s)$ of order $r < n$, with all poles in $\mathrm{Re}\,[s] < 0$, such that

$$J = \max_\omega \overline{\sigma}\{[C(j\omega) - C_r(j\omega)]W(j\omega)\} \qquad (10.2\text{-}1)$$

is minimum. Of course, the choice of W for the controller reduction problem can be pursued as described in the last section. Note also that there is no real loss of generality in requiring $C(s)$ and $C_r(s)$ to be stable—if $C(s)$ is unstable, its unstable part is copied into $C_r(s)$.

We shall not solve the problem in the form given. Rather, we shall give a construction for a $C_r(s)$, in general not minimizing J, which nevertheless proves attractive in many examples, and is partly motivated by results for the case when $W(s) = I$. We shall first review this identity weighting case.

Consider (10.2-1) with $W(j\omega) = I$ and the problem of minimizing J. Let the

Hankel singular values of $C(s)$ in decreasing magnitude be σ_1, σ_2, ..., σ_n, and suppose $\sigma_r \neq \sigma_{r+1}$ (see Appendix B). Then it is possible, see [5], to prove that all stable $C_r(s)$ of degree r satisfy

$$\max_\omega \overline{\sigma}\{[C(j\omega) - C_r(j\omega)]\} \geq \sigma_{r+1} \qquad (10.2\text{-}2)$$

while the procedure of balanced realization truncation (reviewed below and noted in Appendix B) leads to

$$\max_\omega \overline{\sigma}\{[C(j\omega) - C_r(j\omega)]\} \leq 2(\sigma_{r+1} + \cdots + \sigma_n) \qquad (10.2\text{-}3)$$

(this result being established in [5] and [6].) An alternative procedure of [5], called Hankel singular value approximation, will lead to the removal of the multiplier 2 in (10.2-3) but with a possible penalty being incurred: if $C(s)$ is strictly proper, $C_r(s)$ will not be. Now if $\sigma_{r+2}, \ldots, \sigma_n$ are much smaller than σ_{r+1}, from (10.2-2) and (10.2-3) it is clear that balanced realization truncation necessarily comes close to the optimum. By "close," we mean that the error is effectively within a factor of 2 of the optimum. If the optimum is a constrained one, requiring $C_r(s)$ to be strictly proper if $C(s)$ is, then the lower bound in (10.2-2) is conjectured to be $2\sigma_{r+1}$. In this case, balanced realization truncation is likely to give something very near the optimum if $\sigma_{r+2}, \ldots, \sigma_n$ are very small. Comparisons on practical examples suggest that balanced realization has its own in-built frequency shaping which appears to enhance more often than degrade a controller approximation. This reason and its relative simplicity are behind our current preference for the balanced realization approach.

The discussion so far suggests that we should seek to replace the minimization of J in (10.2-1) by a variant on balanced realization truncation, which somehow allows incorporation of a frequency weight. This we shall do. To assist in understanding this variant, we shall first recall the procedure for balanced realization truncation.

Suppose that

$$C(s) = \overline{H}'(sI - \overline{F})^{-1}\overline{G} \qquad (10.2\text{-}4)$$

with $\{\overline{F}, \overline{G}, \overline{H}\}$ minimal. Of course, $\text{Re } \lambda_i(\overline{F}) < 0$ since $C(s)$ is stable. Let $\overline{P}$, $\overline{Q}$ be the infinite time controllability and observability gramians, satisfying

$$\overline{F}\overline{P} + \overline{P}\overline{F}' + \overline{G}\overline{G}' = 0 \qquad (10.2\text{-}5)$$

$$\overline{Q}\overline{F} + \overline{F}'\overline{Q} + \overline{H}\overline{H}' = 0 \qquad (10.2\text{-}6)$$

Then there exists a nonsingular matrix T such that in a new coordinate basis with $x = T^{-1}\overline{x}$, $F = T^{-1}\overline{F}T$, $G = T^{-1}\overline{G}$ and $H' = \overline{H}'T$, there holds

$$F\Sigma + \Sigma F' + GG' = 0 \qquad (10.2\text{-}7)$$
$$\Sigma F + F'\Sigma + HH' = 0 \qquad (10.2\text{-}8)$$

where

$$\Sigma = \text{diag}\,[\sigma_1, \sigma_2, \ldots, \sigma_n] \qquad \sigma_i \geq \sigma_{i+1} \qquad (10.2\text{-}9)$$

with the σ_i^2 the eigenvalues of $\bar{P}\bar{Q}$; that is, the σ_i are the Hankel singular values. Most control system software design packages contain procedures for "balancing" a linear system realization. See Problem 10.2-1 also for insight into the construction of T.

Unweighted approximation of $C(s)$ is easy: one simply selects the first r rows and columns of F, and the first r rows of G and of H, to define submatrices F_r, G_r, H_r for which

$$C_r(s) = H_r'(sI - F_r)^{-1}G_r \qquad (10.2\text{-}10)$$

Provided that $\sigma_r > \sigma_{r+1}$, it turns out that Re $\lambda_i(F_r) < 0$ for all i. In case $\sigma_r = \sigma_{r+1}$, there is some nonuniqueness in the construction of T and thus in $\{F_r, G_r, H_r\}$. The nonuniqueness can be exploited to obtain Re $\lambda_i(F_r) < 0$ for all i, but if it is not exploited, then there is a possibility that Re $\lambda_i(F_r) = 0$ for some i; see [7] for a full discussion. Let us assume here that $\sigma_r \neq \sigma_{r+1}$. Then, as proved in [6], the bound of (10.2-3) holds.

Now let us define a procedure, frequency-weighted balanced truncation, designed to generate a reduced order $C_r(s)$ such that $C_r(j\omega)W(j\omega)$, as a function of ω, is close to $C(j\omega)W(j\omega)$. Thus minimization of J in (10.2-1) is not achieved. But it is hoped that the procedure yields an acceptable $C_r(s)$ in a simple way. The idea was first developed in [6], and extends the procedure above.

Suppose that $C(s)$ is as in (10.2-4) and further that

$$W(s) = D_w + H_w'(sI - F_w)^{-1}G_w \qquad (10.2\text{-}11)$$

This means that

$$C(s)W(s) = [\bar{H}' \quad 0]\begin{bmatrix} sI - \bar{F} & -\bar{G}H_w' \\ 0 & sI - F_w \end{bmatrix}^{-1}\begin{bmatrix} \bar{G}D_w \\ G_w \end{bmatrix} \qquad (10.2\text{-}12)$$

The formula on the right is obtained by "cascading" the two given realizations of $C(s)$ and $W(s)$. The state vector of the combined realization is

$$\bar{x} = \begin{bmatrix} \bar{x}_c \\ x_w \end{bmatrix} \qquad (10.2\text{-}13)$$

where $\bar{x}_c$, x_w are state vectors associated with individual realizations of $C(s)$ and $W(s)$. Set up the equations defining the controllability and observability gramians for (10.2-12). Call these matrices $\bar{P}$, $\bar{Q}$ and partition them:

$$\bar{P}\begin{bmatrix} \bar{F}' & 0 \\ H_w\bar{G}' & F_w' \end{bmatrix} + \begin{bmatrix} \bar{F} & \bar{G}H_w' \\ 0 & F_w \end{bmatrix}\bar{P} + \begin{bmatrix} \bar{G}D_w \\ G_w \end{bmatrix}[D_w'\bar{G}' \quad G_w'] = 0 \qquad (10.2\text{-}14)$$

$$\begin{bmatrix} \bar{F}' & 0 \\ H_w\bar{G}' & F_w' \end{bmatrix}\bar{Q} + \bar{Q}\begin{bmatrix} \bar{F} & \bar{G}H_w' \\ 0 & F_w \end{bmatrix} + \begin{bmatrix} \bar{H} \\ 0 \end{bmatrix}[\bar{H}' \quad 0] = 0 \qquad (10.2\text{-}15)$$

$$\bar{P} = \begin{bmatrix} \bar{P}_{cc} & \bar{P}_{cw} \\ \bar{P}_{cw}' & \bar{P}_{ww} \end{bmatrix} \qquad \bar{Q} = \begin{bmatrix} \bar{Q}_{cc} & \bar{Q}_{cw} \\ \bar{Q}_{cw}' & \bar{Q}_{ww} \end{bmatrix} \qquad (10.2\text{-}16)$$

The top left corner $\bar{P}_{cc}$ of $\bar{P}$ can be thought of as the weighted controllability gramian for $C(s)$. The top left corner of $\bar{Q}$, viz. $\bar{Q}_{cc}$, satisfies

$$\overline{F}'\overline{Q}_{cc} + \overline{Q}_{cc}\overline{F} + \overline{H}\overline{H}' = 0 \qquad (10.2\text{-}17)$$

and is evidently unaffected by the weight. Obviously, it is the observability gramian for $C(s)$. Now we find a coordinate basis change matrix for $\bar{x}_c$, but not for x_w. Thus $x_c = T^{-1}\bar{x}_c$, and

$$x = \begin{bmatrix} T^{-1} & 0 \\ 0 & I \end{bmatrix}\bar{x} \qquad (10.2\text{-}18)$$

such that in the new coordinate basis $P_{cc} = Q_{cc} = \text{diag}\ [\lambda_1, \lambda_2, \ldots, \lambda_n]$, $\lambda_i \geq \lambda_{i+1}$. (The procedure explained in Problem 10.2-1 can again be used for this.) The λ_i^2 are the eigenvalues of $\overline{P}_{cc}\overline{Q}_{cc}$ and can be thought of as weighted singular values of $C(s)$. With $F = T^{-1}\overline{F}T$, $G = T^{-1}\overline{G}$, $H' = \overline{H}T$, the triple $\{F, G, H\}$ is termed a frequency-weighted balanced realization of $C(s)$. A frequency-weighted degree r approximant $C_r(s)$ of $C(s)$ is obtained by eliminating all but the first r rows and columns of F, and all but the first r rows of G, H. The resulting F_r, G_r, H_r define $C_r(s)$ by (10.2-10), and Re $\lambda_i(F_r) < 0$ provided $\lambda_r \neq \lambda_{r+1}$. (The arguments are virtually the same as in the unweighted case.)

We stress again that this procedure does not ensure that $C_r(s)$ minimizes the index (10.2-1). Indeed, we do not even have available analogues for the frequency-weighted case of the error formulas (10.2-2) and (10.2-3). We can justify the scheme only by appealing to its simplicity and its efficacy as displayed by examples.

The procedure just described dealt with "input" weighting; that is, $W(j\omega)$ affected the controllability, not observability gramian of C, and in (10.2-1), multiplication of a vector by $(C - C_r)W$ implies multiplication first by W, then by $C - C_r$. It is easy to formulate a dual procedure allowing output weighting (see Problem 10.2-3). It is in fact possible to formulate a procedure allowing simultaneous input and output weighting (as would be required if the index associated with closed-loop transfer function approximation is used; see [6]. This index is described in the previous section.)

It is instructive to note the size of the matrix equation involved in weighted balanced truncation. Suppose $C(s)$ is of order n (and is stable) and $W(s)$ is of order l. Then $\overline{P}$ is defined by an equation of dimensions $(n + l) \times (n + l)$. Now if W is the weighting associated with the first (stability based) approximation procedure noted in the previous section, then $W = P(I + CP)^{-1}$ and if P has order n, one expects W to have order $l = 2n$. Thus $\overline{P}$ is defined by a $3n \times 3n$ equation so the search for an order 10 approximation to an order 50 controller may involve an equation of dimension 150×150, which is sizeable.

Actually, when $P(s) = H'(sI - F)^{-1}G$ and $C(s) = K'(sI - F - GK' - K_eH')^{-1}K_e$ (as in a linear quadratic design), and also Re $\lambda_i(F + GK' + K_eH') < 0$ for all i, then the equation for P, the controllability gramian, can be decomposed easily into three $n \times n$ equations (see Problem 10.2-2).

Another rather ad hoc procedure for easing the dimensionality burden involves initially finding unweighted approximations of P and C for the purposes of obtaining a lower-order weight W_r, which is then used with the original C, or even an unweighted reduction of C if there is negligible "error" in this, to determine a

weighted approximation. It seems likely that as long as this lower-order W_r captures the gross characteristics of W, there will be little effect on the resulting C_r.

The controller reduction scheme presented above has been applied to a number of examples, [6]. These include a controller for a plant comprising four spinning disks. The disks are connected by a flexible rod, a motor applies torque to the third disk, and the angular displacement of the first disk is the variable of interest. The plant transfer function is

$$G(s) = \frac{1}{4s^2} \frac{\frac{s^2 + 2\zeta_0\omega_0 s + \omega_0^2}{\omega_0^2} \cdot \frac{s^2 + 2\zeta_1\omega_1 s + \omega_1^2}{\omega_1^2} \cdot \frac{s + a}{a}}{\frac{s^2 + 2\zeta_2\omega_2 s + \omega_2^2}{\omega_2^2} \cdot \frac{s^2 + 2\zeta_3\omega_3 s + \omega_3^2}{\omega_3^2} \cdot \frac{s^2 + 2\zeta_4\omega_4 s + \omega_4^2}{\omega_4^2}}$$

with

$$\zeta_0 = 0.02 \qquad \omega_0 = 1$$
$$\zeta_1 = -0.4 \qquad \omega_1 = 5.65$$
$$\zeta_2 = \zeta_3 = \zeta_4 = 0.02 \qquad \omega_2 = 0.765$$
$$a = 4.84 \qquad \omega_3 = 1.41$$
$$\omega_4 = 1.85$$

Note that the system is nonminimum phase (because $\zeta_1 < 0$). A minimal realization in modal coordinates is provided by

$$F = \text{diag}\left\{ \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} -0.015 & 0.765 \\ -0.765 & -0.015 \end{bmatrix}, \begin{bmatrix} -0.028 & 1.410 \\ -1.410 & -0.028 \end{bmatrix}, \begin{bmatrix} -0.04 & 1.85 \\ -1.85 & -0.04 \end{bmatrix} \right\}$$

$$g = \begin{bmatrix} 0.026 \\ -0.251 \\ 0.033 \\ -0.886 \\ -4.017 \\ 0.145 \\ 3.604 \\ 0.280 \end{bmatrix} \qquad h = \begin{bmatrix} -0.996 \\ -0.105 \\ 0.261 \\ 0.009 \\ -0.001 \\ -0.043 \\ 0.002 \\ -0.026 \end{bmatrix}$$

Now the loop shape constraints imposed by performance (low-frequency constraint) and robustness in the face of unstructured uncertainty (high-frequency constraint) turn out to require the loop gain to lie outside the shaded region in Fig. 10.2-1. Note that the unity gain cross-over frequency can be kept well below the frequency associated with the nonminimum phase zeros, so they should present little problem in securing an adequate design.

The first step is to design a state feedback law k. Note that, because the high-frequency constraint rolls off at 40 dB/decade, while $k'(j\omega I - F)^{-1}g$ can roll off at only 20 dB/decade, this state feedback law cannot result in $k'(j\omega I - F)^{-1}g$
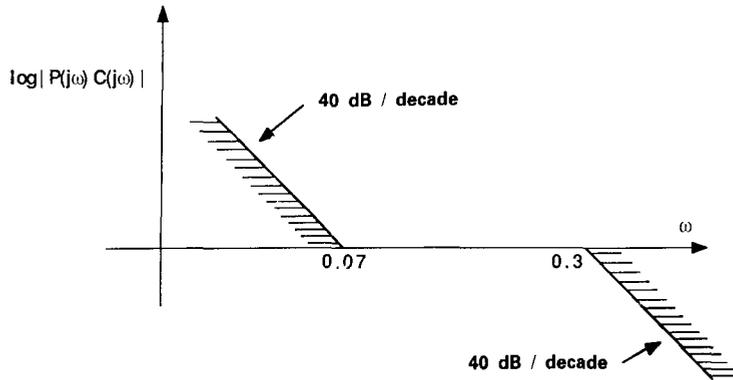
Figure 10.2-1   Constraints on loop gain for disk example.

satisfying the high-frequency constraint. We must rely on an additional roll-off being provided by the estimator.

We determine the gain k by trial and error, using a state weighting matrix Q that is diagonal. Entries of Q are adjusted, up or down as required, to give a loop gain $k'(j\omega I - F)^{-1}g$ which in some way meets the constraint as far as possible; in particular, the low-frequency constraint is met. The choice of Q is

$$Q = \text{diag}\{2 \times 10^{-3}, 2 \times 10^{-3}, 8 \times 10^{-2}, 8 \times 10^{-2}, 8 \times 10^{-3}, 8 \times 10^{-3}, 3 \times 10^{-3}, 3 \times 10^{-3}\}$$

while $r = 1$. The resulting k is

$$k' = [4.47 \times 10^{-2} \quad 6.61 \times 10^{-1} \quad 4.14 \times 10^{-3} \quad 3.59 \times 10^{-1}$$
$$1.03 \times 10^{-1} \quad -3.70 \times 10^{-2} \quad -4.38 \times 10^{-2} \quad 3.38 \times 10^{-2}]'$$

The Nyquist plot for $-k'(j\omega I - F)^{-1}g$ is depicted in Fig. 10.2-2. Note the avoidance of the disk of center $-1 + j0$ and radius 1. The magnitude of the loop gain is plotted in Fig. 10.2-3. As foreshadowed above, the magnitude constraint is not met at high frequencies.

Next, an estimator gain is determined. Using a noise covariance matrix

$$\hat{Q} = Sgg'S$$

where

$$S = \text{diag}[0.346 \quad 0.346 \quad 0.024 \quad 0.024 \quad 0.042 \quad 0.042 \quad 0.077 \quad 0.077]$$

and $\hat{R} = 1$ results in

$$k_c' = [4.111 \times 10^{-1} \quad 8.70 \times 10^{-2} \quad 3.78 \times 10^{-4} \quad -3.66 \times 10^{-3}$$
$$-8.24 \times 10^{-3} \quad 8.72 \times 10^{-3} \quad -7.41 \times 10^{-5} \quad 1.41 \times 10^{-2}]$$

The effect of S can be explained in the following way. The process noise is coupled into modes with different intensities as changed by S. The coupling into the modes
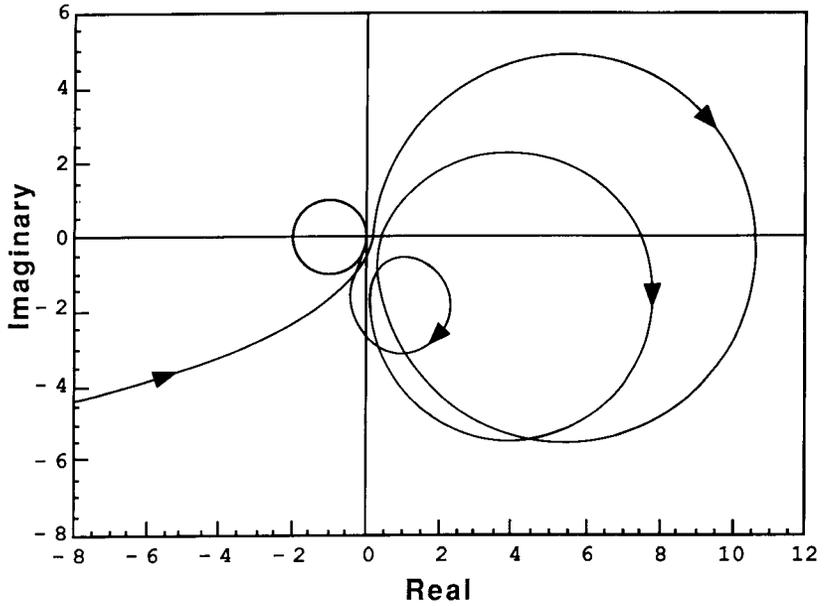
Figure 10.2-2   Nyquist plot of state feedback design gain.
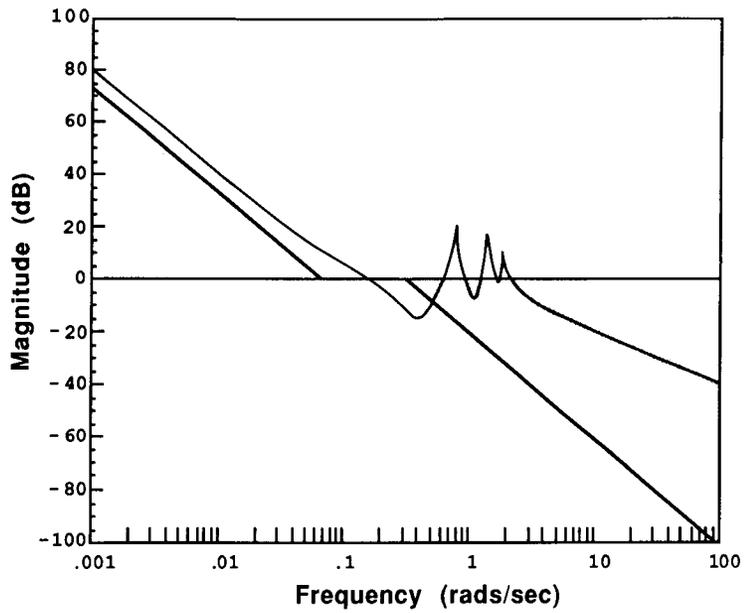
Figure 10.2-3   Bode plot of state feedback design gain.

associated with poles at the origin is high, while the coupling associated with high-frequency modes is much lower. As a result, there is more suppression by the estimator of high frequencies; that is, the loop gain is effectively decreased at high frequencies.

The loop gain when the full-order control is used is

$$|P(j\omega)C(j\omega)| = |h'(j\omega I - F)^{-1}gk'(j\omega I - F - gk' - k_e h')^{-1}k_e|$$

and is plotted in Fig. 10.2-4. Notice that the constraints are met. For controller reduction, we need to form the weighting function

$$W(j\omega) = P(j\omega)[I + C(j\omega)P(j\omega)]^{-1}$$

The magnitude of this function is plotted in Fig. 10.2-5. Of course, large values correspond to frequencies where more accurate controller approximation is desired.

When the controller is reduced to a dimension of four, we obtain

$$C_r(j\omega) = \frac{0.0513(j\omega)^3 + 0.00424(j\omega)^2 + 0.0296(j\omega) + 0.00157}{(j\omega)^4 + 0.693(j\omega)^3 + 0.779(j\omega)^2 + 0.293(j\omega) + 0.0739}$$

The loop gain $P(j\omega)C_r(j\omega)$ is plotted in Fig. 10.2-6, together with the constraints. Actually, the constraints are violated in a very minor way. The two closed-loop transfer function magnitudes obtainable from C and C_r are plotted for comparison in Fig. 10.2-7. The two step responses are virtually indistinguishable in a
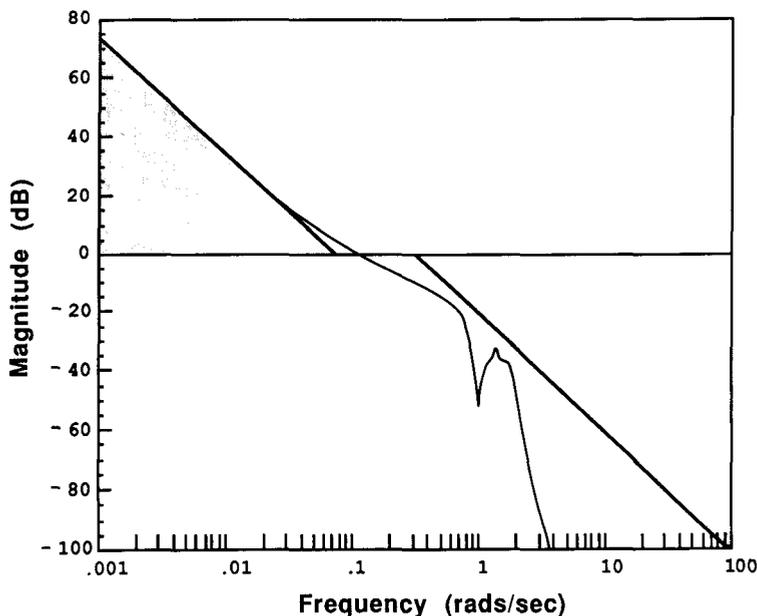
Figure 10.2-4   Bode plot of LQG design gain.
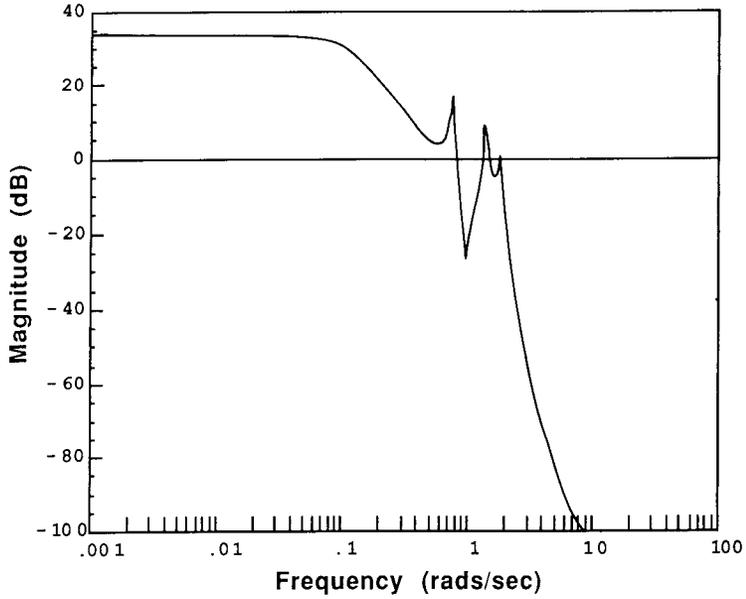
Figure 10.2-5   Weighting function for controller reduction.
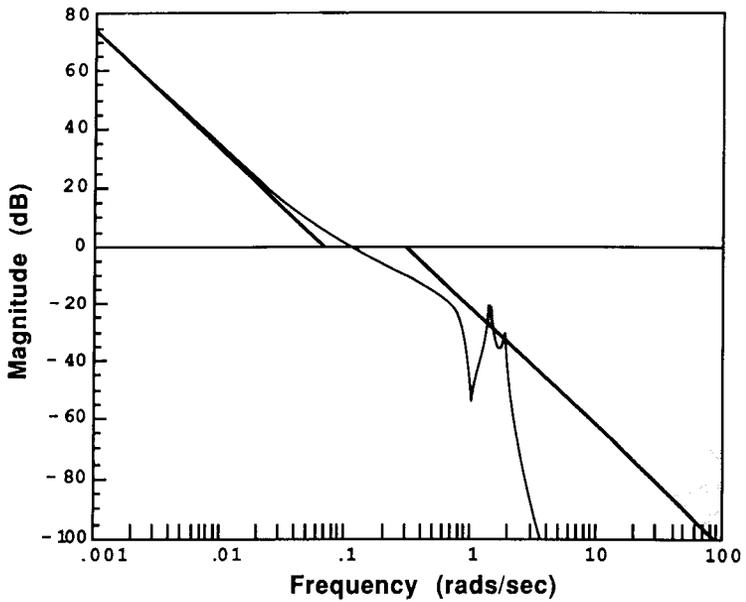
Figure 10.2-6   Bode plot of reduced order design gain.
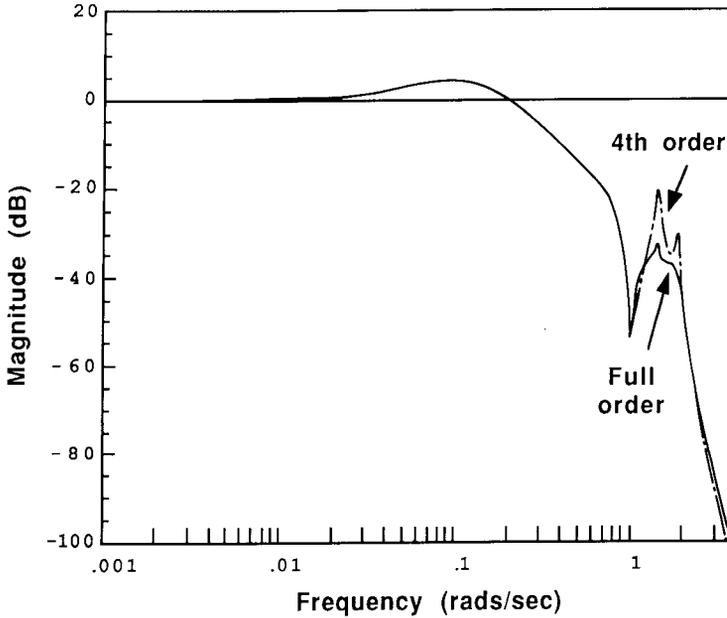
Figure 10.2-7   Closed-loop transfer function comparison.

graphical presentation. Two other measures of comparison come from the gain margin and phase margin:

Gain margin with C, C_r = 8.94 (dB), 8.6 (dB)

Phase margin with C, C_r = 37.53 (deg), 37.88 (deg)

Main points of the section.   An algorithm that is an extension of balanced realization truncation can be used to solve in an approximate manner the frequency-weighted minimization problem associated with controller reduction.

Problem 10.2-1.   Suppose that for a minimal triple $\bar{F}, \bar{G}, \bar{H}$ with Re $\lambda_i(\bar{F}) <$ 0, there holds $\bar{P}\bar{F}' + \bar{F}\bar{P} + \bar{G}\bar{G}' = 0$, $\bar{Q}\bar{F} + \bar{F}'\bar{Q} + \bar{H}\bar{H}' = 0$. Perform a Cholesky decomposition of $\bar{Q}$, that is, $\bar{Q} = R'R$, with R upper triangular. Then $R\bar{P}R'$ is positive definite symmetric. Find an orthogonal U and positive definite diagonal Σ such that $R\bar{P}R' = U\Sigma^2 U'$. Set $T = \Sigma^{-1/2} U'R$. Show that if $F = T\bar{F}T^{-1}$, $G = T\bar{G}$, $H' = \bar{H}'T^{-1}$, there holds $\Sigma F' + F\Sigma + GG' = 0$, $\Sigma F + F'\Sigma + HH' = 0$.

Problem 10.2-2.   Let $W = P(I + CP)^{-1}$ where $P = \bar{H}'(sI - \bar{F})^{-1}\bar{G}$, $C = \bar{K}'(sI - \bar{F} - \bar{G}\bar{K}' - \bar{K}_e\bar{H}')^{-1}\bar{K}_e$, with $\bar{F} + \bar{G}\bar{K}'$, $\bar{F} + \bar{K}_e\bar{H}'$, and $\bar{F} + \bar{G}\bar{K}' + \bar{K}_e\bar{H}'$ all possessing stable eigenvalues. Show that

$$H_w' = (sI - F_w)^{-1}G_w = [\bar{H}' \quad 0]\begin{bmatrix} sI - \bar{F} & -\overline{GK}' \\ \overline{K_e\bar{H}'} & sI - \bar{F} - \overline{GK}' - \bar{K}_e\bar{H}' \end{bmatrix}^{-1}\begin{bmatrix} \bar{G} \\ 0 \end{bmatrix}$$

Consider the Lyapunov equation for the controllability gramian of CW, viz.

$$\bar{P}\begin{bmatrix} (\bar{F} + \overline{GK}' + \bar{K}_e\bar{H}')' & 0 \\ H_w\bar{K}_e' & F_w' \end{bmatrix} + \begin{bmatrix} (\bar{F} + \overline{GK}' + \bar{K}_e\bar{H}') & \bar{K}_e H_w' \\ 0 & F_w \end{bmatrix}\bar{P}$$

$$+ \begin{bmatrix} 0 \\ G_w \end{bmatrix}\begin{bmatrix} 0 & G_w' \end{bmatrix} = 0$$

where $W = H_w'(sI - F_w)^{-1}G_w$. Write the equation as

$$\bar{P}\bar{A}' + \bar{A}\bar{P} + \bar{B}\bar{B}' = 0$$

Consider the transformations $\bar{A} \to \hat{A} = T\bar{A}T^{-1}$, $\bar{B} \to \hat{B} = T\bar{B}$, $\bar{P} \to \hat{P} = T\bar{P}T'$, where

$$T = \begin{bmatrix} I & 0 & I \\ 0 & I & 0 \\ 0 & -I & I \end{bmatrix}$$

Show that $\hat{P}$ has first block row and column equal to zero and that three $n \times n$ matrix equations define the other entries of $\hat{P}$.

Problem 10.2-3. The index associated with output weighting is $J_s' = \max_\omega \bar{\sigma}\{[I + P(j\omega)C(j\omega)]^{-1}P(j\omega)[C(j\omega) - C_r(j\omega)]\}$. Set $W = [I + PC]^{-1}P = P[I + CP]^{-1}$, and let $\{F_w, G_w, H_w\}$ and $\{\bar{F}, \bar{G}, \bar{H}\}$ define minimal realizations of $W(s)$ and $C(s)$. Describe how frequency-weighted balanced truncation can take place.

10.3 APPROACHES TO CONTROLLER REDUCTION VIA FRACTIONAL REPRESENTATIONS

Various other approaches to controller reduction exist, some quite distinct conceptually from those discussed in the last section. For example, methods of [8] attempt to match impulse response and covariance data. In this section, we concentrate on a method that is close in spirit to the scheme of the last two sections. The ideas are drawn from [9–12] in the main.

The key difference is that we represent the controller transfer function matrix in a different manner. Effectively, in the last two sections, each controller transfer function matrix is decomposed additively into a stable part and an unstable part, and the stable part is reduced. In this section, we represent the controller transfer function matrix as a fraction of transfer function matrices that are themselves stable. Then we reduce both the numerator and denominator of this fraction. The reduced numerator and reduced denominator together define a fractional representation of the new controller.

What are the reasons for doing this? First, the method of the past two sections seems overly restrictive, in that the unstable part of the controller is not varied at all. One would imagine that, even if the number of unstable poles of the controller were maintained in a reduction procedure, preservation of their locations and residues is

unlikely to be optimum. Second, the derivation of the stability based indices J_s, J_s' involves use of sufficient conditions for stability. The derivation of the corresponding indices in this section similarly involves sufficient conditions of stability, but modern treatments—see, for example, [13]—suggest that the conditions used in this section are less conservative. Third, the methods of the past two sections all involve nonconstant weighting matrices in the various indices. In contrast, the index derived from spectral considerations in this section involves no such weighting. Such simplicity is appealing, but is by no means a particularly compelling reason for preferring one method over another. Fourth, no one method is universally the best method, so it seems desirable to allow a designer the opportunity to use more than one method.

We need to note one restriction of the methods of this section. In contrast to the situation applying earlier in the chapter, we shall assume that the plant is defined by

$$P(s) = H'(sI - F)^{-1}G \qquad (10.3\text{-}1)$$

with $\{F, G, H\}$ minimal, and the controller is defined by

$$C(s) = K'(sI - F - GK' - K_eH')^{-1}K_e \qquad (10.3\text{-}2)$$

where $F + GK'$ and $F + K_eH'$ both have all eigenvalues in Re $[s] < 0$.

It is possible, as we have seen earlier, to give fractional representations of the plant, as follows:

$$P(s) = H'(sI - F - GK')^{-1}G[I + K'(sI - F - GK')^{-1}G]^{-1}$$
$$= B_R(s)A_R^{-1}(s) \qquad (10.3\text{-}3)$$

$$P(s) = [I + H'(sI - F - K_eH')^{-1}K_e]^{-1}H'(sI - F - K_eH')^{-1}G$$
$$= A_L^{-1}(s)B_L(s) \qquad (10.3\text{-}4)$$

Similarly, it is possible to write

$$C(s) = K'(sI - F - GK')^{-1}K_e[I - H'(sI - F - GK')^{-1}K_e]^{-1}$$
$$= Y_R(s)X_R^{-1}(s) \qquad (10.3\text{-}5)$$

$$C(s) = [I - K'(sI - F - K_eH')^{-1}G]^{-1}K'(sI - F - K_eH')^{-1}K_e$$
$$= X_L^{-1}(s)Y_L(s) \qquad (10.3\text{-}6)$$

As noted in Appendix B, (10.3-3) and (10.3-5) are termed right fractions and (10.3-4) and (10.3-6) are left fractions. Note that each denominator and numerator is stable, because of the eigenvalue restrictions on $F + GK'$ and $F + K_eH'$.

Derivation of a noise-induced index.   Consider the arrangement of Fig. 10.3-1 in which the controller with transfer function matrix $K'(sI - F - GK' - K_eH')^{-1}K_e$, is depicted in a certain way. In fact, the controller is depicted as a feedback system with an open-loop representation

Figure 10.3-1 Controller-plant interconnection exhibiting controller as a one-input, two-output system with one output feedback.

$$\begin{bmatrix} u \\ z \end{bmatrix} = \begin{bmatrix} K'(sI - F - GK')^{-1}K_e \\ H'(sI - F - GK')^{-1}K_e \end{bmatrix} v \tag{10.3-7}$$

The input to the controller is $r - y$, and $v = z + (r - y)$. This way of thinking about the controller is reflected in the fractional representation (10.3-5).

Now suppose that K_e is an optimal Kalman filter gain. This means that the signal v, which is the innovation signal as described in Chapter 7, has a white spectrum; that is, for a positive definite $\hat{R}$,

$$E[v(t)v'(s)] = \hat{R}\delta(t - s) \qquad \Phi_{vv}(j\omega) = \hat{R} \tag{10.3-8}$$

Now if we think of the controller as being defined by (10.3-7) together with an interconnection rule [viz., generate v by $v = (r - y) + z$], the reduction problem becomes one of finding a stable pair $\hat{X}_R(s)$, $\hat{Y}_R(s)$ such that

$$J_n = \max_{\omega} \left\| \begin{bmatrix} Y_e(j\omega) - \hat{Y}_R(j\omega) \\ X_R(j\omega) - \hat{X}_R(j\omega) \end{bmatrix} \hat{R}^{1/2} \right\| \tag{10.3-9}$$

is minimized, and such that $[\hat{Y}_R'(s) \quad \hat{X}_R'(s)]'$ has prescribed order. With $A = F + GK'$, this is simply a matter of approximating

$$V(s) \triangleq \begin{bmatrix} K' \\ H' \end{bmatrix}(sI - A)^{-1}K_e\hat{R}^{1/2} \tag{10.3-10}$$

[derived from (10.3-7) with weighting introduced] by a lower-order strictly proper transfer function matrix $\hat{V}(s)$. With $\hat{V}(s) = [\hat{V}_1(s) \quad \hat{V}_2(s)]'$ there follows $\hat{X}_R(s) = I - \hat{V}_2(s)$, $\hat{Y}_R(s) = \hat{V}_1(s)$ and $C_r(s) = \hat{Y}_R(s)\hat{X}_R^{-1}(s)$. Note that the order of $C_r(s)$ will be identical with the order of $\hat{V}(s)$.

Of course, balanced realization truncation provides a very convenient tool for obtaining $\hat{V}(s)$ from $V(s)$, but the resulting $\hat{V}(s)$ will not in general minimize J_n.

There are several further points to note. First, knowledge of the plant is included, albeit implicitly, in the reduction procedure—after all, one could not form the fractional representation (10.3-5) knowing $C(s)$ as a transfer function, and nothing about the plant. Thus the dictum that the plant needs to be taken into account in the reduction process is not violated.

Second, in contrast to the scheme of the last section, the number of unstable poles of $C_r(s)$ may not be the same as the number of unstable poles of $C(s)$—these numbers being determined by the number of right half-plane zeros of $I - \hat{V}_2(s)$ and $I - V_2(s)$, respectively. (Of course, if $\hat{V}_2$ and V_2 are close, the numbers are likely to be the same). If the numbers are different, this does not of itself imply that $C_r(s)$ will not be stabilizing.

Third, the replacement of C by C_r will cause the signal v to lose its white character; thus one of the premises behind the approximation becomes invalid in the process of doing that approximation. Naturally, if C_r is close to C there should be little problem—but there is a warning implicit in the observation.

Fourth, even if v is a scalar, there is a scaling possibility open. Increase of the weighting on u in (10.3-7) would imply a view of the designer that it is more important to get the plant input accurately approximated than an internal feedback in the compensator; but note that, precisely because of this feedback, inaccuracies in z will produce inaccuracies in u, even if $K'[sI - (F + GK')]^{-1}K_e$ could be approximated with zero error. The transfer function from v to u is associated with controller zeros, and that from v to z with controller poles. So variation in scaling could be regarded as change of emphasis on the approximation of zeros as opposed to poles. Generally speaking, unstable poles and zeros need to be well approximated.

Example.   To illustrate the idea, we use the example of the last section. With the values of F, g, h, k and k_e presented earlier, we work with the single-input, two-output transfer function

$$V_1(s) = \begin{bmatrix} k' \\ h' \end{bmatrix}[sI - (F + gk')]^{-1}k_e$$

As a first attempt at scaling, let us approximate not $V_1(s)$ but

$$V_\alpha(s) = \begin{bmatrix} \alpha k' \\ h' \end{bmatrix}[sI - (F + gk')]^{-1}k_e$$

where

$$\alpha = \frac{\max_\omega |1 - h'[j\omega I - (F + gk')]^{-1}k_e|}{\max_\omega |k'[j\omega I - (F + gk')]^{-1}k_e|}$$

This is simply an attempt to give equal weighting after scaling to approximation of the two signal paths within the compensator. As it turns out, α is approximately 30. The choice $\alpha = 30$ with balanced truncations of $V_{30}(s)$ yields a third-order approximation of the eighth-order original as

$$\hat{V}_{30}(s) = \begin{bmatrix} 30\hat{k}' \\ \hat{h}' \end{bmatrix}(sI - \hat{A})^{-1}\hat{k}_e$$

where

$$\hat{k} = \begin{bmatrix} 0.0373 \\ 0.0083 \\ -0.0191 \end{bmatrix} \qquad \hat{h} = \begin{bmatrix} -0.7126 \\ 0.7079 \\ -0.0160 \end{bmatrix} \qquad \hat{k}_e = \begin{bmatrix} 1.3255 \\ 0.7504 \\ 0.5737 \end{bmatrix}$$

$$\hat{A} = \begin{bmatrix} -0.0681 & 0.0627 & 0.0556 \\ -0.1042 & -0.0504 & 0.0490 \\ -0.0647 & -0.0888 & -0.1220 \end{bmatrix}$$

Of course, what is more important is the controller transfer function

$$C_r(s) = \frac{\hat{k}'(sI - \hat{A})^{-1}\hat{k}_e}{1 - \hat{h}'(sI - \hat{A})^{-1}\hat{k}_e}$$

$$= \frac{0.0446s^2 + 0.0133s + 0.00049}{s^3 + 0.6630s^2 + 0.2270s + 0.02482}$$

The loop gain with the low-order controller appears in Fig. 10.3-2. Figure 10.3-3 compares the closed-loop transfer functions resulting from use of $C(s)$ and $C_r(s)$, while Fig. 10.3-4 compares the corresponding step responses. Gain and phase margin comparisons are

Gain margin with C, C_r = 8.94 dB, 9.32 dB

Phase margin with C, C_r = 37.53 deg, 35.57 deg

Derivation of a stability-criterion-induced index.   In the first section of this chapter, we argued that one way of capturing the controller reduction problem was to formulate a frequency-weighted transfer function matrix approximation problem. Different weights could be advanced, depending on the
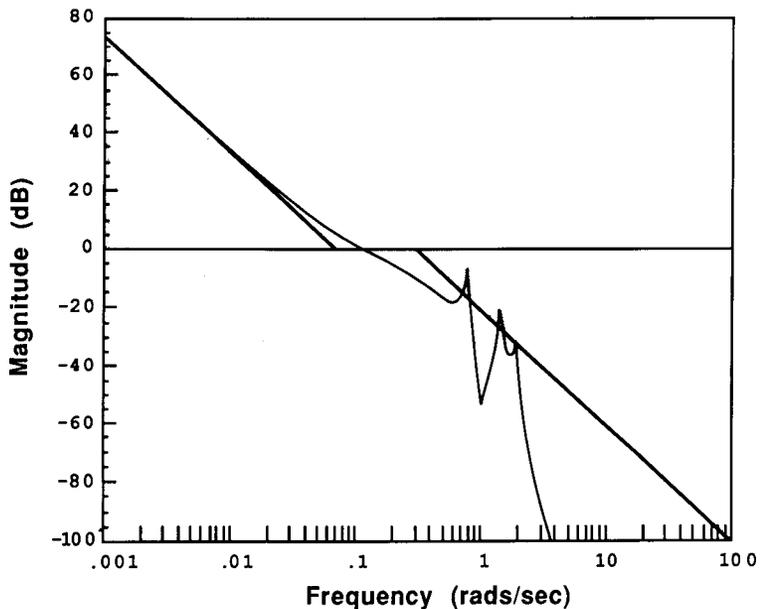
Figure 10.3-2   Loop gain with third-order controller.
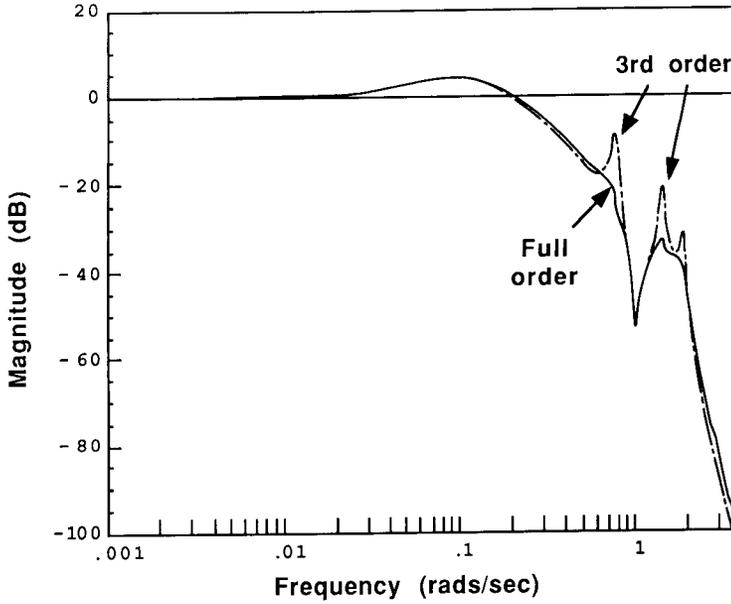
Figure 10.3-3   Closed-loop transfer function magnitude.
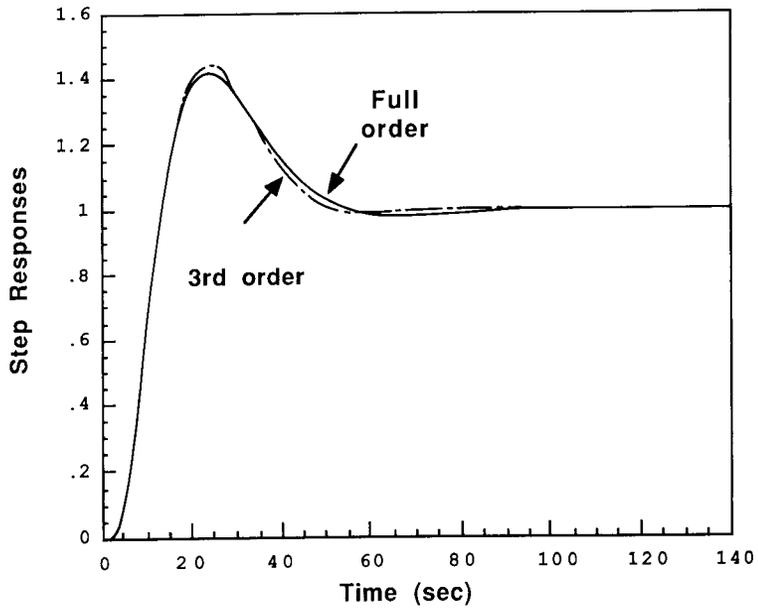
Figure 10.3-4   Step response comparison.

particular aspect of closed-loop performance to which attention was most directed. In particular, one could obtain weights induced by noise considerations, and by considerations of stability. The same is, of course, true when we work with fractional representations of controllers. Having just considered the noise-induced weight, we turn now to consider stability-induced weights.

Lying behind the construction of an associated index is a so-called Bezout equation. In terms of the definitions in (10.3-3) through (10.3-6) of A_L, B_L, A_R, B_R, X_L, Y_L, X_R, and Y_R, one can write the following; see [14] and also Appendix B.

$$\begin{bmatrix} X_L & Y_L \\ -B_L & A_L \end{bmatrix}\begin{bmatrix} A_R & -Y_R \\ B_R & X_R \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \tag{10.3-11}$$

The 1-2 and 2-1 block terms yield $X_L Y_R = Y_L X_R$ and $B_L A_R = A_L B_R$; the first is immediate from the pair (10.3-5) and (10.3-6), and the second is immediate from the pair (10.3-3) and (10.3-4). The identities for the 1-1 and 2-2 terms follow by direct calculation; see Problem 10.3-1.

Now let us return to the controller as depicted in Fig. 10.3-1 and represented by a right fraction. We shall tie together the Bezout identity and a stability robustness result. We can redraw Fig. 10.3-1 as in Fig. 10.3-5.

We can think of there being a new "plant" $\hat{P}(s)$ and "controller" $\hat{C}(s)$ defined by

$$\hat{P}(s) = [P(s) \quad -I] = [H'(sI - F)^{-1}G \quad -I] \tag{10.3-12}$$

$$\hat{C}(s) = \begin{bmatrix} K' \\ H' \end{bmatrix}(sI - F - GK')^{-1}K_e \tag{10.3-13}$$

Now if we take the stability weighting point of view set out in Section 10.1, we should seek $\hat{C}_r(s)$ to minimize

$$J_s' = \max_\omega \bar{\sigma}\{\hat{P}(j\omega)[I + \hat{C}(j\omega)\hat{P}(j\omega)]^{-1}[\hat{C}(j\omega) - \hat{C}_r(s)]\} \tag{10.3-14}$$

Now observe that we can easily express $\hat{P}$, $\hat{C}$ in terms of A_L, B_L, X_R, and Y_R. Then

$$\begin{aligned}
\hat{P}[I + \hat{C}\hat{P}]^{-1} &= [I + \hat{P}\hat{C}]^{-1}\hat{P} \\
&= \left\{I + [A_L^{-1}B_L \quad -I]\begin{bmatrix} Y_R \\ -X_R + I \end{bmatrix}\right\}^{-1}[A_L^{-1}B_L \quad -I] \\
&= \left\{A_L + [B_L \quad -A_L]\begin{bmatrix} Y_R \\ -X_R + I \end{bmatrix}\right\}^{-1}[B_L \quad -A_L] \\
&= [B_L \quad -A_L]
\end{aligned} \tag{10.3-15}$$
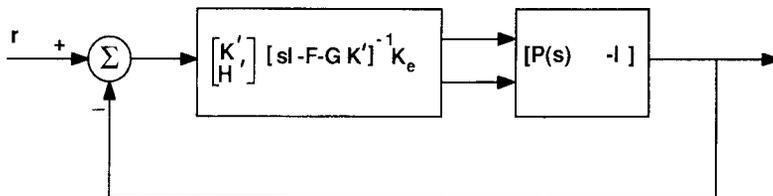
Figure 10.3-5 Redrawing of controller-plant interconnection to allow derivation of stability robustness index.

(The last line follows when the Bezout identity is used.) Now with the full-order $\hat{C}(s)$ given by

$$\hat{C}(s) = \begin{bmatrix} Y_R(s) \\ -X_R(s) + I \end{bmatrix}$$

the reduced order $\hat{C}(s)$ is given by

$$\hat{C}_r(s) = \begin{bmatrix} \hat{Y}_R(s) \\ -\hat{X}_R(s) + I \end{bmatrix}$$

and

$$\hat{C}(s) - \hat{C}_r(s) = \begin{bmatrix} Y_R(s) - \hat{Y}_R(s) \\ -[X_R(s) - \hat{X}_R(s)] \end{bmatrix} \tag{10.3-16}$$

The index J_s' is accordingly

$$J_s' = \max_{\omega} \overline{\sigma} \left\{ \begin{bmatrix} -B_L(j\omega) & A_L(j\omega) \end{bmatrix} \begin{bmatrix} -Y_R(j\omega) + \hat{Y}_R(j\omega) \\ X_R(j\omega) - \hat{X}_R(j\omega) \end{bmatrix} \right\} \tag{10.3-17}$$

The parallel with the formation of the 2-2 term of the Bezout identity should be obvious: the new controller $\hat{Y}_R(j\omega)\hat{X}_R^{-1}(j\omega)$ should leave the 2-2 terms of the Bezout identity as little changed from I as possible when $\hat{X}_R$, $\hat{Y}_R$ replaces X_R, Y_R. This view of stability robustness is actually set out in [13], with further conceptual and intuitive backing.

In place of the index J_s', we could have worked with the index J_s, but the conceptual insight is not so attractive.

Summarizing to this point, we have argued that with plant and controller fractional representation as in (10.3-3) through (10.3-6), robust stability arguments suggest that we seek to approximate

$$\begin{bmatrix} Y_R(s) \\ -X_R(s) + I \end{bmatrix} = \begin{bmatrix} K' \\ H' \end{bmatrix} (sI - F - GK')^{-1} K_e$$

using a weighting (on the left), of

$$\begin{bmatrix} -B_L & A_L \end{bmatrix} = \begin{bmatrix} -H'(sI - F - K_eH')^{-1}G & I + H'(sI - F - K_eH')^{-1}K_e \end{bmatrix}$$

A dual construction is possible. Start with the description of the controller via (10.3-6). Figure 10.3-6 illustrates the controller depicted as a two-input, one-output system with feedback. The associated transfer function matrix is

$$\hat{C}(s) = K'(sI - F - K_eH')^{-1} \begin{bmatrix} -G & K_e \end{bmatrix} = \begin{bmatrix} X_L(s) - I & Y_L(s) \end{bmatrix} \tag{10.3-18}$$

Define also

$$\hat{P}(s) = \begin{bmatrix} I \\ P(s) \end{bmatrix} \tag{10.3-19}$$

We can redraw the arrangement of Fig. 10.3-6 as in Fig. 10.3-7, in order to write down an index reflecting stability robustness. With
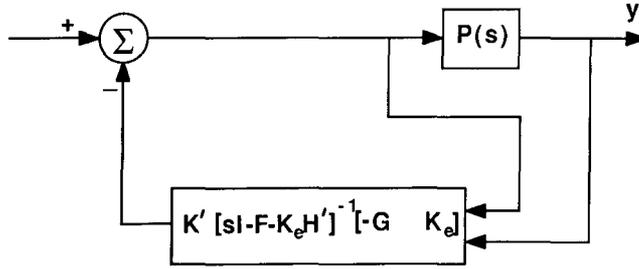
Figure 10.3-6   Representation of controller in dual fashion to Fig. 10.3-5.

$$J_s = \max_{\omega} \overline{\sigma} \{ [\hat{C}(j\omega) - \hat{C}_r(j\omega)] \hat{P}(j\omega)[I + \hat{C}(j\omega)\hat{P}(j\omega)]^{-1} \qquad (10.3\text{-}20)$$

we can show (see Problem 10.3-2) that

$$\hat{P}(s)[I + \hat{C}(s)\hat{P}(s)]^{-1} = \begin{bmatrix} A_R \\ B_R \end{bmatrix} \qquad (10.3\text{-}21)$$

Consequently, J_s can be written as

$$J_s = \max_{\omega} \overline{\sigma} \left\{ [X_L(j\omega) - \hat{X}_L(j\omega) \quad Y_L(j\omega) - \hat{Y}_L(j\omega)] \begin{bmatrix} A_R(j\omega) \\ B_R(j\omega) \end{bmatrix} \right\} \qquad (10.3\text{-}22)$$

Evidently, this index looks at what happens to the 1-1 block of the Bezout identity (10.3-11) when X_L, Y_L are approximated.

Of course, as in the last section, we do not necessarily seek to minimize J_s' in (10.3-14) or J_s in (10.3-20) exactly. Rather, we use the form of the index to set up a frequency-weighted balanced realization and then truncate that realization. Of course, all transfer function matrices that arise are stable. If we work with the $\hat{C}$, $\hat{P}$ of (10.3-18) and (10.3-19), the weighted transfer function matrix $\hat{C}\hat{P}[I + \hat{C}\hat{P}]^{-1}$ has order $2n \times 2n$ and so apparently a $2n \times 2n$ Lyapunov matrix equation has to be solved for the controllability gramian. However, it is not hard to check that actually, this $2n \times 2n$ equation is equivalent to a single $n \times n$ equation. (See Problem 10.3-3.)

   Example.   When we apply two procedures (using a right and left co-prime factorization with weighting) to the example treated to this point, the results are far more attractive with right factorization than with left. The frequency-weighted Hankel singular values for the two cases are
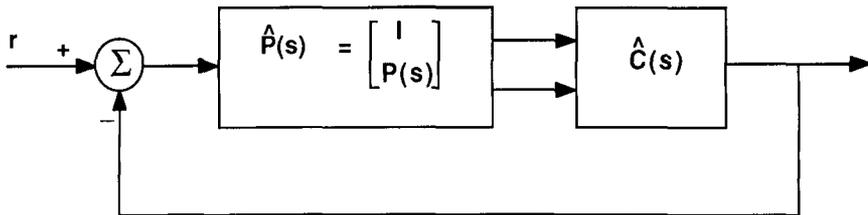
Figure 10.3-7   Redrawing of Fig. 10.3-6.

Right	Left
1.095×10^{1}	1.936
5.295×10^{-1}	1.735
1.240×10^{-2}	1.270
6.035×10^{-3}	1.180
5.887×10^{-4}	9.290×10^{-1}
3.298×10^{-4}	3.417×10^{-1}
1.231×10^{-4}	3.188×10^{-1}
1.073×10^{-4}	9.851×10^{-2}

A reduction to order two proves possible with the right factorization, while even a reduction to order seven with the left factorization produces instability. This sort of behavior is suggested by the weighted Hankel singular values.

Figures 10.3-8, 10.3-9, and 10.3-10 depict for a second-order controller the loop gain, closed-loop transfer function, and step response, with the latter two showing the full-order case for comparison.

The gain margin and phase margin with the reduced order controller are, respectively, 10.46 dB and 37.42 deg, with the corresponding figures for the full-order controller being 8.94 dB and 37.53 deg.

Approximating other variables.   There is a straightforward embellishment of the reduction procedure associated with (10.3-18); see also Fig. 10.3-6. Suppose that it is desired that a certain linear functional or collection of
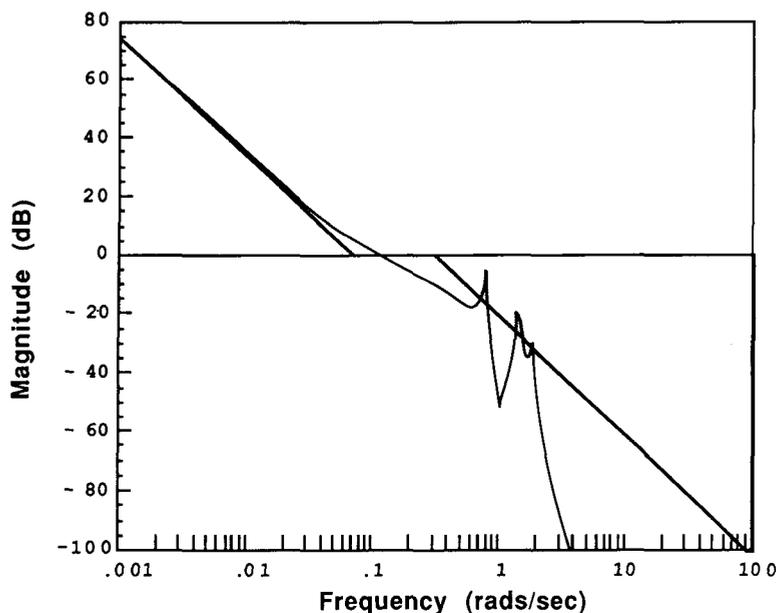
Figure 10.3-8   Loop gain with second-order controller.
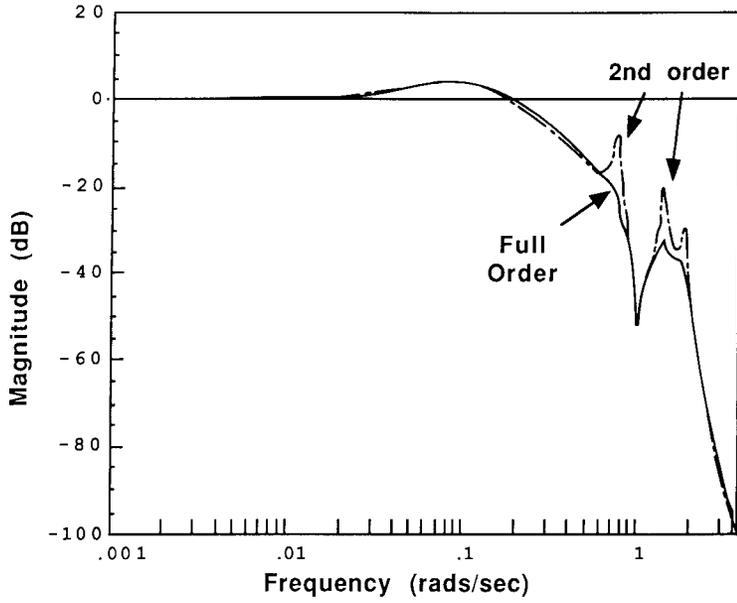
Figure 10.3-9   Closed-loop transfer function magnitude.
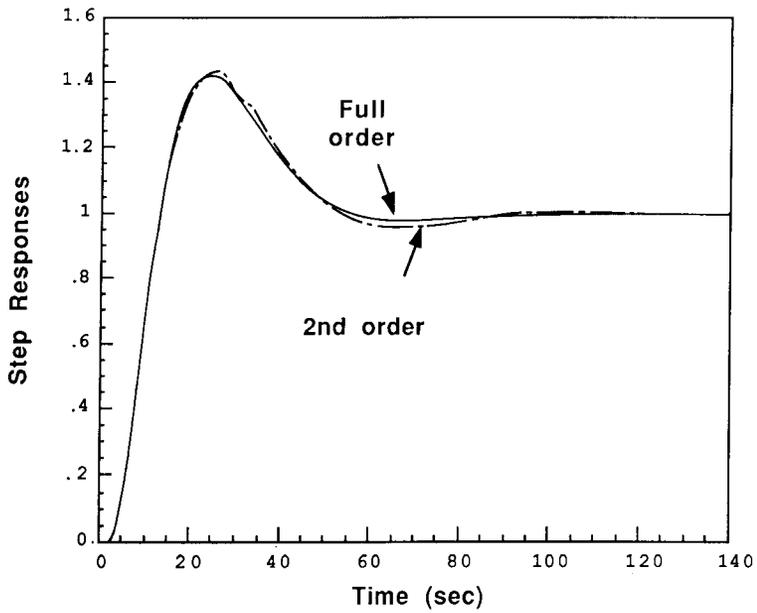
Figure 10.3-10   Step response comparison.

linear functionals of the plant, $L'x$, be little affected by the controller approximation process. Note that $L'x$ may well include or be identical with $y = H'x$. Now the controller is a combination of an estimator and a state feedback law and $L'x$ can be asymptotically recovered by drawing an additional output out of the estimator part of the controller; see Fig. 10.3-11. This suggests that there may be value in approximating not

$$K'(sI - F - K_eH')^{-1}[-G \quad K_e]$$

but

$$\begin{bmatrix} K' \\ L' \end{bmatrix}(sI - F - K_eH')^{-1}[-G \quad K_e]$$

When $L' = H'$, we observe that this transfer function matrix is precisely

$$\begin{bmatrix} X_L(s) - I & Y_L(s) \\ -B_L(s) & A_L(s) - I \end{bmatrix}$$

Apart from the $-I$ terms, which reappear in the approximation, this matrix is precisely one of the two product matrices in the Bezout identity. By introducing a scaling constant α, and working with

$$\begin{bmatrix} K' \\ \alpha L' \end{bmatrix}(sI - F - K_eH')^{-1}[-G \quad K_e]$$

one can examine a range of situations with $\alpha = 0$ corresponding to a stability-oriented reduction procedure, and $\alpha \gg 1$ corresponding to a performance-oriented reduction procedure. Some examples of this procedure can be found in [10].

Main points of the section.   Alternative procedures for controller reduction are obtained when one works with fractional representations of controllers, and the controllers are obtained by combining an estimator with state estimate feedback. Taking a viewpoint suggested by the whiteness of the innovations process in a Kalman filter, one is led to a reduction problem without weighting. Otherwise,
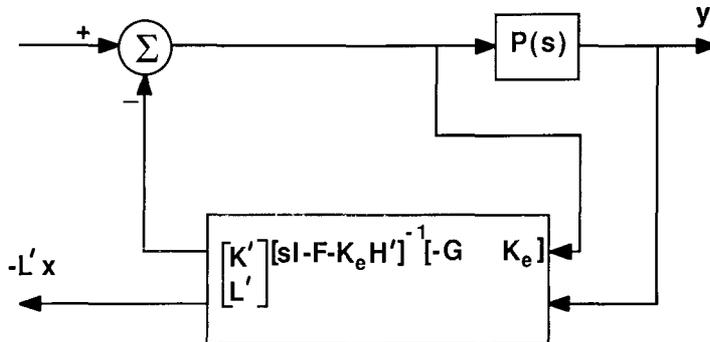
Figure 10.3-11   Introduction of further signals that should be well approximated.

taking a robust stability viewpoint leads to weights which appear in the so-called Bezout equation. This relates fractional representations of the plant and controller.

Problem 10.3-1.   Define

$$\begin{bmatrix} X_L(s) & Y_L(s) \\ -B_L(s) & A_L(s) \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} + \begin{bmatrix} K' \\ H' \end{bmatrix} (sI - F - K_eH')^{-1}[-G \quad K_e]$$

$$\begin{bmatrix} A_R(s) & -Y_R(s) \\ B_R(s) & X_R(s) \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} + \begin{bmatrix} K' \\ H' \end{bmatrix} (sI - F - GK')^{-1}[G \quad -K_e]$$

Show that these two matrices are inverses of one another.

Problem   10.3.2.   Suppose   that $P(s) = B_R(s)A_R^{-1}(s) = A_L^{-1}(s)B_L(s)$ and $C(s) = X_L^{-1}(s)Y_L(s) = Y_R(s)X_R^{-1}(s)$, and that the Bezout identity

$$\begin{bmatrix} X_L & Y_L \\ -B_L & A_L \end{bmatrix}\begin{bmatrix} A_R & -Y_R \\ B_R & X_R \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$$

holds. Suppose further that

$$\hat{C}(s) = [X_L(s) - I \quad Y_L(s)] \qquad \hat{P}(s) = \begin{bmatrix} I \\ P(s) \end{bmatrix}$$

Show that

$$\hat{P}[I + \hat{C}\hat{P}]^{-1} = \begin{bmatrix} A_R \\ B_R \end{bmatrix}$$

Problem 10.3-3.   Suppose that

$$\hat{C}(s) = K'(sI - F - K_eH')^{-1}[-G \quad K_e]$$

$$\hat{W}(s) = \hat{P}(s)[I + \hat{C}(s)\hat{P}(s)]^{-1} = \begin{bmatrix} I + K'(sI - F - GK')^{-1}G \\ H'(sI - F - GK')^{-1}G \end{bmatrix}$$

A realization of $\hat{C}(s)\hat{W}(s)$ is provided by

$$\hat{F} = \begin{bmatrix} F + K_eH' & K_eH' - GK' \\ 0 & F + GK' \end{bmatrix} \qquad \hat{G} = \begin{bmatrix} -G \\ G \end{bmatrix}$$

$$\hat{H}' = [K' \quad 0]$$

Suppose that $\hat{\Pi}$ solves $\hat{\Pi}\hat{F}' + \hat{F}\hat{\Pi} + \hat{G}\hat{G}' = 0$. Show that $\hat{\Pi}$ is obtained by solving an $n \times n$ Lyapunov equation. [*Hint:* With $T = \begin{bmatrix} I & 0 \\ -I & I \end{bmatrix}$, replace $\hat{F}$ by $T^{-1}\hat{F}T$, and so forth.]

Problem 10.3-4.   Consider the left fraction representation of the usual compensator $C(s) = [I - K'(sI - F - K_eH')^{-1} G]^{-1} K'(sI - F - K_eH')^{-1} K_e = X_L^{-1}(s)Y_L(s)$. Suppose that K_e is designed by a loop recovery process with measurement noise covariance matrix $= I$ and that $H'(sI - F)^{-1}G$ is minimum phase and nonsingular. Show that as the process noise covariance $\rho GG'$ changes with $\rho \to \infty$,

$$X_L(s) \to I$$

$$Y_L(s) \to -K'(sI - F)^{-1}G[H'(sI - F)^{-1}G]^{-1}$$

Notice that

$$[X_L(j\omega) \quad Y_L(j\omega)]\begin{bmatrix} A_R(j\omega) \\ B_R(j\omega) \end{bmatrix} \to -[K'(j\omega I - F)^{-1}GP^{-1}(j\omega)]B_R(j\omega)$$

and show that the procedure of the previous section leads to a problem of reducing $K'(j\omega I - F)^{-1}GP^{-1}(j\omega)$ weighted by $B_R(j\omega)$ also.

10.4 DIRECT DESIGN OF LOW-ORDER CONTROLLERS

As noted earlier, it is possible to contemplate a search *ab initio* for a low-order controller minimizing a quadratic performance index. In this section, we shall introduce the reader to these ideas. We shall begin by considering a very simple problem, that of the determination of a constant feedback gain from the output to the input of a system; see [15].

We consider a time-invariant system

$$\dot{x} = Fx + Gu \qquad y = H'x \tag{10.4-1}$$

where the control is constrained to be of the form

$$u = K_0'y = K_0'H'x \tag{10.4-2}$$

with performance index

$$V(x(0), u(\cdot)) = \int_0^\infty (u'Ru + x'Qx) \, dt \tag{10.4-3}$$

The performance index value can readily be determined provided (10.4-2) is stabilizing. It is

$$V(x(0), u(\cdot)) = x'(0)\bar{P}x(0) \tag{10.4-4}$$

where

$$\bar{P}(F + GK_0'H') + (F' + HK_0G')\bar{P} + HK_0RK_0'H' + Q = 0 \tag{10.4-5}$$

In general, the K_0 that minimizes (10.4-4) will depend on $x(0)$. Clearly, we want a single gain as the solution to an optimization problem. So, a little arbitrarily, we adopt as the index tr $(\bar{P})$, and seek K_0 to minimize tr $(\bar{P})$. It is obvious that a necessary condition for an optimum K_0 to exist is that there exist a stabilizing feedback law for (10.4-1). (At this stage, only rather complicated algorithms not discussed here are available for answering the existence question; see [16, 17]).

We shall now derive necessary conditions on K_0 to achieve optimality. Suppose that in (10.4-5), K_0 is perturbed to $(K_0 + \delta K_0)$, causing a perturbation of $\bar{P}$ to

$(\bar{P} + \delta\bar{P})$. Writing down a perturbed version of (10.4-5), subtracting from it (10.4-5), and neglecting second-order terms yields

$$(\delta\bar{P})(F + GK_0'H') + (F + GK_0'H')'(\delta\bar{P}) + (\bar{P}G + HK_0R)(\delta K_0)'H'$$
$$+ H(\delta K_0)(G'\bar{P} + RK_0'H') = 0$$

whence

$$\delta\bar{P} = \int_0^\infty \exp[(F + GK_0'H')'t]\{(\bar{P}G + HK_0R)(\delta K_0)'H' + H(\delta K_0)(G'\bar{P} + RK_0'H')\}$$
$$\times \exp[(F + GK_0'H')t]\, dt$$

Taking the trace, and using some of its simple properties yields

$$\mathrm{tr}\,(\delta\bar{P}) = 2\,\mathrm{tr}\int_0^\infty (\delta K_0)(G'\bar{P} + RK_0'H')\exp[(F + GK_0'H')t]$$
$$\times \exp[(F + GK_0'H')'t]H\, dt$$
$$= 2\,\mathrm{tr}\,[(\delta K_0)(G'\bar{P} + RK_0'H')MH] \tag{10.4-6}$$

where

$$M(F + GK_0'H')' + (F + GK_0'H')M + I = 0 \tag{10.4-7}$$

A necessary condition for $\mathrm{tr}\,(\bar{P})$ to be minimized for all δK_0 is then

$$H'M(\bar{P}G + HK_0R) = 0 \tag{10.4-8}$$

Equations (10.4-5), (10.4-7), and (10.4-8) together implicitly define the optimal control law. The number of unknowns is equal to the number of (independent) scalar equations, but the equations are nonlinear. The structure of the solution set is far from clear, and in particular, it is not certain whether there may be a multiplicity of solutions. The equations must be satisfied by every local minimum or local maximum, and so that extraction of the correct solution is by no means straight-forward. Suggestions for solution methods are contained in [15] and the predecessor of this book, [18].

A characteristic of the solution procedures is that they require the initial identification of a stabilizing gain; an iterative procedure is then used which recovers a sequence of gains. If at any stage one of these gains fails to be stabilizing, the algorithm fails, and no simple scheme of recovery is available. A procedure to avoid much of this difficulty is due to Ly, [19]. It is based on the following observation. The two matrices $\bar{P}$ and M provided K_0 is stabilizing are given by

$$\bar{P} = \int_0^\infty \exp[(F' + HK_0G')t][HK_0RK_0'H + Q]\exp[(F + GK_0'H')t]\, dt \tag{10.4-9}$$

$$M = \int_0^\infty \exp[(F + GK_0'H')t]\exp[(F' + HK_0G')'t]\, dt \tag{10.4-10}$$

In case K_0 is not stabilizing, the upper limit of ∞ is replaced by T for some fixed T. The formula (10.4-6) is then used to iterate on the current value of K_0 to reduce trace $(\delta \bar{P})$; thus a gradient search is used (which should terminate in a local minimum) rather than a procedure based on seeking directly a solution to the nonlinear equation satisfied by the optimum K_0. The value of T can be increased during the successive iterations.

Clearly, a comparatively simple problem statement has created a computationally complicated set of equations to study. The situation is even more formidable when we consider dynamic controllers.

Optimal fixed-order dynamic compensators.
We shall outline some of the ideas of [2]; these can be regarded as significant extensions of the theory just presented. We are given an n-dimensional stochastic system

$$\dot{x} = Fx + Gu + v \tag{10.4-11a}$$

$$y = H'x + w \tag{10.4-11b}$$

with $E[v(t)v'(s)] = \hat{Q}\delta(t-s)$, $E[w(t)w'(s)] = \hat{R}\delta(t-s)$. We seek a fixed order time-invariant dynamic compensator of order n_c

$$\dot{x}_c = A_c x_c + B_c y \tag{10.4-12a}$$

$$u = C_c x_c \tag{10.4-12b}$$

such that the closed-loop system is stable, and such that the following performance index is minimized:

$$J(A_c, B_c, C_c) = \lim_{t \to \infty} E[x'(t)Qx(t) + u'(t)Ru(t)] \tag{10.4-13}$$

We shall not derive, but simply state here the equations yielding necessary conditions on A_c, B_c, C_c, referring the reader to [2] for a derivation. Of course, the basic idea behind the derivations is like that used for the constant output feedback problem. To present the equations, we define the concept of a $(\Lambda, \Theta, \Gamma)$ factorization of a product of two $n \times n$ nonnegative definite matrices M and M_e. Suppose the product has rank n_c. Then there exist Λ, Γ both $n_c \times n$ and an $n_c \times n_c$ matrix Θ with diagonal Jordan form and positive eigenvalues such that

$$MM_e = \Lambda'\Theta\Gamma \qquad \Gamma\Lambda' = I_{n_c} \tag{10.4-14}$$

We also define the (oblique) projection operator

$$\tau = \Lambda'\Gamma \tag{10.4-15}$$

(Notice that $\tau^2 = \tau$, but in general, $\tau \neq \tau'$). Then if (A_c, B_c, C_c) minimize J with a stable closed-loop system, there exist $n \times n$ nonnegative definite symmetric matrices P, P_e such that

$$A_c = \Gamma(F - GR^{-1}G'P - P_e H\hat{R}^{-1}H')\Lambda' \tag{10.4-16a}$$

$$B_c = \Gamma P_e H\hat{R}^{-1} \tag{10.4-16b}$$

$$C_c = -\hat{R}^{-1}G'P\Lambda' \qquad (10.4\text{-}16c)$$

while P, P_e, M, and M_e satisfy equations given below. The similarity of (10.4-16) with the theory of full-order compensators is striking; indeed, if we had $\Gamma = \Lambda = I$, and if P, P_e had their earlier significance, we would exactly recover the earlier theory.

Now let us consider the equations satisfied by P, P_e, M, and M_e. These are

$$P(F - GR^{-1}G'P\tau) + (F - GR^{-1}G'P\tau)'P$$
$$+ \tau'PGR^{-1}G'P\tau + Q = 0 \qquad (10.4\text{-}17a)$$

$$P_e(F - \tau P_e H\hat{R}^{-1}H')' + (F - \tau P_e H\hat{R}^{-1}H')P_e$$
$$+ \tau P_e H\hat{R}^{-1}H'P_e\tau' + \hat{Q} = 0 \qquad (10.4\text{-}17b)$$

$$\tau[M(F - GR^{-1}G'P)' + (F - GR^{-1}G'P)M + P_e H\hat{R}^{-1}H'P_e] = 0 \qquad (10.4\text{-}17c)$$

$$[M_e(F - P_e H\hat{R}^{-1}H') + (F - P_e H\hat{R}^{-1}H')'M_e + PGR^{-1}G'P]\tau = 0 \qquad (10.4\text{-}17d)$$

In the special case that $n_c = n$, one can show that $\tau = I$, and P and P_e can be identified with matrices arising in the standard LQG problem.

As noted in [2], the above remarks do *not* address several issues, including: (1) conditions for the existence of a stabilizing compensator of a certain order; (2) sufficiency conditions, that is, conditions additional to the above which single out the global minimum; (3) numerical algorithms.

As it turns out, the number of distinct solutions of equations (10.4-14), (10.4-15), and (10.4-17) can be enormous. Methods for solving them so as to extract the desired global minimum are being developed, using homotopy theory; see [20]. This reference suggests there can be up to $n!/[(n - n_c)!\, n_c!]$ solutions of the equations, while with homotopy methods, there is either only one solution that is obtained, or with $\alpha = \dim$ (unstable subspace of F) and $\beta = \min$ (dim u, dim y) and with $n_c < \beta$, there are $\binom{\beta - \alpha}{n_c - \alpha}$ candidate solutions, one of which yields the global minimum.

Another hitherto unexplored possibility could involve using as a first iterate to the solution a controller obtained by one of the schemes presented earlier in this chapter. Assuming such methods become sufficiently well developed as to allow their inclusion in widely used design packages, there will then be two very distinct routes to the design of low-order controllers for high order systems. However, the extent to which the robustness ideas associated with the return difference equations and the techniques for Q, R selection can be carried over to the methods of this section unknown. So even with good numerical methods for solving a given problem, performance index selection may still be a major issue.

A number of other constrained controller optimizations can be handled by similar techniques to those above. For example, if one sought a decentralized controller for a multiple-input, multiple-output system, one could evaluate the gradient of a performance index (deterministic as for the first problem considered above, or stochastic as in the second problem) with respect to the elements of the controller.

Naturally, questions of solvability, and determination of a global minimum among many extremes or local minima again arise.

Other types of controller constraints, such as bounds on magnitudes of gains, can be handled in principle by invoking more sophisticated ideas of nonlinear optimization in conjunction with the above.

Main points of the section.   Sets of coupled nonlinear equations can be found that are satisfied by the solutions of a constrained order linear quadratic problem. Extraction of the globally minimizing solutions is a separate difficult task.

Problem 10.4-1.   Show that if H is invertible (H' has a left inverse) in the first problem considered in this section, the solution is essentially equivalent to the normal regulator problem.

Problem 10.4-2.   Show that if $n_c = n$ in the second problem considered in this section, the controller obtained is that predicted by the usual LQG theory.

REFERENCES

[1] D. Gangsaas, et al., "Application of Modern Synthesis to Aircraft Control: Three Case Studies," *IEEE Trans. Auto. Control,* Vol. AC-31 (1986), pp. 995–1104.

[2] D. S. Bernstein and D. C. Hyland, "The Optimal Projection Equations for Fixed-order Dynamic Compensation," *IEEE Trans. Auto. Control,* Vol. AC-29 (1984), pp. 1034–1037.

[3] M. C. Davis, "Factoring the Spectral Matrix," *IEEE Trans. Auto. Control,* Vol. AC-8 (1963), pp. 296–305.

[4] B. D. O. Anderson, K. L. Hitz, and N. Diem, "Recursive Algorithms for Spectral Factorization," *IEEE Trans. Circuits and Systems,* Vol. CAS-21 (1974), pp. 742–750.

[5] K. Glover, "All Optimal Hankel-norm Approximation of Linear Multivariable Systems and their L^{∞}-error bounds," *Int. J. Control,* Vol. 39 (1984), pp. 1115–1193.

[6] D. Enns, "Model Reduction for Control System Design," Ph.D. Thesis, Department of Aeronautics and Astronautics, Stanford University, 1984.

[7] L. Pernebo and L. M. Silverman, "Model Reduction via Balanced State Space Representations," *IEEE Trans. Auto. Control,* Vol. AC-27 (1982), pp. 382–387.

[8] A. Yousuff, and R. E. Skelton, "Controller Reduction by Component Cost Analysis," *IEEE Trans. Auto. Control,* Vol. AC-29, No. 6 (June 1984), pp. 520–530.

[9] Y. Liu and B. D. O. Anderson, "Controller Reduction via Stable Factorization and Balancing," *Int. J. Control,* Vol. 44 (1986), pp. 507–531.

[10] B. D. O. Anderson and Y. Liu, "Controller Reduction: Concepts and Approaches," *Proc. 1987 Amer. Control Conf.,* Minneapolis, pp. 1–9.

[11] Y. Liu, B. D. O. Anderson, and U.-L. Ly, "Coprine Factorization Controller Reduction with Bezout Identity Induced Frequency Weighting," *Automatica,* to appear.

[12] J.B. Moore, U.-L. Ly, and A. Telford, "Controller Reduction Methods Maintaining Performance and Robustness," *Proc. 27th IEEE Conf. on Decision and Control,* Austin, Texas, 1988, pp. 1159–1164.

[13] M. Vidyasagar, *Control System Synthesis: A Factorization Approach.* Cambridge, Mass.: MIT Press, 1985.

[14] C. N. Nett, C. A. Jacobson, and M. J. Balas, "A Connection between State-Space and Doubly Coprime Fractional Representation," *IEEE Trans. Auto. Control,* Vol. AC-29 (1984), pp. 831–832.

[15] W. S. Levine and M. Athans, "On the Determination of the Optimal Constant Output-Feedback Gains for Linear Multivariable Systems," *IEEE Trans. Auto. Control,* Vol. AC-15, No. 1 (February 1970), pp. 44–48.

[16] B. D. O. Anderson, N. K. Bose, and E. I. Jury, "Output Feedback Stabilization and Related Problems—Solution via Decision Methods," *IEEE Trans. Auto. Control,* Vol. AC-20, No. 1 (February 1975), pp. 53–66.

[17] B. D. O. Anderson and R. W. Scott, "Output Feedback Stabilization—Solution by Algebraic Geometry Methods," *Proc. IEEE,* Vol. 65, No. 6 (June 1977), pp. 849–861.

[18] B. D. O. Anderson and J. B. Moore, *Linear Optimal Control.* Englewood Cliffs, N.J.: Prentice-Hall Inc., 1971.

[19] U.-L. Ly, "A Design Algorithm for Robust Low-Order Controllers," Ph.D. dissertation, Department of Aeronautics and Astronautics, Stanford University, 1982.

[20] S. W. Greely, D. C. Hyland, and S. Richter, "Reduced Order Compensation: LQG Reduction versus Optimal Projection using a Homotopic Continuation Method," *Proc. 26th IEEE Conf. on Decision and Control,* Los Angeles, 1987, pp. 742–747.

11

Digital Controllers

11.1 CONTROLLER IMPLEMENTATION

To this point, we have focused almost exclusively on the problem of designing analog controllers for analog plants. Our concern now is with the problem of implementing these controllers in digital form.

If the controller is ultimately to be implemented in digital form, it is reasonable to ask why a digital design, using a sampled-data representation of the plant, might not be done ab initio. In favor of initially doing an analog design, we note that physical insight concerning an analog plant will often be lost when a sampled-data representation is introduced—a sparse F matrix with entries associated with physically meaningful parameters transforms into a possibly nonsparse exp (Fh) in the sampled data matrix, in which the physical parameters are much more buried. Here h is the sampling interval. Evaluation of a design, and iterative adjustment of design parameters such as performance index weighting matrices is made easier when physical insight is preserved. Insight related to frequency domain notions is also more straightforward to achieve in continuous time, due to an absence of aliasing effects and possible distortion of frequencies in the mapping from continuous to sampled-data descriptions—points that are discussed later. Of course, easier insight means easier design iteration. Are there any clear-cut disadvantages of doing analog design first? One is that certain design freedoms may be lost: deadbeat responses are not obtainable by taking an analog controller and implementing it in digital form, while direct digital design can secure such responses. Again, it may be

323

that multirate digital designs offer freedoms not achievable with analog designs, while proceeding from an analog design to a multirate digital design throws up additional difficulties, for example, in the choice of anti-aliasing filters, which are discussed subsequently. However, experience and intuition both indicate that if it is not possible to obtain a good design in continuous time, it will not be possible to do so in discrete time, and vice versa. In practice, both transformation to digital form of an analog design and direct digital design can be found.

The structure that replaces an analog compensator is shown in Fig. 11.1-1. Key issues that need to be addressed include:

1. Choice of sampling rate
2. Role and design of the analog prefilter
3. The determination of the discrete-time transfer function matrix implemented in the computer
4. The choice of the state-variable realization for the transfer function matrix referred to in 3.

No digital compensation can exactly mimic an analog controller, and it is important to understand the broad trade-off between costs in performance and costs of implementation of the digital compensator. At the broadest level, one may note that if the sampling rate becomes infinitely fast, and infinite precision arithmetic is used, the digital controller will (in principle) duplicate the behavior of the analog one. The cost incurred stems from the requirement to provide hardware to do the calculations extremely fast (infinitely fast being out of the question). This cost is *not* just proportional to the sampling frequency, but rises additionally because the higher the sampling frequency, the greater the word lengths usually required to maintain accuracy. Of course, there are also practical (hardware cost
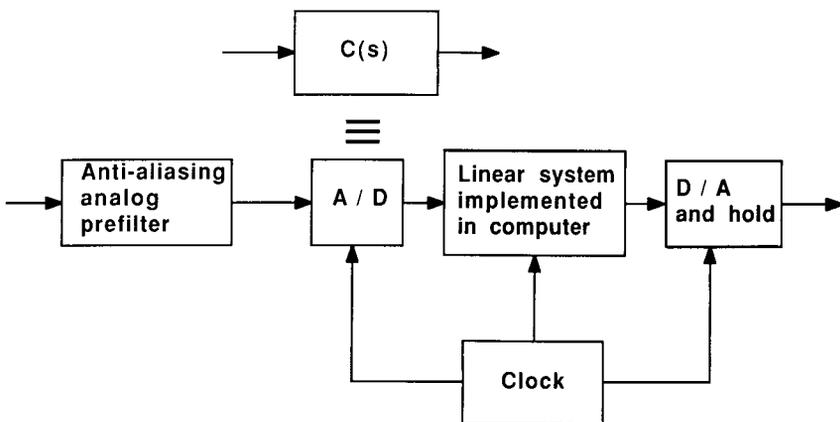
Figure 11.1-1   Digital implementation of a compensator

constrained) upper limits to the word lengths within the compensator and also in the A/D and D/A converters.

Problem 11.1-1.   Consider the arrangement depicted in Fig. 11.1-1 and suppose that the discrete-time linear system implementation on the computer is perfect and the A/D, D/A, and hold operations are perfect (i.e., there is no quantization error anywhere). The system from input of the anti-aliasing analog prefilter to output of the D/A converter and hold is *not* a linear time-invariant system. Is it nonlinear, and if so, in what special way? Is it time-varying, and if so, in what special way?

11.2 SAMPLING TIME SELECTION

Underlying the issue of sample time selection is the Nyquist sampling theorem: Let $s(t)$ for $t \in (-\infty, \infty)$ be an analog signal strictly bandlimited to frequency ω_N, and suppose samples of $s(t)$ are obtained at a frequency of ω_S. Then one can reconstruct $s(t)$ for all t from these samples if and only if $\omega_S > 2\omega_N$. The frequency $2\omega_N$ is termed the Nyquist frequency. For a proof and discussion, almost any book on digital filters can be consulted; for example, [1, 2].

The theorem deals with the potential loss of information when sampling occurs, and specifies circumstances under which no loss will occur: Ideal bandlimiting (which is only possible with infinite dimensional filters), noiseless samples (impossible), and no bound on the time required for reconstruction (again impossible). Practical utilization of the sampling theorem then demands at the least that the sampling frequency be significantly greater than twice the maximum frequency of interest. What does this mean for a compensator implementation? Let T_r be the rise time of a signal and h be the sampling interval, so that $N_r = T_r/h$ is the number of samples per rise time. Considering typical step responses of a first-order system or a second-order system, we see that it appears reasonable to take $N_r = 2$ to 4. This leads to $\omega_s/\omega_0 \simeq 2$ to 4 for first-order systems (where ω_0 is the 3-dB bandwidth) and $\omega_s/\omega_0 \simeq 6$ to 12 for second-order systems (with ω_0 the resonant frequency, and a damping ratio of 0.7).

Perhaps a more fundamental pointer to the required sampling time is provided by the closed-loop bandwidth, call it ω_0. There is no universally agreed relationship between ω_s and ω_0, but figures of ω_s/ω_0 in the interval 4 to 20 have been suggested [3, 4] as a basis for capturing the essential content of the sampling theorem by using a digital compensator. Such a choice should be the starting point in the selection of sampling frequency, but there are a number of other factors that need to be taken into account, which might drive an initial choice upwards. Any advantages from increase in sampling rate must be weighed against hardware cost increases, including the possible need for higher-precision calculations. We now list some of the factors that might dictate an increase in sampling frequency.

Regulator effectiveness.   One task of a regulator is to suppress the effect of random disturbances. In a digital controller, there will be a delay in acting to oppose random disturbances of, on average, one-half the sampling interval. Consequently, the suppression will not be as effective as with the analog compensator. Essentially, the compensator acts in open-loop mode between one sample and the next.

Sensitivity to plant parameter variation.   The point at issue is the ability of a compensator to suppress the effects of plant parameter variation away from the nominal value used for controller design. Examples exist—see [4]— which show that the lower the sampling frequency, the greater the deleterious effect on performance of a particular parameter variation in the plant.

Responsiveness to command changes.   Often, the interconnection of a controller and plant will provide a unity feedback system in which the plant output should follow externally applied reference signals. One effect of a digital compensator is to introduce a delay between the application of the externally applied reference and the *commencement* of the plant response. The delay might be unacceptable.

Smoothness of response to command change.   The D/A converter and hold typically replaces a discrete-time signal by a piecewise constant continuous time signal equal to the most recent discrete value. (By contrast, a first-order hold does a linear extrapolation of the last two discrete values, and so produces a signal with piecewise constant derivative. The zero-order hold is much more common.) With the plant input a piecewise constant signal, two distinct unwanted effects can arise. First, the response may be unacceptably jerky, despite the obvious filtering produced by the plant itself, and perhaps its actuator. Second, unwanted resonances, that is, lightly damped, high-frequency modes, may well be excited, even when these are outside the closed-loop bandwidth. The solution is to decrease the sampling interval.

Some properties of the zero-order hold and first-order hold are examined in the problems.

Implications for closed-loop stability.   We have already noted that one of the effects of use of a digital compensator is to introduce some time delay. On average, this will be one-half the sampling interval. Indeed, a transfer function model of the zero-order hold uses

$$\frac{1 - e^{-sh}}{s}$$

where h is the sampling interval. Observe that for ωh not large

$$\frac{1 - e^{-j\omega h}}{j\omega h} \simeq \frac{1 - \left[1 - j\omega h + \frac{(j\omega h)^2}{2!} - \cdots \right]}{j\omega h}$$

$$\simeq 1 - \frac{j\omega h}{2}$$

$$\simeq e^{-j\omega h/2}$$

This is another way of looking at the time delay introduction. It follows that the phase margin has to be available to accommodate this delay. If it is not, h has to be reduced (or a redesign must be performed).

Actually, the zero-order hold and first-order hold display amplitude variation with ω as well as phase shift. In [3], it is noted that with a sampling frequency 20 times a driving frequency, the percentage errors in tracking a sine wave are 15 percent and 5 percent, respectively, for a zero-order hold and first-order hold. This error represents the combined effects of phase and amplitude variation.

As an alternative to increasing the sampling frequency, one approach to dealing with the delay of a hold is to allow for it ahead of time; that is, the continuous-time plant model is augmented by the inclusion of further dynamics approximating the delay before the design of the analog controller.

Accommodating the anti-aliasing analog filter.   We shall see in the next section that the sampling frequency choice may in fact be dictated by the presence of an anti-aliasing filter.

Performance index.   The performance index with a sampled data controller will take a value greater than that with a continuous-time controller; as the sampling time becomes smaller, the error will become smaller, and the error may influence the choice of sampling time.

Main points of the section.   The sampling frequency ω_s should be chosen as 4 to 20 times the system bandwidth ω_0. As a first step, account must also be taken of regulator effectiveness, sensitivity to plant parameter variations, delay and smoothness of the external response to command changes, closed-loop stability retention in the presence of additional delay, inclusion of the anti-aliasing analog filter, and performance index value.

Problem 11.2-1.   The zero-order hold sets

$$f(t) = f(kh) \qquad kh \le t < (k + 1)h$$

and the first-order hold sets

$$f(t) = f(kh) + \frac{(t - kh)}{h}[f(kh) - f(\overline{k - 1}\, h)] \qquad kh \le t < (k + 1)h$$

Suppose there is an underlying signal $g(t)$ with $g(kh) = f(kh)$. Then $f(t)$ should approximate $g(t)$ in $(kh, \overline{k + 1}\, h)$. Show that with the zero-order hold,

$$\sup_{t \in [kh, \overline{k+1}h]} |g(t) - f(t)| \le h \max_t |\dot{g}(t)|$$

and for the first-order hold,

$$\sup_{t \in [kh, \overline{k+1}h]} |g(t) - f(t)| \le h^2 \max_t |\ddot{g}(t)|$$

Problem 11.2-2. The first-order hold transfer function has the property that if injected at time $-h$ with an impulse $u_{-1}\delta(t + h)$ and at time 0 with an impulse $u_0\delta(t)$, it produces over $[0, h)$ the (nonimpulsive) output

$$u_0 + \frac{t(u_0 - u_{-1})}{h}$$

(i.e., it extrapolates from u_0 with a slope defined by the last two samples). Show that the impulse response is

$$g(t) = 1 + \frac{t}{h} \qquad 0 \le t < h$$

$$= -\frac{(t - 1)}{h} \qquad h \le t < 2h$$

$$= 0 \qquad\qquad 2h \le t$$

and that the associated transfer function is

$$G(s) = h\left[\frac{1 - e^{-sh}}{hs}\right]^2 (hs + 1)$$

Show that the phase characteristic has zero slope at zero frequency (DC).

11.3   ANTI-ALIASING ANALOG PREFILTER

To understand the role of the anti-aliasing filter, a key property of the sampling process must be understood. Let $f(t)$ be a continuous time signal, with Fourier transform $F(j\omega)$. Let h be the sampling interval, ω_s the sampling frequency in radians per second, and $f(kh)$, $k = 0, \pm 1, \pm 2, \ldots$ the discrete samples of $f(t)$. Suppose that the z-transform of $f(kh)$ is evaluated on the unit circle. Thus

$$F_s(e^{j\omega h}) = \sum_{-\infty}^{+\infty} f(kh)e^{-jkh\omega} \qquad (11.3\text{-}1)$$

Then $F_s(e^{j\omega h})$ and $F(j\omega)$ are related, by the formula

$$F_s(e^{j\omega h}) = \frac{1}{h} \sum_{-\infty}^{+\infty} F(j\omega + k\omega_s) \tag{11.3-2}$$

This is a standard result in signal processing; see, for example, [1, 2]. The sampling theorem deals with the case when $F(j\omega) = 0$ for $\omega \geq \frac{\pi}{h} = \frac{\omega_s}{2}$. Then the sum in (11.3-2) contains for $\omega < \frac{\omega_s}{2}$ just one term and

$$F_s(e^{j\omega h}) = \frac{1}{h} F(j\omega) \tag{11.3-3}$$

Thus the discrete-time spectrum is a copy of the continuous-time spectrum.

But what happens if the sampling theorem condition is not fulfilled? Then the discrete-time spectrum at ω is a sum of values of the continuous-time spectrum at $\omega + k\omega_s$, $k = 0, \pm 1, \pm 2, \ldots$. Many frequencies before sampling map into one frequency after sampling, in a way which would not allow their subsequent disentangling, no matter what form of signal processing was used. One says that the frequency ω is the alias of $\omega + k\omega_s$, $k = \pm 1, \pm 2, \ldots$.

Now suppose the A/D converter in the digital compensator is operated with no anti-aliasing filter ahead of it. Then the converter will map frequencies in excess of $\omega_s/2$ into frequencies less than $\omega_s/2$. If ω_s is, say, 5–20 times the closed-loop system bandwidth, is this likely to be a problem? Yes, it is. Sensor noise, which is generally wideband, would go straight into the A/D converter and be in effect amplified through the aliasing process. Also, there may be ripples or slowly decreasing rapid oscillations in the plant output outside the closed-loop bandwidth which are there because the plant is excited with a piecewise constant signal. These too might be aliased downwards.

We need to prevent aliasing of these unwanted signals, especially if aliasing into the passband of the closed-loop system is involved, because this will have a deleterious effect on performance. The solution, not a perfect one, of course, is to use an anti-aliasing filter, normally a low-pass first-order, second- or higher-order filter with transfer function $\omega_f (s + \omega_f)^{-1}$ or $\omega_f^2(s^2 + \sqrt{2}\omega_f s + \omega_f^2)^{-1}$, and so on.

The introduction of such filters brings another problem: an increase in phase lag in the closed-loop. The phase lag at the frequency ω_0 defining the system bandwidth must be such as not to use up all the phase margin of the system. How, then, should the cut-off frequency ω_f of the anti-aliasing filter be chosen, in order to satisfy the requirements that it must have significant attenuation at $\omega_s/2$, but modest phase shift at ω_0?

The conservative approach is to select the multiple ω_s/ω_0 sufficiently large that there is no problem in fulfilling the two possibly conflicting objectives. The prefilter cut-off frequency ω_f is chosen well above ω_0 so that the phase shift at ω_0 is small, and ω_s is chosen as 5–10 times ω_f. Consequently, ω_s may be 20 to 100 times ω_0. The prefilter is independent of the design of the analog loop in the first instance.

The alternative is to lower the ratio ω_f/ω_0 substantially, and to design the analog loop with extra phase margin to accommodate the phase lag introduced by the analog prefilter. The ratio ω_s/ω_f can also be reduced, the cost being progressive degradation of performance due to the aliasing of undesired signals into the system passband.

Main points of the section.   The analog anti-aliasing prefilter is used to avoid the aliasing of undesirable signals such as sensor noise into the pass-band. The cut-off frequency ω_f of the filter is related to the sampling frequency ω_s (to secure adequate anti-aliasing) and to the system bandwidth ω_0 (to avoid excessive phase shift). If phase shift introduced by the filter is a problem, it can (ideally) be avoided by taking ω_s/ω_0 sufficiently large, or by designing the analog loop to provide extra phase margin that will be absorbed by the prefilter phase lag.

Problem 11.3-1.   Consider a prefilter with transfer function $\omega_f^2(s^2 + \sqrt{2}\omega_f s + \omega_f^2)^{-1}$. Suppose that at the system bandwidth ω_0, the phase shift should be 5 deg while the attenuation at the Nyquist frequency $\omega_s/2$ should be 20 dB. Express ω_f and ω_s in terms of ω_0.

11.4 THE DISCRETE-TIME TRANSFER FUNCTION

Suppose that the sampling rate and anti-aliasing filter have been determined. We now consider the case of proceeding from the continuous-time transfer function $C(s)$ to a discrete-time $D(z)$, to be implemented on the digital computer, so that the cascade of A/D converter, $D(z)$ and the D/A converter and hold behaves like $C(s)$.

We begin the discussion assuming the compensator is strictly proper. Suppose that

$$C(s) = H'(sI - F)^{-1}G \qquad (11.4\text{-}1)$$

with $\{F, G, H\}$ a minimal triple. The variables e and u are chosen to designate the compensator input and output, since often the input will be an error between a reference signal and the plant output, while the compensator output will serve as the plant input. Then to approximate the equations

$$\dot{x} = Fx + Ge \qquad (11.4\text{-}2a)$$

$$u = H'x \qquad (11.4\text{-}2b)$$

we can use

$$x\overline{(k+1h)} = \exp(Fh)x(kh) + \left[\int_0^h \exp(Fs)\,ds\right]Ge(kh) \qquad (11.4\text{-}3a)$$

$$u(kh) = H'x(kh) \qquad (11.4\text{-}3b)$$

This is the most straightforward approach. Other approaches are discussed subsequently. Associated with (11.4-3) is the transfer function

$$D(z) = H'(zI - A)^{-1}B \qquad (11.4\text{-}4)$$

where

$$A = \exp\,(Fh) \qquad B = \left[\int_0^h \exp\,(Fs)\,ds\,\right]G \qquad (11.4\text{-}5)$$

Each eigenvalue of F in Re $[s] < 0$, Re $[s] = 0$ and Re $[s] > 0$ corresponds to an eigenvalue of A in $|z| < 1$, $|z| = 1$, or $|z| > 1$. In fact, if s_i is an eigenvalue of F or a pole of $C(s)$, the corresponding eigenvalue of A or poles of $D(z)$, call it z_i, is given by $z_i = \exp\,(s_i h)$. Observe that Re $s_i < 0 \Leftrightarrow |z_i| < 1$. Hence if the poles of $C(s)$ are stable, so are the poles of $D(z) = H'(zI - A)^{-1}B$. When h is small, the finite zeros s_i *of* $C(s)$ map into zeros z_i *of* $D(z)$ with $z_i \simeq e^{s_i h}$. (An exact formula is not available.) Hence left half-plane zeros map into the interior of the unit circle.

In general, though, we must also deal with the infinite zeros of $C(s)$. For scalar $C(s)$, the number of such zeros is the difference between the denominator degree and the numerator degree. When $C(s)$ has one infinite zero, so does $D(z)$. When $C(s)$ has two, $D(z)$ has one infinite zero at infinity, and one zero inside the unit circle which approaches -1 as h$\to$ 0. When $C(s)$ has more than two zeros at $s = \infty$, $D(z)$ necessarily has for small h one or more zeros outside $|z| = 1$, [3]. Note that such zeros are nonminimum phase.

At least three other approaches to the discretization of (11.4-2) have been proposed. These are, in contrast to (11.4-3), not exact when $e\,(kh)$ is piecewise constant. But of course, $e\,(t)$ will rarely if ever be piecewise constant, and so (11.4-3) is also an approximation. The first two of these other approximations use a forward difference or a backward difference; that is, $\dot{x}\,(t)$ is replaced by

$$\dot{x}\,(t) \simeq \frac{1}{h}\,[x(k + 1)h - x(kh)] \qquad (11.4\text{-}6)$$

and

$$\dot{x}\,(t) \simeq \frac{1}{h}\,[x(kh) - x(\overline{k - 1}\,h)] \qquad (11.4\text{-}7)$$

with, in each case, the rest of (11.4-2) being evaluated at $t = kh$. This is equivalent to using transformations in the frequency domain; for the forward difference, this is

$$z = 1 + sh \qquad s = \frac{z - 1}{h} \qquad (11.4\text{-}8)$$

so that

$$D(z) = C\left[\frac{z - 1}{h}\right] \qquad C(s) = D(1 + sh) \qquad (11.4\text{-}9)$$

For the backward difference

$$z = \frac{1}{1 - sh} \qquad s = \frac{1 - z^{-1}}{h} \tag{11.4-10}$$

with

$$D(z) = C\left[\frac{1 - z^{-1}}{h}\right] \qquad C(s) = D\left[\frac{1}{1 - sh}\right] \tag{11.4-11}$$

The forward difference maps Re $s < 0$ into Re $z < 1$ and conversely. This seems unattractive—but if the sampling rate is high enough, a stable pole will map into $|z| < 1$, in fact map into a point close to $z = 1$.

The third alternative approach rests on a trapezoidal integration formula, rather than a forward or backward difference. As such, it lies between the former two approaches. The frequency domain transformation is

$$z = \frac{1 + sh/2}{1 - sh/2} \qquad s = \frac{2}{h}\frac{z - 1}{z + 1} \tag{11.4-12}$$

This bilinear transformation certainly maps Re $[s] < 0$ onto $|z| < 1$. However, if $C(j\omega_0) = 0$, it does not follow that $D(\exp j\omega_0 h) = 0$. Rather, $D(\exp j\overline{\omega}_0 h) = 0$ where $\overline{\omega}_0 = \frac{z}{h}\tan^{-1}\frac{\omega h}{2}$. Thus there is some frequency domain warping. It is possible to eliminate warping at one frequency, ω_1 say, by using

$$s = \frac{\omega_1}{\tan \omega_1 h/2}\frac{z - 1}{z + 1} \tag{11.4-13}$$

This ensures that $C(j\omega_1) = 0$ implies $D(\exp j\omega_1 h) = 0$. However, distortion remains at other frequencies. These ideas are all discussed in [3, 4].

Allowing for computation time.   Implicit in our discussion to this point has been the assumption that the calculation of $x(\overline{k + 1}\ h)$ for $x(kh)$ and $e(kh)$ can be completed before time $(k + 1)h$, in order that the correct value of $u(\overline{k + 1}\ h)$ can be generated and applied beginning at time $(k + 1)h$. This assumption will be valid if we match hardware capability to the requirement, and this is in principle possible. But the situation changes as soon as we allow $C(s)$ to be nonzero at $s = \infty$:

$$C(s) = H'(sI - F)^{-1}G + J \tag{11.4-14}$$

Equation (11.4-2a) is unaltered, as is (11.4-3a). On the other hand, we have

$$u = H'x + Je \tag{11.4-15}$$

and

$$u(kh) = H'x(kh) + Je(kh) \tag{11.4-16}$$

Equation (11.4-16) apparently demands that $u(kh)$ be available instantaneously with the sample $e(kh)$. If the computation time is very small, the assumption implicit in (11.4-16) may be harmless. But otherwise, there is here another source of time delay in the system, and possible problems of synchronization. There is, however, a way out. Suppose that δ is the computation time, always assumed less than h. Then we aim to produce a sequence $u(kh + \delta)$, instead of $u(kh)$ and introduce clock skew to the output D/A converter. Of course, the sequence $u(kh + \delta)$ must be produced with information available at time kh. Now with $e(t)$ constant over $[kh, kh + \delta]$, the basic state differential equation implies

$$x(kh + \delta) = \exp(F\delta)x(kh) + \left[\int_0^\delta \exp(Fs)\,ds\right]Ge(kh)$$

This means that from (11.4-15) that

$$u(kh + \delta) = H'x(kh + \delta) + Je(kh)$$

$$= H'\exp(F\delta)x(kh) + \left[H'\int_0^\delta \exp(Fs)\,dsG + J\right]e(kh) \qquad (11.4\text{-}17)$$

$$= \bar{H}'x(kh) + \bar{J}e(kh)$$

for some new pair $\bar{H}, \bar{J}$ defined in an obvious way.

Main points of the section.   There are several approaches to forming discrete-time linear equations from continuous-time equations. No one procedure appears to be uniformly preferable. It is possible to make allowance for computation time by introducing clock skew into the D/A converter at the compensator's output.

Problem 11.4-1.   When h is very small, approximations of $A = \exp Fh$ and $B = [\int_0^h \exp(Fs)\,ds]G$ are provided by $I + Fh$ and hG. Also, $1 + s_ih$ is an approximation of $\exp(s_ih)$. Show that the finite zeros of $C(s) = H'(sI - F)^{-1}G$ map into zero z_i of $D(z) = H'(zI - A)^{-1}B$ via $z_i \simeq e^{s_ih}$ when h is small, by using the above approximation.

Problem 11.4-2.   Suppose that

$$D(z) = C\left[\frac{2}{h}\frac{z - 1}{z + 1}\right]$$

and $C_s = H'(sI - F)^{-1}G$. Show that a state variable realization for $D(z)$ is provided by

$$D(z) = \frac{h}{2}H'\left[I - \frac{hF}{2}\right]^{-1}G + hH'\left[I - \frac{hF}{2}\right]^{-1}\left\{zI - \left[I - \frac{hF}{2}\right]^{-1}\right.$$

$$\left. \times \left[I + \frac{hF}{2}\right]\right\}^{-1}\left[I - \frac{hF}{2}\right]^{-1}G$$

Problem 11.4-3.   Find the discrete time equations corresponding to application of the forwards and backwards integration rules to $\dot{x} = Fx + Gu$.

11.5 STATE-VARIABLE IMPLEMENTATION OF THE DISCRETE-TIME TRANSFER FUNCTION

Were it practical to do all calculations in the digital computer with infinite precision arithmetic, and to use infinitely long words in the A/D and D/A converters, the question would not arise of how one should implement the discrete-time transfer function. In practice, there are at least two issues that must be addressed—what word lengths and state-variable realizations should be used?

Word length requirements are linked to quantization errors, and quantization errors occur for one of several reasons:

1. The A/D converters replace an analog signal by a digital signal with finite word length.
2. Coefficients in the linear discrete system are quantized.
3. Arithmetic operations in the computer can lead to small quantization errors (due to round-off or truncation) or large errors, due to overflow. (Designs ought to be made which always avoid overflow.)

Quantization errors can also cause limit cycles.

Many of these issues are addressed in treatments of digital filtering; see, for example, [5–8]. These treatments also explain that the particular state-variable realization employed can make a great difference in the word length required to secure a given amount of accuracy in the compensator's output. Canonical form realizations tied to companion matrices are particularly to be avoided. On the other hand, particularly for high-order systems, a relatively sparse set of matrices in the state-variable equations is desirable, in order not to overload the processor with too many arithmetic operations per clock cycle. These considerations have led to the conclusion that generally, a scalar transfer function should be realized as a cascade or parallel connection (or mixture) of second-order sections, together with a first-order section if required. The determination of scaling issues, and word lengths to secure a prescribed level of quantization error, and the avoidance of limit cycles under zero input conditions all become relatively straightforward.

It is, however, not straightforward to ensure avoidance of limit cycles under nonzero inputs, and there seems little alternative to testing in the presence of nonzero inputs a design that is free of limit cycles under zero inputs.

Because of the general desirability of minimizing the computation time, which absolutely must be less than the sampling interval, attention can also be given to structures that allow parallel computation. Many of these issues are discussed in [5].

Much less is known about the realizations of matrix transfer functions. Also, should the open-loop compensator be unstable, a possibility in a control problem,

digital filtering concepts may be less relevant, since some digital filtering results presuppose a stable open-loop system. Scaling is, however, easy to address—one simply uses the closed loop, assumed stable, as the basis for scaling.

It is fair to say that all these ideas are in a current state of development, and it will be some time before the key concepts will be presentable in a matured form.

Main points of the section.   The best practical realizations are those based on cascading and/or paralleling second-order sections, which are optimized for scaling and round-off error.

REFERENCES

[1] L. R. Rabiner and B. Gold, *Theory and Application of Digital Signal Processing*. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1975.

[2] D. W. Stanley, *Digital Signal Processing*. Reston, Va.: Reston Pub. Co., 1975.

[3] K. J. Åström and B. Wittenmark, *Computer Controlled Systems: Theory and Design*. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1984.

[4] G. F. Franklin and J. D. Powell, *Digital Control of Dynamic Systems*. Reading, Mass.: Addison-Wesley Pub. Co., 1980.

[5] P. Moroney, *Issues in the Implementation of Digital Feedback Compensators*. Cambridge, Mass.: MIT Press, 1983.

[6] T. W. Parks and C. S. Burrus, *Digital Filter Design*. New York: John Wiley and Sons, Inc., 1987.

[7] R. A. Roberts and C. T. Mullis, *Digital Signal Processing*. Reading, Mass.: Addison-Wesley, 1987.

[8] L. B. Jackson, *Digital Filters and Signal Processing*. Boston, Mass.: Kluiver, 1986.

A

Brief Review
of Some Results
# of Matrix Theory

The purpose of this appendix is to provide a rapid statement of those particular results of matrix theory used in this book. For more extensive treatments standard textbooks—for example, [1–3]—should be consulted. Increasingly, there have become available specialized software packages dealing with linear algebra manipulations, and familiarity with one or more of these packages is also desirable.

1. Matrices and vectors. An $m \times n$ matrix A consists of a collection of mn quantities† $a_{ij}(i = 1, 2, \ldots, m; j = 1, 2, \ldots, n)$ written in an array of m rows and n columns:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

Sometimes, one simply writes

$$A = (a_{ij})$$

The quantity a_{ij} is an *entry* (the $(i\text{-}j)$th entry, in fact) of A.

---

† The a_{ij} will be assumed real in most of our discussions.

An *m vector*, or, more fully, a *column m vector*, is a matrix with 1 column and *m* rows; thus,

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_m \end{bmatrix}$$

defines *x* as column *m* vector, whose *i*th entry is the quantity x_i. A row *n* vector is a matrix with 1 row and *n* columns.

2. Addition, subtraction, and multiplication by a scalar.

Two matrices *A* and *B* with the *same number of rows and also the same number of columns* may be added, subtracted, or individually multiplied by a scalar. With k_1, k_2, scalar, the matrix

$$C = k_1 A + k_2 B$$

is defined by

$$c_{ij} = k_1 a_{ij} + k_2 b_{ij}$$

Thus, to add two matrices, one simply adds corresponding entries; to subtract two matrices, one simply subtracts corresponding entries, and so forth. Of course, addition is commutative—that is,

$$A + B = B + A$$

3. Multiplication of matrices.

Consider two matrices *A* and *B*, with *A* an $m \times p$ matrix and *B* a $p \times n$ matrix. Thus, the number of columns of *A* equals the number of rows of *B*. The product *AB* is an $m \times n$ matrix defined by

$$C = AB$$

with

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}$$

Notice that *C* has the same number of rows as *A*, and the same number of columns as *B*.

The product of three (or more) matrices can be defined by

$$D = ABC = (AB)C = A(BC)$$

In other words, multiplication is associative. However, multiplication is not commutative—that is, it is *not* in general true that

$$AB = BA$$

In fact, although AB can be formed, the product BA may not be capable of being formed.

For any integer p, the $p \times p$ matrix

$$I = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & 0 \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

possessing p rows and columns is termed the *identity matrix of order p*. It has the property that with A any $m \times p$ matrix,

$$AI = A$$

Likewise, the identity matrix of order m has the property that

$$IA = A$$

Any matrix consisting entirely of entries that are zero is termed *the zero matrix*. Its product with any matrix produces the zero matrix, whereas if it is added to any matrix, it leaves that matrix unaltered.

Suppose A and B are both $n \times n$ matrices (A and B are then termed square matrices). Then AB is square. It can be proved then that

$$|AB| = |A| \, |B|$$

where $|A|$ is the determinant of A.

[The definition of the determinant of a square matrix is standard. One way of recursively defining $|A|$ for A an $n \times n$ matrix is to expand A by its first row; thus

$$|A| = a_{11} \begin{vmatrix} a_{22} & \cdots & a_{2n} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ a_{n2} & \cdots & a_{mn} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} & a_{24} & \cdots & a_{2n} \\ a_{31} & a_{33} & a_{34} & \cdots & a_{3n} \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ a_{n1} & a_{n3} & a_{n4} & \cdots & a_{mn} \end{vmatrix}$$

$$+ \, a_{13} \begin{vmatrix} a_{21} & a_{22} & a_{24} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{34} & \cdots & a_{3n} \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ a_{n1} & a_{n2} & a_{n4} & \cdots & a_{nn} \end{vmatrix} - \cdots$$

This expresses $|A|$ in terms of determinants of $(n-1) \times (n-1)$ matrices. In turn, these determinants may be expressed by using determinants of $(n-2) \times (n-2)$ matrices, and so on.]

4. Direct sum of two matrices.

Let A be an $n \times n$ matrix and B an $m \times m$ matrix. The *direct* sum of A and B, written $A \dotplus B$, is the $(n + m) \times (n + m)$ matrix

$$\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$$

5. Transposition.

Suppose A is an $m \times n$ matrix. The *transpose* of A, written A', is an $n \times m$ matrix defined by

$$B = A'$$

where

$$b_{ij} = a_{ji}$$

It is easy to establish the important result

$$(AB)' = B'A'$$

which extends to

$$(ABC)' = C'B'A'$$

and so on. Also, trivially, one has

$$(A + B)' = A' + B'$$

6. Singularity and nonsingularity.

Suppose A is an $n \times n$ matrix. Then A is said to be *singular* if $|A|$ is zero. Otherwise, A is termed nonsingular.

7. Rank of a matrix.

Let A be an $m \times n$ matrix. The rank of A is a positive integer q such that some $q \times q$ submatrix of A, formed by deleting $(m - q)$ rows and $(n - q)$ columns, is nonsingular, whereas no $(q + 1) \times (q + 1)$ submatrix is nonsingular.

The rank of A is also the maximum number of linearly independent rows of A and the maximum number of linearly independent columns of A.

It can be shown that

$$\text{rank } (AB) \leq \min[\text{rank } A, \text{rank } B]$$

If rank A is equal to the number of columns *or* the number of rows of A, A is often said to have *full rank*. If A is $n \times n$, the statement rank $A = n$ is equivalent to the statement A is nonsingular. If, for an arbitrary matrix A, rank $A = 0$, then A is the zero matrix.

8. Range space and null space of a matrix.

Let A be an $m \times n$ matrix. The range space of A, written $\mathfrak{R}[A]$, is the set of all vectors Ax, where x ranges over the set of all n vectors. The range space has dimension equal to

the rank of A—that is, the maximal number of linearly independent vectors in $\mathcal{R}[A\,]$ is rank A. The null space of A, written $\mathcal{N}[A\,]$, is the set of vectors y for which $Ay = 0$.

An easily proved property is that $\mathcal{R}[A\,']$ and $\mathcal{N}[A\,]$ are orthogonal—that is, if $y_1 = A\,'x$ for some x, and if y_2 is such that $Ay_2 = 0$, then $y_1'y_2 = 0$.

9. Inverses and pseudoinverses.

Let A be a square matrix. If, but only if, A is nonsingular, there exists a unique matrix, call it B, termed the *inverse of A*, with the properties

$$BA = AB = I$$

The inverse of A is generally written A^{-1}. There are many computational procedures for passing from a prescribed A to its inverse A^{-1}. A formula is, in fact, available for the entries of $B = A^{-1}$, obtainable as follows.

Define the cofactor of the $i - j$ entry of A as $(-1)^{i+j}$ times the determinant of the matrix obtained by deleting from A the ith row and jth column, that is, the row and column containing a_{ij}. Then

$$b_{ij} = \frac{1}{|A|} \times \text{cofactor of } a_{ji}$$

It easily follows that

$$(A^{-1})' = (A')^{-1}$$

If A_1 and A_2 are two $n \times n$ nonsingular matrices, it can be shown that

$$(A_1 A_2)^{-1} = A_2^{-1} A_1^{-1}$$

When A is singular, one can define a unique object $A^{\#}$, the Moore-Penrose pseudoinverse of A, such that $A^{\#}A$ acts as the identity matrix on as large a set of vectors as practical [viz. $R(A\,')$]. The following provide the definition.

$$A^{\#}Ax = x \qquad \forall x \in \mathcal{R}(A\,')$$

$$A^{\#}x = 0 \qquad \forall x \in \mathcal{N}(A\,')$$

(Should A be nonsingular, then $A^{\#} = A^{-1}$).

There are many key properties, for example, $(Ax - y)'(Ax - y)$ for fixed y is minimized with respect to x by $x = A^{\#}y$; if A is scalar and nonzero, then $A^{\#} = A^{-1}$, otherwise $A^{\#} = 0$; if Λ is diagonal, $\Lambda^{\#} = \text{diag } \lambda_i^{\#}$; if $A = T'\Lambda T$ with T nonsingular and Λ diagonal, $A^{\#} = T^{-1}\Lambda^{\#}(T')^{-1}$, and

$$(A^{\#})^{\#} = A \qquad A^{\#}AA^{\#} = A^{\#} \qquad AA^{\#}A = A$$

10. Powers of a square matrix.

For positive m, A^m for a square matrix A is defined as $AA \ldots A$, there being m terms in the product. For negative m, let $m = -n$, where n is positive; then $A^m = (A^{-1})^n$. It follows that $A^p A^q = A^{p+q}$ for any integers p and q, positive or negative, and likewise that $(A^p)^q = A^{pq}$.

A polynomial in A is a matrix $p(A) = \sum_{i=0}^{r} a_i A^i$ where the a_i are scalars. Any

two polynomials in the same matrix commute—that is, $p(A)q(A) = q(A)p(A)$, where p and q are polynomials. It follows that $p(A)q^{-1}(A) = q^{-1}(A)p(A)$, and that such *rational functions* of A also commute.

11. Exponential of a square matrix.   Let A be a square matrix. Then it can be shown that the series

$$I + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \cdots$$

converges, in the sense that the i-j entry of the partial sums of the series converges for all i and j. The sum is defined as e^A. It follows that

$$e^{At} = I + At + \frac{1}{2!}A^2t^2 + \cdots$$

Other properties are: $p(A)e^{At} = e^{At}p(A)$ for any polynomial A, and $e^{-At} = [e^{At}]^{-1}$.

12. Differentiation and integration.   Suppose A is a function of a scalar variable t, in the sense that each entry of A is a function of t. Then

$$\frac{dA}{dt} = \left(\frac{da_{ij}}{dt}\right)$$

It follows that

$$\frac{d}{dt}(AB) = \frac{dA}{dt}B + A\frac{dB}{dt}$$

Also, from the definition of e^{At}, one has

$$\frac{d}{dt}(e^{At}) = Ae^{At} = e^{At}A$$

The integral of a matrix is defined in a straightforward way as

$$\int A\,dt = \left(\int a_{ij}\,dt\right)$$

Suppose ϕ is a scalar function of a vector x. Then

$$\frac{d\phi}{dx} = \text{a vector whose } i\text{th entry is} \frac{\partial\phi}{\partial x_i}$$

Suppose ϕ is a scalar function of a matrix A. Then,

$$\frac{d\phi}{dA} = \text{a matrix whose } i\text{-}j \text{ entry is } \frac{\partial\phi}{\partial a_{ij}}$$

Suppose z is a vector function of a vector x. Then,

$$\frac{dz}{dx} = \text{a matrix whose } i\text{-}j \text{ entry is } \frac{\partial z_i}{\partial x_j}$$

13. Eigenvalues and eigenvectors of a square matrix.

Let A be an $n \times n$ matrix. Construct the polynomial $|sI - A|$. This is termed the *characteristic polynomial* of A; the zeros of this polynomial are the *eigenvalues* of A. If λ_i is an eigenvalue of A, there always exists at least one vector x satisfying the equation

$$Ax = \lambda_i x$$

The vector x is termed an *eigenvector* of the matrix A. If λ_i is not a repeated eigenvalue—that is, if it is a simple zero of the characteristic polynomial, to within a scalar multiple x is unique. If not, there *may* be more than one eigenvector associated with λ_i. If λ_i is real, the entries of x are real, whereas if λ_i is complex, the entries of x are complex.

If A has zero entries everywhere off the main diagonal—that is, if $a_{ij} = 0$ for all i, j, with $i \neq j$, then A is termed *diagonal*. (*Note:* Zero entries are still permitted on the main diagonal.) It follows trivially from the definition of an eigenvalue that the diagonal entries of the diagonal A are precisely the eigenvalues of A.

It is also true that for a general A,

$$|A| = \prod_{i=1}^{n} \lambda_i$$

If A is singular, A possesses at least one zero eigenvalue.

The eigenvalues of a rational function $r(A)$ of A are the numbers $r(\lambda_i)$, where λ_i are the eigenvalues of A. The eigenvalues of e^{At} are $e^{\lambda_i t}$.

14. Trace of a square matrix A.   Let A be $n \times n$. Then the trace

of A, written tr $[A]$, is defined as

$$\text{tr}[A] = \sum_{i=1}^{n} a_{ii}$$

An important property is that

$$\text{tr}[A] = \sum_{i=1}^{n} \lambda_i$$

where the λ_i are eigenvalues of A. Other properties are

$$\text{tr}[A + B] = \text{tr}[B + A] = \text{tr}[A] + \text{tr}[B]$$

and, assuming the multiplications can be performed to yield square product matrices

$$\text{tr}[AB] = \text{tr}[B'A'] = \text{tr}[BA] = \text{tr}[A'B']$$

$$\text{tr}[A'A] = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}^2$$

15. Orthogonal, symmetric, skew-symmetric, unitary, hermitian and skew-hermitian matrices, and their eigenvalue properties.

If a square A is such that $AA' = I$, and thus $A'A = I$, A is termed *orthogonal*. The eigenvalues of A then have a magnitude of unity. If $A = A'$, A is termed *symmetric*, and the eigenvalues of A are all real. Moreover, if x_1 is an eigenvector associated with λ_1, x_2 with λ_2, and if $\lambda_1 \ne \lambda_2$, then $x_1' x_2 = 0$. The vectors x_1 and x_2 are termed *orthogonal*. (*Note:* Distinguish between an orthogonal matrix and an orthogonal pair of vectors.) If $A = -A'$, A is termed *skew* or *skew symmetric*, and the eigenvalues of A are pure imaginary.

The corresponding properties for complex matrices are important. Let a superscript asterisk denote complex conjugate transpose. A square U with $U^*U = I$ (and thus $UU^* = I$) is termed *unitary*, and all eigenvalues have magnitude unity. If $U = U^*$, U is termed *hermitian*, and all eigenvalues are real. Moreover, if x_1 and x_2 are eigenvectors associated with differing eigenvalues λ_1 and λ_2, then $x_1^* x_2 = 0$. If $U = -U^*$, U is skew hermitian, and all eigenvalues are pure imaginary.

16. The Cayley-Hamilton theorem.

Let A be a square matrix, and let $|sI - A| = s^n + \alpha_1 s^{n-1} + \cdots + \alpha_n$; then,

$$A^n + \alpha_1 A^{n-1} + \cdots + \alpha_n I = 0$$

From the Cayley-Hamilton theorem, if follows that A^m for any $m \ge n$ and e^A are expressible as a linear combination of $I, A, \ldots, A^{n-1}$.

17. Similar matrices and diagonalizability.

Let A and B be $n \times n$ matrices. If there exists a nonsingular $n \times n$ matrix T such that $B = T^{-1}AT$, the matrices A and B are termed *similar*. Similarity is an equivalence relation. Thus:

1. A is similar to A.
2. If A is similar to B, then B is similar to A.
3. If A is similar to B and B is similar to C, then A is similar to C.

Similar matrices have the same eigenvalues. This may be verified by observing that

$$sI - B = T^{-1}sIT - T^{-1}AT = T^{-1}(sI - A)T$$

Therefore,

$$|sI - B| = |T^{-1}||sI - A||T| = |sI - A||T^{-1}||T|$$

But $T^{-1}T = I$ so that $|T^{-1}||T| = 1$. The result is then immediate.

If, for a given A, a matrix T can be formed such that

$$\Lambda = T^{-1}AT$$

is diagonal, then A is termed *diagonalizable*, the diagonal entries of Λ are eigenvalues of A, and the columns of T turn out to be eigenvalues of A. Even when A is real, T may be necessarily complex.

Not all square matrices are diagonalizable, but matrices that have no repeated eigenvalues are diagonalizable, as are orthogonal, symmetric, and skew-symmetric matrices, and unitary, hermitian, and skew-hermitian matrices.

In fact, when A is real symmetric, it can be diagonalized by a real orthogonal matrix, and when it is unitary, hermitian, or skew-hermitian, it can be diagonalized by a unitary matrix. When it is orthogonal or skew symmetric, a real orthogonal matrix can be found that will almost diagonalize A, in fact $T^{-1}AT = T'AT$ becomes a direct sum of 2×2 matrices of the form

$$\begin{bmatrix} \cos\theta_i & -\sin\theta_i \\ \sin\theta_i & \cos\theta_i \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 0 & \alpha_i \\ -\alpha_i & 0 \end{bmatrix}$$

respectively, together with possibly I and $-I$ for orthogonal A.

18. Jordan form.   Not all square matrices are diagonalizable. But it is always possible to get very close to diagonal matrix via a *similarity transformation*. In fact, there always exists a matrix T such that

$$T^{-1}AT = \begin{bmatrix} \lambda_1 & 1 & & & & & & & \\ & \lambda_1 & & & & & & & \\ & & \lambda_1 & & & & & & \\ & & & \lambda_2 & 1 & & & & \\ & & & & \lambda_2 & & & & \\ & & & & & \lambda_3 & 1 & & \\ & & & & & & \lambda_3 & 1 & \\ & & & & & & & \lambda_3 & \\ & & & & & & & & \ddots \\ & & & & & & & & & \ddots \end{bmatrix}$$

or something similar. Here, all blank entries are zero, the eigenvalues of A occur on the main diagonal, and there *may or may not* be entries of 1 above and to the right of repeated eigenvalues—that is, on the superdiagonal. For any A, the distribution of 1s and 0s on the superdiagonal is fixed, but different A yield different distributions. The preceding almost-diagonal matrix is called the *Jordan canonical form* of A. The *Jordan blocks* of A are the matrices

$$\begin{bmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{bmatrix} \quad \begin{bmatrix} \lambda_1 \end{bmatrix} \quad \begin{bmatrix} \lambda_2 & 1 \\ 0 & \lambda_2 \end{bmatrix} \quad \begin{bmatrix} \lambda_3 & 1 & 0 \\ 0 & \lambda_3 & 1 \\ 0 & 0 & \lambda_3 \end{bmatrix} \quad \text{etc.}$$

If A is real and also the matrix T is restricted to being real, one can obtain a "real" Jordan form. Eigenvalues necessarily occur in complex conjugate pairs $\alpha_i \pm j\beta_i$, and instead of there being, say, two 1×1 Jordan blocks $[\alpha_i + j\beta_i]$ and $[\alpha_i - j\beta_i]$, these are replaced in the real Jordan form by a real 2×2 block

$$\begin{bmatrix} \alpha_i & -\beta_i \\ \beta_i & \alpha_i \end{bmatrix}$$

The idea extends to multiple eigenvalues.

19. Schur form.
With A an arbitrary real matrix, there exists a unitary U such that U^*AU is upper triangular. The eigenvalues of A necessarily appear on the diagonal, but with arbitrary ordering. Thus although one talks of "the Schur form" in relation to U^*AU, it is not unique. When A is real, and U is restricted to being real orthogonal, one can find U so that $U'AU$ is block upper triangular, with the diagonal blocks being either 1×1, or 2×2 of the form

$$\begin{bmatrix} \alpha_i & -\beta_i \\ \beta_i & \alpha_i \end{bmatrix}$$

20. Positive and nonnegative definite matrices.
Suppose A is $n \times n$ and real and symmetric. Then A is termed positive definite, if for all nonzero real vectors x the scalar quantity $x'Ax$ is positive. Also, A is termed nonnegative definite if $x'Ax$ is simply nonnegative for all nonzero x. Negative definite and nonpositive definite are defined similarly. The quantity $x'Ax$ is termed a *quadratic form*, because when written as

$$x'Ax = \sum_{i,j=1}^{n} a_{ij} x_i x_j$$

it is quadratic in the entries x_i *of* x.

There are simple tests for positive and nonnegative definiteness. For A to be positive definite, all *leading minors* must be positive—that is,

$$a_{11} > 0 \qquad \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix} > 0 \qquad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{vmatrix} > 0 \text{ etc.}$$

For A to be nonnegative definite, all minors whose diagonal entries are diagonal entries of A must be nonnegative. That is, for a 3×3 matrix A,

$$a_{11}, a_{22}, a_{33} \geq 0 \qquad \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix}, \begin{vmatrix} a_{11} & a_{13} \\ a_{13} & a_{33} \end{vmatrix}, \begin{vmatrix} a_{22} & a_{23} \\ a_{23} & a_{33} \end{vmatrix} \geq 0$$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{vmatrix} \geq 0$$

A symmetric A is positive definite if and only if its eigenvalues are positive, and nonnegative definite if and only if its eigenvalues are nonnegative.

If D is an $n \times m$ matrix, then $A = DD'$ is nonnegative definite, and positive definite if and only if D has rank n. An easy way to see this is to define a vector y by $y = D'x$. Then $x'Ax = xDD'x = y'y = \Sigma\, y_i^2 \geq 0$. The inequality becomes an equality if and only if $y = 0$ or $D'x = 0$, which is impossible for nonzero x if D has rank n.

If A and B are nonnegative definite, so is $A + B$, and if one is positive definite, so is $A + B$. If A is nonnegative definite and $n \times n$, and B is $m \times n$, then BAB' is nonnegative definite.

If A is a symmetric matrix and λ_{max} is the maximum eigenvalue of A, then $\lambda_{max}I - A$ is nonnegative definite.

If A is nonnegative definite, there exists a matrix B that is a *symmetric square root* of A; it is also nonnegative definite. It has the property that

$$B^2 = A$$

and is often denoted by $A^{1/2}$. If A is positive definite, so is $A^{1/2}$, and $A^{1/2}$ is then unique.

Given a positive definite symmetric A, it is easy to construct a lower triangular B with positive diagonal entries such that

$$BB' = A$$

(The nonzero entries of B can be determined successively, beginning with the first row, then the second, third, and so forth working from the first entry through to the diagonal entry in each row). This construction is termed a *Cholesky decomposition* and B is termed a *Cholesky factor*. One can also demand that

$$B \Sigma B' = A$$

where Σ is diagonal positive definite, and B is lower triangular with 1's on the diagonal.

Virtually all the above notions remain valid with but minor change for positive definite hermitian matrices.

21. Singular value decomposition. Let A be a real or complex $n \times n$ matrix. The eigenvalues of A^*A are all real and nonnegative, and positive if A is nonsingular. The square roots $\lambda^{1/2}(A^*A)$ are termed the *singular values* of A. There exist unitary matrices U, V such that

$$UAV = \text{diag}[\lambda_i^{1/2}(A^*A)]$$

and if unitary U, V yield UAV to be a diagonal nonnegative definite matrix, the diagonal entries are necessarily the singular values of A.

22. Norms of vectors and matrices. The norm of a vector x, written $\|x\|$, is a measure of the size or length of x. There is no unique definition, but the following postulates must be satisfied.

1. $\|x\| \geq 0$ for all x with equality if and only if $x = 0$.
2. $\|ax\| = |a| \|x\|$ for any scalar a and for all x.
3. $\|x + y\| \leq \|x\| + \|y\|$ for all x and y.

If $x = (x_1, x_2, \ldots, x_n)$, three common norms are

$$\|x\| = \left[\sum_{i=1}^{n} x_i^2\right]^{1/2}, \qquad \|x\| = \max_i |x_i| \quad \text{and} \quad \|x\| = \sum_{i=1}^{n} |x_i|$$

the first being the Euclidean norm.

When the Euclidean norm is used, the Schwartz inequality states that $|x'y| \leq \|x\|\|y\|$ for arbitrary x and y; with equality if and only if $x = \lambda y$ for some scalar λ.

The norm of an $m \times n$ matrix A is defined in terms of an associated vector norm by

$$\|A\| = \max_{\|x\|=1} \|Ax\|$$

The particular vector norm used must be settled to fix the matrix norm. Corresponding to the three vector norms listed, the matrix norms become, respectively, $[\lambda_{\max}(A'A)]^{1/2}$, $\max_i (\sum_{j=1}^{n}|a_{ij}|)$ and $\max_j (\sum_{i=1}^{n}|a_{ij}|)$. Note that $[\lambda_{max}(A'A)]^{1/2}$ is the largest singular value of A. Important properties of matrix norms are

$$\|Ax\| \leq \|A\|\|x\|, \qquad \|A + B\| \leq \|A\| + \|B\|$$

and

$$\|AB\| \leq \|A\|\|B\|$$

23. Kronecker Product and Vec.   Let A, B be $m \times n$ and $p \times r$ matrices. The $mp \times nr$ matrix C, defined as

$$C = \begin{bmatrix} a_{11}B & \ldots & a_{1n}B \\ a_{21}B & \ldots & a_{2n}B \\ \cdot & & \\ \cdot & & \\ \cdot & & \\ a_{m1}B & \ldots & a_{mn}B \end{bmatrix}$$

and written $C = A \otimes B$ is termed the *Kronecker product* of A and B. In case A and B are square, the set of eigenvalues of C is given by $\lambda_i(A)\,\lambda_j(B)$ for all i, j. The Kronecker product is associative, $(A \otimes B)(C \otimes D) = AC \otimes BD$, and $(A \otimes B)' = A' \otimes B'$.

Let A be an $m \times n$ matrix. The column mn-vector, obtained by stacking column 2 of A after column 1, column 3 after column 2, and so forth, is termed vec A.

If M, N are matrices for which the product MN can be formed, then

$$\text{vec}(MN) = [I \otimes M]\,\text{vec}\,N$$

$$= [N' \otimes I]\,\text{vec}\,M$$

24. Linear matrix equations.   If A, B, and C are known matrices, of dimension $n \times n$, $m \times m$, and $n \times m$, respectively, we can form the following equation for an unknown $n \times m$ matrix X:

$$AX + XB + C = 0$$

This equation is merely a condensed way of writing a set of mn simultaneous equations for the entries of X. It is solvable to yield a unique X if and only if $\lambda_i(A) + \lambda_j(B) \neq 0$ for any i and j—that is, the sum of any eigenvalue of A and any eigenvalue of B is nonzero.

The vec operation yields

$$[I \otimes A + B' \otimes I] \text{vec}\, X = -\text{vec}\, C$$

and it can be shown that the eigenvalues of $[I \otimes A + B' \otimes I]$ are precisely the collection $\lambda_i(A) + \lambda_j(B)$.

If C is positive definite and $A = B'$, the *lemma of Lyapunov* states that X is positive definite symmetric if and only if all eigenvalues of B have negative real parts.

The equation

$$X - A \times B = C$$

is equivalent to

$$[I - B' \otimes A] \text{vec}\, X = \text{vec}\, C$$

and has a unique solution if and only if $\lambda_i(A)\,\lambda_j(B) \neq 1$ for any i, j. If $B = A'$ and $|\lambda_i(A)| < 1$ for all i, the equation has a solution for all C which is symmetric if C is symmetric, and positive definite if $C = DD'$ with $[A, D]$ completely controllable.

25. Strengthened version of Lemma of Lyapunov.
The Lemma of Lyapunov states that for positive definite C, there exists a unique positive definite P such that $PA + A'P + C = 0$ if and only if Re $\lambda_i(A) < 0$. The first strengthening states that if $[A, D]$ is completely observable, there exists a unique positive definite P such that $PA + A'P = -DD'$ if and only if Re $\lambda_i(A) < 0$. The second strengthening states that if $[A, D]$ is completely detectable, there exists a unique nonnegative P such that $PA + A'P = -DD'$ if and only if Re $\lambda_i(A) < 0$. In all cases where P exists,

$$P = \int_0^\infty e^{A't}DD'e^{At}\, dt$$

26. Matrix inversion lemma and block matrix inversion.   Suppose that A and C are nonsingular matrices (not necessarily of the same dimension), and B, D are such that $A + BCD$ can be formed and is nonsingular. Then

$$(A + BCD)^{-1} = A^{-1} - A^{-1}B(DA^{-1}B + C^{-1})^{-1}DA^{-1}$$

with $DA^{-1}B + C^{-1}$ guaranteed to be nonsingular.

As a special case, we have that if F is $n \times n$, G and K are $n \times m$, then

$$[I - K'(sI - F)^{-1}G]^{-1} = I + K'(sI - F - GK')^{-1}G$$

Suppose that the following matrix is invertible:

$$X = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

Suppose further that A^{-1}, D^{-1} exist. Then

$$X^{-1} = \begin{bmatrix} (A - BD^{-1}C)^{-1} & -(A - BD^{-1}C)^{-1}BD^{-1} \\ -D^{-1}C(A - BD^{-1}C)^{-1} & (D - CA^{-1}B)^{-1} \end{bmatrix}$$

and

$$(A - BD^{-1}C)^{-1}BD^{-1} = A^{-1}B(D - CA^{-1}B)^{-1}$$

27. Common differential equations involving matrices.

The equation

$$\frac{d}{dt}x(t) = A(t)x(t) \qquad x(t_0) = x_0$$

commonly occurs in system theory. Here, A is $n \times n$, and x is an n vector. If A is constant, the solution is

$$x(t) = \exp[A(t - t_0)]x_0$$

If A is not constant, the solution is expressible in terms of the solution of

$$\frac{dX(t)}{dt} = A(t)X(t) \qquad X(t_0) = I$$

where now X is an $n \times n$ matrix. The solution of this equation cannot normally be computed analytically, but is denoted by the *transition matrix* $\Phi(t, t_0)$, which has the properties

$$\Phi(t_0, t_0) = I \qquad \Phi(t_2, t_1)\Phi(t_1, t_0) = \Phi(t_2, t_0)$$

and

$$\Phi(t, t_0)\Phi(t_0, t) = I$$

The vector differential equation has solution

$$x(t) = \Phi(t, t_0)x_0$$

The solution of

$$\frac{dx(t)}{dt} = A(t)x(t) + B(t)u(t) \qquad x(t_0) = x_0$$

where $u(t)$ is a forcing term is

$$x(t) = \Phi(t, t_0)x_0 + \int_{t_0}^{t} \Phi(t, \tau)B(\tau)u(\tau)\,d\tau$$

The matrix differential equation

$$\frac{dX}{dt} = AX + XB + C(t) \qquad X(t_0) = X_0$$

also occurs commonly. With A and B constant, the solution of this equation may be written as

$$X(t) = \exp[A(t - t_0)]X_0 \exp[B(t - t_0)]$$

$$+ \int_{t_0}^{t} \exp[A(t - \tau)]C(\tau)\exp[B(t - \tau)]\,d\tau$$

A similar result holds when A and B are not constant.

When A, B, and C are constant and A, B have eigenvalues with negative real parts, then $X(t) \to \overline{X}$ as $t \to \infty$, where

$$\overline{X} = \int_{0}^{\infty} e^{At}Ce^{Bt}\,dt$$

and also

$$A\overline{X} + \overline{X}B + C = 0$$

The Lemma of Lyapunov formula (see 25 above) is a special case.

28. Several manipulative devices.   Let $f(A)$ be a function of A such that

$$f(A) = \sum_{i=0}^{\infty} a_i A^i$$

and $f(z)$, where z is a scalar, is analytic. Then,

$$T^{-1}f(A)T = f(T^{-1}AT)$$

This identity suggests one technique for computing $f(A)$, if A is diagonalizable. Choose T so that $T^{-1}AT$ is diagonal. Then $f(T^{-1}AT)$ is readily computed, and $f(A)$ is given by $Tf(T^{-1}AT)T^{-1}$. It also follows from this identity that the eigenvalues of $f(A)$ are $f(\lambda_i)$ where λ_i are eigenvalues of A; the eigenvectors of A and $f(A)$ are the same.

For n vectors x and y, and A any $n \times n$ matrix, the following trivial identity is often useful:

$$x'Ay = y'A'x$$

If A is $n \times m$, B is $m \times n$, I_m denotes the $m \times m$ unit matrix, and I_n the $n \times n$ unit matrix, then

$$|I_n + AB| = |I_m + BA|$$

If A is a column vector a and B a row vector b', then this implies

$$|I + ab'| = 1 + b'a$$

Next, if A is nonsingular and a matrix function of time, then

$$\frac{d}{dt}[A^{-1}(t)] = -A^{-1}\frac{dA}{dt}A^{-1}$$

(This follows by differentiating $AA^{-1} = I$.)

If $\Phi(t, t_0)$ is the transition matrix associated with $A(t)$, then

$$\frac{d}{dt}[|\Phi(t, t_0)|] = \mathrm{tr}\,[A(t)]\,|\Phi(t, t_0)|$$

$$|\Phi(t, t_0)| = \exp\left[\int_{t_0}^{t} \mathrm{tr}\,[A(s)]\,ds\right]$$

If P is an $n \times n$ symmetric matrix, we note the value of grad $(x'Px)$, often written just $(\partial/\partial x)(x'Px)$, where the use of the partial derivative occurs since P may depend on another variable, such as time. As may be easily checked by writing each side in full,

$$\frac{\partial}{\partial x}(x'Px) = 2Px$$

The derivative $\frac{\partial\varphi}{\partial X}$ of a scalar function φ of a matrix X is readily defined, as a matrix with i-j element $\frac{\partial\varphi}{\partial x_{ij}}$. If X is square and nonsingular,

$$\frac{\partial}{\partial X}\log|X| = X^{-1}$$

and

$$\frac{\partial}{\partial X}(\mathrm{tr}\,WX^{-1}) = -X^{-1}WX^{-1}$$

If X is square, nonsingular and $n \times n$,

$$\log|X| \le \mathrm{tr}\,X - n$$

with equality if and only if $X = I$.

If X, Y are symmetric, nonnegative definite and $n \times n$,

$$\mathrm{tr}\,X\,\mathrm{tr}\,Y \ge \left[\sum_i \lambda_i^{1/2}(XY)\right]^2$$

$$\mathrm{tr}\,X + \mathrm{tr}\,Y \ge 2\left[\sum_i \lambda_i^{1/2}(XY)\right]$$

with equality in the first case when $X = \rho Y$, ρ scalar and in the second case when $X = Y$.

REFERENCES

[1] F. R. Gantmacher, *The Theory of Matrices*, Vols. 1 and 2. New York: Chelsea Publishing Co., 1959.

[2] S. Barnett, *Matrices in Control Theory*. London: Van Nostrand Reinhold Company, 1971.

[3] R. E. Bellman, *Introduction to Matrix Analysis*, 2nd ed. New York: McGraw-Hill Book Company, 1970.

B

Brief Review
of Some Major Results
of Linear System Theory

This appendix provides a summary of several facts of linear system theory. A basic familiarity is, however, assumed. Source material may be found in, for example, [1] through [3].

1. Passage from state-space equations to transfer function matrix. In system theory, the equations

$$\dot{x} = Fx + Gu$$

$$y = H'x + Ju$$

frequently occur. The Laplace transform may be applied in the same manner as to scalar equations to yield

$$sX(s) = FX(s) + X(0) + GU(s)$$

$$Y(s) = H'X(s) + JU(s)$$

whence

$$Y(s) = [H'(sI - F)^{-1}G + J]U(s)$$

with $x(0) = 0$. The transfer function matrix relating $U(s)$ to $Y(s)$ is $J + H'(sI - F)^{-1}G$.

In discrete time, one has

$$x_{k+1} = Fx_k + Gu_k$$

$$y_k = H'x_k + Ju_k$$

and with $X(z) = \sum_{k=0}^{\infty} x_k z^{-k}$, and so on,

$$zX(z) = FX(z) + x(0) + GU(z)$$

$$Y(z) = H'X(z) + JU(z)$$

Thus

$$Y(z) = [J + H'(zI - F)^{-1}G]U(z)$$

when $x(0) = 0$.

The impulse response in continuous time is $H' \exp (Ft)G$ for $t \geq 0$ and the impulse response sequence $\{w_k\}$ in discrete time is $H'F^{k-1}G$ for $k \geq 1$, J for $k = 0$.

A transfer function matrix that is finite for $s = \infty$ or $z = \infty$ is termed proper. It corresponds to a causal impulse response. If the transfer function matrix is zero at $s = \infty$ or $z = \infty$, it is termed strictly proper.

2. Conditions for complete controllability and observability.

A pair of constant matrices $[F, G]$ with F $n \times n$ and G $n \times r$ is termed completely controllable if the following equivalent conditions hold:

1. Rank $[G \ FG \cdots F^{n-1}G] = n$.
2. $w'e^{Ft}G = 0$ for all t implies $w = 0$.
3. $\int_0^T e^{Ft}GG'e^{Ft} \, dt$ is positive definite for all $T > 0$.
4. There exists an $n \times r$ matrix K such that the eigenvalues of $F + GK'$ can take on arbitrary prescribed values.
5. Given the system $\dot{x} = Fx + Gu$, arbitrary states x_0, x_1, and arbitrary times t_0, t_1, with $t_0 < t_1$, there exists a control taking the system from state x_0 at t_0 to state x_1 at t_1. [In contrast to (1) through (4), this is also valid for time-varying F and G if $t_0 = t_0(t_1, x_0, x_1)$.]
6. There exists no complex λ and nonzero complex n-vector w such that $w^*[\lambda I - F \quad G] = 0$.
7. There exists no state coordinate basis change such that

$$F = \begin{bmatrix} F_{11} & F_{12} \\ 0 & F_{22} \end{bmatrix} \qquad G = \begin{bmatrix} G_1 \\ 0 \end{bmatrix}$$

   with F_{22} of nontrivial dimension.
8. $w'F^iG = 0$ for all i implies $w = 0$.

A pair of matrices $[F, H]$ with F $n \times n$ and H $n \times r$ is termed completely observable if $[F', H]$ is completely controllable.

Complete controllability is preserved under state variable feedback; that is, $[F, G]$ is completely controllable if and only if $[F + GK', G]$ is completely controllable. Likewise, $[F, H]$ is completely observable if and only if $[F + LH', H]$ is completely observable.

3. Complete stabilizability and detectability.
The pair $[F, G]$ is completely stabilizable if all uncontrollable modes are asymptotically

stable. More precisely, complete stabilizability is equivalent to any of the following: (The continuous time version only will be stated)

1. $w^*[\lambda I - F \quad G] = 0$ for some $w^* \neq 0$ implies $\text{Re } \lambda < 0$.
2. There exists a coordinate basis change such that

$$F = \begin{bmatrix} F_{11} & F_{12} \\ 0 & F_{22} \end{bmatrix} \qquad G = \begin{bmatrix} G_1 \\ 0 \end{bmatrix}$$

with the pair $[F_{11}, G_1]$ completely controllable, and, if F_{22} has nontrivial dimension, $\text{Re } \lambda_i(F_{22}) < 0$ for all i.
3. There exists a K such that

$$\text{Re } \lambda_i(F + GK') < 0 \ \forall i$$

Complete detectability is the dual, so that $[F, H]$ is completely detectable if and only if $[F', H]$ is completely stabilizable.

As with controllability and observability, $[F, G]$ is stabilizable if and only if $[F + GK, G]$ is stabilizable, and similarly for detectability.

4. Minimality.   If a transfer function matrix $W(s)$ is related to a matrix triple F, G, H by

$$W(s) = H'(sI - F)^{-1}G,$$

then F has minimal dimension if and only if $[F, G]$ is completely controllable and $[F, H]$ is completely observable. The triple F, G, H is termed a minimal realization of $W(s)$. Given two minimal realizations of $W(s)$—call them F_1, G_1, H_1, and F_2, G_2, H_2—there always exists a nonsingular T such that

$$TF_1T^{-1} = F_2 \qquad TG_1 = G_2 \qquad (T^{-1})'H_1 = H_2$$

5. Passage from transfer function matrix to state space equations.   The determination of state-space equations corresponding to a transfer function, as distinct from transfer function matrix, is straightforward. Given a transfer function

$$W(s) = \frac{b_n s^{n-1} + b_{n-1}s^{n-2} + \cdots + b_1}{s^n + a_n s^{n-1} + \cdots + a_1}$$

state-space equations $\dot{x} = Fx + gu, y = h'x$ yield the same transfer function relating $U(s)$ to $Y(s)$ if

$$F = \begin{bmatrix} 0 & 1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & 1 & & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & & & \cdot & \cdot \\ \cdot & \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & \cdot & & & & 1 \\ -a_1 & -a_2 & -a_3 & & \cdot & \cdot & -a_n \end{bmatrix} \qquad g = \begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 1 \end{bmatrix} \qquad h = \begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_n \end{bmatrix}$$

or

$$
F = \begin{bmatrix} 0 & 0 & \cdot & \cdot & \cdot & -a_1 \\ 1 & 0 & \cdot & \cdot & \cdot & -a_2 \\ 0 & 1 & \cdot & \cdot & \cdot & -a_3 \\ \cdot & \cdot & \cdot & & & \cdot \\ \cdot & \cdot & \cdot & & & \cdot \\ \cdot & & \cdot & \cdot & 1 & -a_n \end{bmatrix} \qquad g = \begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_n \end{bmatrix} \qquad h = \begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 1 \end{bmatrix}
$$

These formulas are valid irrespective of whether the numerator and denominator of $W(s)$ have common factors. The first pair of F and g is completely controllable, and there exists a coordinate basis transformation taking any other set of $F, g,$ and h which are completely controllable to the prescribed form. The second pair of F and h is completely observable, and there exists a coordinate basis transformation taking any other completely observable set to the prescribed form.

If the numerator and denominator of $W(s)$ have no common factor, both sets are simultaneously completely controllable and observable.

When $W(s)$ is no longer scalar, procedures for securing a minimal realization are much more complicated. See, for example, [1] and [2] for a collection of methods, including a discussion of the multivariable generalization of the above canonical forms.

6. Hankel Matrix, Markov Matrix Parameters, and Realizations.

When $W(s)$ is a matrix—say, $p \times m$—an algorithm due to Ho [4] provides a convenient route to determining matrices $F, G,$ and H given certain data concerning $W(s)$. First, $W(s)$ is assumed to be strictly proper, that is, zero at $s = \infty$. It is then expanded as

$$
W(s) = \frac{A_0}{s} + \frac{A_1}{s^2} + \frac{A_2}{s^3} + \cdots
$$

where the A_i are termed *Markov matrices*. In discrete time, the Markov matrices are nothing but the entries of the impulse response matrix, that is, $A_k = w_{k+1}$ where $W(z) = \Sigma\, z^{-k} w_k$. Then the A_i are arranged to form truncated *Hankel* matrices H_N as follows:

$$
H_N = \begin{bmatrix} A_0 & A_1 & \cdot & \cdot & \cdot & A_{N-1} \\ A_1 & A_2 & & & & A_N \\ \cdot & & & & & \cdot \\ \cdot & & & & & \cdot \\ \cdot & & & & & \cdot \\ A_{N-1} & A_N & \cdot & \cdot & \cdot & A_{2N-1} \end{bmatrix}
$$

The next step requires the checking of the ranks of H_N for different N, to determine the first integer r such that rank $H_r = $ rank $H_{r+1} = $ rank $H_{r+2} = \cdots$. If $W(s)$ is rational, there always exists such an r, and it is a classical result (at least for the scalar case) that $r = $ dimension of minimal realization of $W(s)$.

A realization can be constructed as follows. Nonsingular matrices P and Q are found so that

$$PH_rQ = \begin{bmatrix} I_n & 0 \\ 0 & 0 \end{bmatrix}$$

where n = rank H_r. The following matrices realize $W(s)$, in the sense that $W(s) = H'(sI - F)^{-1}G$:

$$G = n \times m \text{ top left corner of } PH_r$$
$$H' = p \times n \text{ top left corner of } H_rQ$$
$$F = n \times n \text{ top left corner of } P(\sigma H_r)Q$$

where

$$\sigma H_r = \begin{bmatrix} A_1 & A_2 & \cdot & \cdot & \cdot & A_r \\ A_2 & A_3 & & & & A_{r+1} \\ \cdot & & & & & \cdot \\ \cdot & & & & & \cdot \\ \cdot & & & & & \cdot \\ A_r & A_{r+1} & \cdot & \cdot & \cdot & A_{2r-1} \end{bmatrix}$$

Moreover, $[F, G]$ is completely controllable and $[F, H]$ is completely observable.

7. Controllability and observability gramians.   Suppose $W(s) = H'(sI - F)^{-1}G$. The matrices

$$P(0, T) = \int_0^T e^{Ft}GG'e^{F't}dt$$

$$Q(0, T) = \int_0^T e^{F't}HH'e^{Ft}dt$$

are termed controllability and observability gramians, and are nonsingular when $\{F, G, H\}$ is minimal for all $T > 0$. If Re $\lambda_i(F) < 0$, the gramians are defined for $T = \infty$, and also are given as the solutions of

$$FP + PF' + GG' = 0$$

$$QF + F'Q + HH' = 0$$

8. Balanced realization of a stable transfer function matrix.   Let $W(s) = H'(sI - F)^{-1}G$ with $\{F, G, H\}$ minimal and Re $\lambda_i(F) < 0$. A realization in which the (infinite time) gramians satisfy

$$P = Q = \text{diag}[\sigma_1, \sigma_2, \ldots, \sigma_n]$$

is termed a *balanced realization*. Generally, $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$, and if strict inequality applies, the balanced realization is unique. If $W(s) = \overline{H}'(sI - \overline{F})^{-1}\overline{G}$ is minimal but not balanced, standard procedures allow construction of a coordinate

basis change to balanced coordinates. The quantities σ_i are determined by $W(s)$ alone, rather than the particular realization, and are also the singular values of the $n \times n$ Hankel matrix defined using the Markov parameters of $W(s)$. They are, therefore, known as the *Hankel singular values*.

If $\{F, G, H\}$ is balanced and

$$F = \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} \qquad G = \begin{bmatrix} G_1 \\ G_2 \end{bmatrix} \qquad H' = [H_1' \quad H_2']$$

with F_{11} $r \times r$, G_1, and H_1 possessing r rows, and if $\sigma_i \geq \sigma_{i+1}$ for all i with $\sigma_r > \sigma_{r+1}$, then

$$W_1(s) = H_1'(sI - F_{11})^{-1}G_1$$

is termed an rth degree balanced approximation of $W(s)$; further Re $\lambda_i(F_{11}) < 0$ and for all real ω

$$\|W(j\omega) - W_1(j\omega)\| \leq 2(\sigma_{r+1} + \cdots + \sigma_n)$$

(The Euclidean norm applies.)

9. Poles and transmission zeros.   Suppose that $W(s) = J + H'(sI - F)^{-1}G$. Any pole of $W(s)$ is necessarily an eigenvalue of F, and conversely if $\{F, G, H\}$ is minimal. A transmission zero of the realization $\{F, G, H, J\}$ is an s_0 for which

$$\text{rank} \begin{bmatrix} s_0I - F & -G \\ H & J \end{bmatrix} < \max_s \text{rank} \begin{bmatrix} sI - F & -G \\ H & J \end{bmatrix}$$

and if $\{F, G, H, J\}$ is minimal, and s_0 is not a pole of $W(s)$, rank $W(s_0) < \max_s$ rank $W(s)$. In case $W(s)$ is scalar, this just says that $W(s_0)$ is zero, and accords with convention. For matrix $W(s)$, $W(s_0)$ drops rank. Thus for a certain input signal $\underline{u} \exp(s_0 t)$, where $\underline{u}$ is a null vector of $W(s_0)$, the steady state output from $W(\cdot)$ will be zero. If $W(s)$ is not scalar, it is possible to have s_0 which are both poles of $W(s)$ and zeros of a minimal realization. If W is square and nonsingular almost everywhere, the transmission zeros of a minimal realization of W are the poles of W^{-1}. If they are all stable, one terms W minimum phase. In case $W(s)$ is not square, and F, G, H, J are generic, there are generically no zeros. In particular, if $W(s) = (sI - F)^{-1}G$, and $[F, G]$ is controllable, there are no zeros.

10. Time domain response.   The solution of

$$\dot{x} = Fx + Gu \qquad y = H'x$$

is

$$y(t) = H'e^{Ft}x(0) + H' \int_0^t [\exp F(t - \tau)]Gu(\tau)\,d\tau$$

If F has the so-called modal form, that is, is a real block diagonal Jordan form, or is a direct sum of 1×1 or 2×2 matrices, the matrices $\exp F(t - \tau)$ are readily computable for different t and τ. With Δ a sampling time, the differential equation can be approximated by

$$x[(k + 1) \Delta] = \exp(F\Delta)x(k\Delta) + \int_0^\Delta \exp(Ft)\,dt\,Gu(k\Delta)$$

$$y(k\Delta) = H'x(k\Delta)$$

and the matrix multiplying $u(k\Delta)$ can be readily computed in case F has real block diagonal Jordan form.

11. Coprime matrix fraction description. It is sometimes convenient to represent a rational transfer function matrix as a left or right quotient of stable, proper transfer function matrices. Thus

$$W(s) = A_L^{-1}(s)B_L(s) = B_R(s)A_R^{-1}(s)$$

where $A_L(s)$, and so on have entries which are finite as $s \to \infty$, and have all poles in Re $(s) < 0$. A left matrix fraction description is termed left coprime if and only if there exist stable proper $X_R(s)$, $Y_R(s)$ such that a Bezout identity

$$A_L(s)X_R(s) + B_L(s)Y_R(s) = I$$

holds, and similarly the right matrix fraction description is termed coprime if and only if there exist stable proper $X_L(s)$, $Y_L(s)$ such that

$$X_L(s)A_R(s) + Y_L(s)B_R(s) = I$$

If

$$W(s) = A_{L1}^{-1}(s)B_{L1}(s) = A_{L2}^{-1}(s)B_{L2}(s)$$

are two coprime realizations, there exists a nonsingular $Z(s)$ such that $Z(s)$ and $Z^{-1}(s)$ are stable and proper, for which

$$Z(s)A_{L1}(s) = A_{L2}(s)$$

$$Z(s)B_{L1}(s) = B_{L2}(s)$$

Suppose that $W(s) = H'(sI - F)^{-1}G$ with $\{F, G, H\}$ minimal. Let K, K_e be such that $F + GK'$ and $F + K_eH'$ have all eigenvalues with negative real parts. Then suitable factorizations are

$$A_R = I + K'[sI - (F + GK')]^{-1}G \qquad B_R = H'[sI - (F + GK')]^{-1}G$$

$$A_L = I + H'[sI - (F + K_eH')]^{-1}K_e \qquad B_L = H'[sI - (F + K_eH')]^{-1}G$$

Also, if

$$X_L = I - K'[sI - (F + K_eH')]^{-1}G \qquad Y_L = K'[sI - (F + K_eH')]^{-1}K_e$$

$$X_R = I - H'[sI - (F + GK')]^{-1}K_e \qquad Y_R = K'[sI - (F + GK')]^{-1}K_e$$

there holds the double Bezout identity [5]:

$$\begin{bmatrix} X_L & Y_L \\ -B_L & A_L \end{bmatrix}\begin{bmatrix} A_R & -Y_R \\ B_R & X_R \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$$

12. Class of all stabilizing controllers.

Let $P(s) = A_L^{-1}(s)B_L(s) = B_R(s)A_R^{-1}(s)$ with A_L, B_L, A_R, B_R stable, proper coprime factorizations. Let $C(s)$ be a transfer function matrix of a proper stabilizing negative feedback controller. The control loop is well posed and stable if and only if

$$\begin{bmatrix} I & C \\ -P & I \end{bmatrix}^{-1} \quad \text{exists and is stable}$$

Equivalently, $(I + PC)^{-1}$, $C(I + PC)^{-1}$, $(I + PC)^{-1}P$ and $(I + CP)^{-1}$ are all stable proper transfer function matrices. Under the conditions above, there exists a right coprime realization $Y_R(s)X_R^{-1}(s)$ for $C(s)$ such that $A_L X_R + B_L Y_R = I$ and left coprime realization $X_L^{-1}(s)Y_L(s)$ such that $X_L A_R + Y_L B_R = I$. Furthermore, the set of all stabilizing controllers for $P(s)$ is given by

$$C(Q) = Y_R(Q)X_R^{-1}(Q) = X_L^{-1}(Q)Y_L(Q)$$

where Q ranges over the set of all proper stable transfer functions and

$$Y_R(Q) = Y_R - A_R Q, \qquad X_R(Q) = X_R + B_R Q$$

$$Y_L(Q) = Y_L - Q A_L, \qquad X_L(Q) = X_L + Q B_L$$

Exploiting the Bezout relationship, another formulation is

$$C(Q) = C - X_L^{-1}Q(I + X_R^{-1}B_R Q)^{-1}X_R^{-1}$$

The linear fractional maps between Q and $C(Q)$ are bijective, and moreover

$$Q = X_L(Q)[C - C(Q)]X_R$$
$$= X_L[C - C(Q)]X_R(Q)$$

13. Stable linear system excitation by white noise.

Consider the linear system

$$\dot{x} = Fx + Gu \qquad y = H'x$$

in which $u(\cdot)$ is a zero mean stationary white noise process, with $E[u(t)u'(\tau)] = Q\delta(t - \tau)$. Suppose Re $\lambda_i(F) < 0$. Then $E[x(t)\ x'(t)] = P$ where P is the unique solution of

$$PF' + FP + GQG' = 0$$

and $E[x(t)x'(s)]$ is given by $[\exp F(t - s)]P$ if $t \geq s$ and $P \exp F'(s - t)$ if $s \geq t$. Further,

$$E[y(t)y'(s)] = H' \exp F(t-s)PH \qquad t \geq s$$
$$= H'P \exp F'(s-t)H \qquad t \leq s$$

The spectrum matrix of $y(\cdot)$, which is the (two-sided) Fourier transform of the covariance $E[y(t)\,y'(0)]$ is

$$\Phi_{yy}(j\omega) = H'(j\omega I - F)^{-1}PH + HP(-j\omega I - F')^{-1}H$$
$$= H'(j\omega I - F)^{-1}GQG'(-j\omega I - F')^{-1}H$$
$$= W(j\omega)QW'(-j\omega)$$

where $W(s) = H'(sI - F)^{-1}G$.

Setting $t = 0$ in the inverse transform leads to the important result

$$E[y(0)y'(0)] = \frac{1}{2\pi} \int_{-\infty}^{+\infty} W(j\omega)QW'(-j\omega)\,d\omega$$

while this expression is also $H'PH$, in view of the formula for $E[y(t)y'(s)]$. Since P, the solution of the Lyapunov equation, is expressible as

$$P = \int_0^{\infty} e^{Ft}GQG'\,e^{F't}\,dt$$

there holds

$$E[y(0)y'(0)] = \int_0^{\infty} H'\,e^{Ft}\,GQG'\,e^{F't}H\,dt$$

which relates to the frequency domain formula by Parseval's theorem.

In case F, G, Q are time-varying, then the state covariance P is time-varying and satisfies

$$\dot{P} = PF' + FP + GQG' \qquad P(t_0) = E[x(t_0)x'(t_0)]$$

14. Controllability, etc. for time-varying systems. The
system $\dot{x} = F(t)x + G(t)u$ is controllable at time t_0 if, given arbitrary $x(t_0)$, there exists a control such that for some t_1, $x(t_1) = 0$ is secured. This will hold if and only if for some t_1, the following controllability gramian is nonsingular:

$$W_c(t_0, t_1) = \int_{t_0}^{t_1} \Phi(t_1, s)G(s)G'(s)\Phi'(t_1, s)\,ds$$

The property is invariant under state variable feedback. If $F(t)$ and $G(t)$ are bounded, the pair is termed uniformly completely controllable if for some $\Delta > 0$ and $\alpha_1 > 0$, there holds $W_c(t_0, t_0 + \Delta) > \alpha_1 I$ for all t_0. This property is invariant under bounded state feedback. A uniformly controllable system has the property that there exists a bounded $K(t)$ such that $\dot{x} = (F + GK')x$ has arbitrary degree of stability; that is, given $\alpha > 0$, one can find $K(t)$ such that $[\exp \alpha t]x(t) \to 0$ for all $x(t_0)$.

REFERENCES

[1] C. T. Chen, *Introduction to Linear System Theory*. New York: Holt Rinehart and Winston, Inc., 1970.

[2] T. Kailath, *Linear Systems*. Englewood Cliffs, N.J.: Prentice Hall, Inc., 1978.

[3] L. A. Zadeh and C. A. Desoer, *Linear System Theory*. New York: McGraw-Hill Book Company, 1963.

[4] B-L. Ho and R. E. Kalman, "Effective Construction of Linear State-Variable Models from Input/Output Functions," *Regelungstechnik*, Vol. 14, No. 12 (1966), pp. 545–548.

[5] C. N. Nett, C. A. Jacobson, and M. J. Balas, "A Connection between State-Space and Doubly Coprime Fractional Representations," *IEEE Trans. Auto. Control*, Vol. AC-29, No. 9 (September 1984), pp. 831–832.

C

The Pontryagin Minimum Principle and Linear Optimal Control

For the sake of completeness, we give treatment in this appendix, albeit in outline, of the regulator problem using the Pontryagin Minimum Principle. Other references include [1], [2].

1. General form of the Minimum Principle.

Consider the system

$$\dot{x} = f(x, u, t) \tag{C1}$$

and performance index

$$V(x(0), u(\cdot)) = \int_0^T l(x(\tau), u(\tau), \tau)\, d\tau + m[x(T)] \tag{C2}$$

Define, with p termed the costate vector,

$$H(x, u, t, p) = p'f(x, u, t) + l(x, u, t) \tag{C3}$$

and

$$H^*(x, t, p) = \min_{u = u(x, t, p)} H(x, u, t, p) \tag{C4}$$

(assuming the minimum exists and at the minimum, $\dfrac{\partial H}{\partial u} = 0$). Then the equations

$$\dot{x} = \frac{\partial H^*}{\partial p} \qquad x(0) \text{ prescribed} \tag{C5a}$$

$$\dot{p} = -\frac{\partial H^*}{\partial x} \qquad p(T) = \frac{\partial m}{\partial \mathbf{x}}\bigg|_{\mathbf{x}(T)} \tag{C5b}$$

are satisfied along the optimal trajectory, and if $x^*(\cdot)$, $p^*(\cdot)$ denote the solution of (C5) corresponding to the optimal trajectory, the optimal $u^*(\cdot)$ is

$$u^*(t) = \arg\min H[x^*(t), u, t, p^*(t)] \tag{C6}$$

(Smoothness assumptions are omitted.) Note that (C5) are coupled ordinary differential equations with two-point boundary conditions. Also, (C6) does not yield an optimal feedback control law, but an optimum open-loop control (time function).

2. Specialization of Minimum Principle equations to the linear quadratic problem.

Take $f(x, u, t) = F(t)x + G(t)u$, $l(x, u, t) = \frac{1}{2}[u'R(t)u + x'Q(t)x]$ (with $R = R' > 0$ and $Q = Q' \geq 0$) and $m[x(T)] = \frac{1}{2}x'(T)Ax(T)$. (The $\frac{1}{2}$ is to yield a "tidier" result). Then

$$H(x, u, t, p) = p'F(t)x + p'G(t)u + \tfrac{1}{2}u'R(t)u + \tfrac{1}{2}x'Q(t)x \tag{C7}$$

$$H^*(x, t, p) = p'F(t)x - \tfrac{1}{2}p'G(t)R^{-1}(t)G'(t)p + \tfrac{1}{2}x'Q(t)x \tag{C8}$$

$$u^*(t) = -R^{-1}(t)G'(t)p^*(t) \tag{C9}$$

$$\dot{x}^* = F(t)x^* - G(t)R^{-1}(t)G'(t)p^* \qquad x(0) \text{ prescribed} \tag{C10}$$

$$\dot{p}^* = -Q(t)x^* - F'(t)p^* \qquad p(T) = Ax(T) \tag{C11}$$

Were it not for the awkward boundary condition in (C10) and (C11), one could in principle solve these equations and use the solution in (C9) to obtain the open-loop optimal control.

3. Solving the Minimum Principle equations in the linear-quadratic case.

With hindsight, define $P(t)$ as the solution of

$$-\dot{P} = PF + F'P - PGR^{-1}G'P + Q \qquad P(T) = A \tag{C12}$$

on $[0, T]$. One can then verify, using (C10) and (C11), that

$$p^*(t) = P(t)x^*(t)$$

for $t \epsilon [0, T]$. (Differentiate $p^*(t) - P(t)x^*(t)$, and use (C10) through (C12) to show it is zero). It follows from (C9) that, as with the Hamilton-Jacobi approach,

$$u^*(t) = -R^{-1}(t)G'(t)P(t)x^*(t)$$

One could, if desired, work out $p^*(t)$ as an explicit function of time, and then $u^*(t)$ as an explicit function of time—but the beauty of the trick in (C12) is that $u^*(t)$ is given as a feedback law.

Deeper theory of Riccati equations throws up an intimate connection between (C12) and the linear equation

$$\begin{bmatrix} \dot{X} \\ \dot{Y} \end{bmatrix} = \begin{bmatrix} F(t) & -G(t)R^{-1}(t)G'(t) \\ -Q(t) & -F'(t) \end{bmatrix} \begin{bmatrix} X(t) \\ Y(t) \end{bmatrix} \tag{C13}$$

where $X(t)$, $Y(t)$ are square matrices. Notice that (C13) is nothing but a rewriting of (C10) and (C11) with x, p replaced by $X(t)$, $Y(t)$.

4. First variation in performance index.

With smooth f, l and m, the Pontryagin principle is obtainable by studying first order variations in V resulting from first-order variations in u, and requiring that these be zero at the optimum. Following the concept of Lagrange multipliers in constrained optimization problems, let $p(t)$ be an n-vector function and seek to minimize, with respect to $u(\cdot), p(\cdot)$

$$\hat{V}(x(0), u(\cdot), p(\cdot)) = \int_0^T \{l(x, u, t) + p'[f(x, u, t) - \dot{x}]\} dt + m[x(T)]$$

$$= \int_0^T [H(x, u, t, p) - p'\dot{x}] dt + m[x(T)] \tag{C14}$$

subject to $\dot{x} - f(x, u, t) = 0$. Let δx, δu denote deviations from the optimum trajectory and control, and $\delta\hat{V}$ the corresponding variation. Then

$$\delta\hat{V} = \int_0^T [H_x'\delta x + H_u'\delta u - p'\delta\dot{x}] dt + m_x\delta x|_T$$

$$= \int_0^T [H_x'\delta x + H_u'\delta u - \dot{p}'\delta x] - p'\delta x|_T + p'\delta x|_0 + m_x'\delta x|_T \tag{C15}$$

Requiring the first-order variation to be zero yields (after some nontrivial argument)

$$H_u = 0 \qquad \dot{p} = -H_x \qquad p(T) = m_x(x(T)) \tag{C16}$$

while the constraint $\dot{x} = f(x, u, t)$ is equivalent to $\dot{x} = H_p$. Let $u = u(x, t, p)$ minimize H (i.e., solve $H_u = 0$), and let H^* be the minimized H as in (C4). Then

$$\left(\frac{\partial H^*}{\partial p}\right)' = \left(\frac{\partial H}{\partial p}\right)' + \left(\frac{\partial H}{\partial u}\right)' \frac{\partial u}{\partial p}\bigg|_{u=u(x, t, p)} = \left(\frac{\partial H}{\partial p}\right)'$$

$$\left(\frac{\partial H^*}{\partial x}\right)' = \left(\frac{\partial H}{\partial x}\right)' + \left(\frac{\partial H}{\partial u}\right)' \frac{\partial u}{\partial x}\bigg|_{u=u(x, t, p)} = \left(\frac{\partial H}{\partial x}\right)'$$

and so (C5) result.

5. Second variation in performance index.

Suppose we have been able to solve for the optimal control $u^*(t)$ and trajectory $x^*(t)$ for the general problem of (C1) and (C2), assuming smooth f, l, and m. A first-order perturbation in the control yields a first-order perturbation in V which is zero, but a second-order perturbation that must be nonnegative. This second-order perturbation is

$$\delta^2 V = \frac{1}{2}\int_0^T [\delta x \quad \delta u]' \begin{bmatrix} H_{xx} & H_{xu} \\ H_{ux} & H_{uu} \end{bmatrix} \begin{bmatrix} \delta x \\ \delta u \end{bmatrix} dt$$

$$+ \frac{1}{2}\delta x'(T)m_{xx}|_{x(T)} \delta x(T)$$

REFERENCES

[1] F. L. Lewis, *Optimal Control.* New York: John Wiley and Sons, 1986.
[2] M. Athans and P. L. Falb, *Optimal Control.* New York: McGraw-Hill, 1966.

D

Lyapunov Stability

---

1. Stability definitions. Lyapunov theory is a technique for studying the stability of free or unforced equations; see [1, 2]. Consider

$$\dot{x} = f(x, t) \qquad\qquad (D1)$$

in which it is assumed that $f(0, t) \equiv 0$, so that $x_e = 0$ is an equilibrium state, and that solutions of (D1) are defined on $[t_0, \infty)$ for all t_0 and $x(t_0)$ of interest. [The set of $x(t_0)$ normally includes a ball containing the origin].

> *Definition.* The equilibrium state $x_e = 0$ is *stable* if for arbitrary t_0 and $\epsilon > 0$, there exists a $\delta(\epsilon, t_0)$ such that $\|x(t_0)\| < \delta$ implies $\|x(t)\| < \epsilon$ for all $t \geq t_0$.

The idea is that one can keep the entire trajectory close to zero by starting off close enough to zero.

> *Definition.* The equilibrium state $x_e = 0$ is *asymptotically stable* if it is stable and if a convergence condition holds: for arbitrary t_0, there exists $\delta_1(t_0)$ such that $\|x(t_0)\| < \delta_1(t_0)$ implies $\|x(t)\| \to \infty$.

Important specializations occur allowing definition of *uniform stability* and *uniform asymptotic stability*: δ and δ_1 must be selectable independently of t_0. If (D1) is autonomous, that is,

$$\dot{x} = f(x) \qquad\qquad (D2)$$

then stability and asymptotic stability (if they hold) are automatically uniform.

Global asymptotic stability arises when δ_1 can be taken arbitrarily large. Exponential asymptotic stability arises when $\|x(t_0)\| < \delta_1(t_0)$ implies additionally

$$\|x(t)\| \leq K\|x(t_0)\| \exp[-\alpha(t - t_0)]$$

for some positive α and K.

2. Lyapunov theorem for autonomous systems (D2).

Let $V(x)$ be a real scalar function of the n-vector x and Ξ be a closed bounded region in $\mathbb{R}^n$ containing the origin.

> **Definition.** $V(x)$ is positive definite (semidefinite) in Ξ, written $V > 0$ ($V \geq 0$), if $V(0) = 0$, $V(x) > 0$ ($V(x) \geq 0$) for all $x \neq 0$ in Ξ. $T(x)$ is negative definite (semidefinite) if and only if $-T(x)$ is positive definite (semidefinite).

> **Theorem 2.1**   (Stability). If there exists in some Ξ containing the origin a positive definite $V(x)$ with $0 \geq \dot{V} = (\text{grad } V)f(x)$, the derivative of V along trajectories of (D2), then $x_e = 0$ is stable.

> **Theorem 2.2**   (Asymptotic Stability). If there exists in some Ξ containing the origin a positive definite $V(x)$ with $\dot{V}(x)$ negative definite, then $x_e = 0$ is asymptotically stable.

> **Theorem 2.3**   (Asymptotic Stability). If there exists in some Ξ containing the origin a positive definite $V(x)$ with $\dot{V}(x) \leq 0$ and with $\dot{V}$ not identically zero along any trajectory except $x \equiv 0$, then $x_e = 0$ is asymptotically stable.

> **Theorem 2.4**   (Global Asymptotic Stability). If in Theorems 2.2 or 2.3, $\Xi = \mathbb{R}^n$ and $V(x) \to \infty$ as $\|x\| \to \infty$, then $x_e = 0$ is globally asymptotically stable.

> **Theorem 2.5**   (Exponential Asymptotic Stability). If in Theorem 2.2 one has $\alpha_1\|x\|^2 \leq V(x) \leq \alpha_2\|x\|^2$ and $-\alpha_3\|x\|^2 \leq \dot{V}(x) \leq -\alpha_4\|x\|^2$ for some positive α_i, then $x_e = 0$ is exponentially asymptotically stable.

A function $V(x)$ which allows proof of a stability result via one of these theorems is termed a Lyapunov function.

3. Lyapunov theory for time-varying systems (D1).

Slight modifications are needed. We consider real scalar functions $V(x,t)$ of the n-vector x and time t in a closed bounded region Ξ containing the origin.

> **Definition.** $V(x, t)$ is positive definite in Ξ, written $V > 0$, if $V(0, t) = 0$ and there exists $W(x)$ with $V(x, t) \geq W(x)$ for all x, t and $W > 0$. $V(x, t)$ is nonnegative definite in Ξ if $V(0, t) = 0$ and $V(x, t) \geq 0$ for all x, t.

Observe then that in relation to (D1),

$$\dot{V} = (\text{grad } V)' f(x, t) + \frac{\partial V}{\partial t}$$

With this change, Theorems 2.1 and 2.2 both hold. If $V(x, t) \leq W_1(x)$ for all t and some positive definite W_1 uniformity holds in these theorems. In Theorem 2.4, if $W_1(x) \geq V(x, t) \geq W(x)$ with $W(x) \to \infty$ as $\|x\| \to \infty$, global uniform asymptotic stability holds. Theorem 2.5 remains valid without change.

REFERENCES

[1] W. Hahn, *Theory and Applications of Lyapunov's Direct Method*. Englewood Cliffs, N.J.: Prentice Hall, Inc., 1963.

[2] M. Vidyasagar, *Nonlinear Systems Analysis*. Englewood Cliffs, N.J.: Prentice Hall, Inc., 1978.

E

The Riccati Equation

---

1. Relation with linear equations.   Consider the Riccati equation (with time-varying coefficient matrices)

$$-\dot{P} = PF + F'P - PGR^{-1}G'P + Q \qquad P(T, T) = A \tag{E1}$$

and the linear equation

$$\begin{bmatrix} \dot{X} \\ \dot{Y} \end{bmatrix} = \begin{bmatrix} F & -GR^{-1}G' \\ -Q & -F' \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix} \qquad \begin{bmatrix} X(T) \\ Y(T) \end{bmatrix} = \begin{bmatrix} I \\ A \end{bmatrix} \tag{E2}$$

Then the solution of (E1) exists on $[t_0, T]$ if $X(t)$ is nonsingular on $[t_0, T]$, and there holds, see [1–4],

$$P(t, T) = Y(t)X^{-1}(t) \tag{E3}$$

(This is straightforward to verify). Conversely, if the solution of (E1) exists on $[t_0, T]$, and $\theta(t, s)$ denotes the transition matrix of $\dot{x}(t) = [F(t) - G(t)R^{-1}(t)G'(t)P(t)]x(t)$, then $X(t) = \theta(t, T)$, $Y(t) = P(t)\theta(t, T)$ is the solution of (E2), with $X(t)$ nonsingular on $[t_0, T]$, and with (E3) holding. This again is straightforward to verify. If $\Phi(t, s)$ is the $2n \times 2n$ transition matrix associated with (E2) and it is partitioned into four $n \times n$ submatrices, then

$$P(t, T) = [\Phi_{21}(t, T) + \Phi_{22}(t, T)A][\Phi_{11}(t, T) + \Phi_{12}(t, T)A]^{-1} \tag{E4}$$

2. Exponential formula for time-invariant problem.
Suppose that F, G, Q, and R are all constant. One can define the so-called Hamiltonian matrix

$$M = \begin{bmatrix} F & -GR^{-1}G' \\ -Q & -F' \end{bmatrix} \tag{E5}$$

With

$$J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$$

it satisfies $M = JM'J = -JM'J^{-1}$. It has no imaginary eigenvalue, given detectability and stabilizability, and if λ is an eigenvalue, so is $-\lambda$ [5]. Thus there exists a real W such that

$$W^{-1}MW = \begin{bmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{bmatrix} \tag{E6}$$

and Λ_1, Λ_2 are real Jordan matrices such that the real parts of all eigenvalues are respectively negative and positive.

It follows that

$$\begin{bmatrix} X(t) \\ Y(t) \end{bmatrix} = W \begin{bmatrix} \exp \Lambda_1 t & 0 \\ 0 & \exp \Lambda_2 t \end{bmatrix} W^{-1} \begin{bmatrix} I \\ A \end{bmatrix}$$

$$= \begin{bmatrix} W_{11} \exp \Lambda_1 t & W_{12} \exp \Lambda_2 t \\ W_{21} \exp \Lambda_1 t & W_{22} \exp \Lambda_2 t \end{bmatrix} \begin{bmatrix} L \\ RL \end{bmatrix}$$

where L is an unimportant matrix and

$$R = -[W_{22} - AW_{12}]^{-1}[W_{21} - AW_{11}] \tag{E7}$$

Moreover,

$$P(t, T) = [W_{21} + W_{22}e^{\Lambda_2(t-T)}Re^{-\Lambda_1(t-T)}][W_{11} + W_{12}e^{\Lambda_2(t-T)}Re^{-\Lambda_1(t-T)}]^{-1} \tag{E8}$$

3. Evaluating the limiting solution.

It follows from (E8) and because $e^{\Lambda_2 t}$ and $e^{-\Lambda_1 t}$ decay to zero as $t \to \infty$, that

$$\lim_{t \to \infty} P(t, T) = \lim_{T \to \infty} P(t, T) = \bar{P} = W_{21}W_{11}^{-1} \tag{E9}$$

(Existence of the inverse can be established). Notice that the limit is approached at an exponential rate equal to twice the smallest real part of any eigenvalue of Λ_2 and is independent of the boundary condition A. From (E6) and (E9) one has

$$M \begin{bmatrix} I \\ \bar{P} \end{bmatrix} = \begin{bmatrix} W_{11}\Lambda_1 W_{11}^{-1} \\ W_{21}\Lambda_1 W_{11}^{-1} \end{bmatrix}$$

and so

$$F - GR^{-1}G'\bar{P} = -W_{11}\Lambda_1 W_{11}^{-1} \tag{E10}$$

So the eigenvalues of Λ_1 are the closed-loop system modes.

4. Transient solution expressed in terms of limiting solution.

With $\bar{P}$ such that $A - \bar{P}$ is nonsingular, $P(t, T)$ can be expressed in terms of $\bar{P}$ as follows. Let $\bar{F} = F - GR^{-1}G'\bar{P}$ and let $\bar{X}$ satisfy

$$\bar{X}\bar{F}' + \bar{F}\bar{X} = GR^{-1}G' \tag{E11}$$

Then

$$P(t, T) = \bar{P} + \exp[-\bar{F}'(t - T)]\{\exp(-\bar{F}(t - T)]\bar{X}\exp[-\bar{F}'(t - T)]$$
$$+ (A - \bar{P})^{-1} - \bar{X}\}^{-1}\exp[-\bar{F}(t - T)] \tag{E12}$$

5. Steady state Riccati equation solution from Hamiltonian matrix.

Formula (E9) expresses the steady state $\bar{P}$ in terms of the eigenvectors of the Hamiltonian matrix M. It is also possible (and numerically preferable) to use a Schur form:

$$M = U'\begin{bmatrix} L_{11} & L_{12} \\ 0 & L_{22} \end{bmatrix} U \tag{E13}$$

where L is a Schur form of M with L_{11} possessing all negative real part eigenvalues, and U is orthogonal. Then

$$\bar{P} = U_{12}'(U_{11}^{-1})' \tag{E14}$$

In case (E13) is any Schur form, without the restriction that the eigenvalues of L_{11} possess negative real parts, the formula $X = U_{12}'(U_{11}^{-1})'$ yields a solution of the quadratic matrix equation $XF + F'X - XGR^{-1}G'X + Q = 0$. To see this note that (E13) implies

$$M\begin{bmatrix} U_{11}' \\ U_{12}' \end{bmatrix} = U'\begin{bmatrix} L_{11} \\ 0 \end{bmatrix}$$

or

$$FU_{11}' - GR^{-1}G'U_{12}' = U_{11}'L_{11} \qquad -QU_{11}' - F'U_{12}' = U_{12}'L_{11}$$

or

$$XF - XGR^{-1}G'X = U_{12}'(U_{11}')^{-1}U_{11}'L_{11}(U_{11}')^{-1}$$
$$= U_{12}'L_{11}(U_{11}')^{-1} = -Q - F'X$$

The eigenvalues of the "closed-loop" system matrix $F - GR^{-1}G'X$ are those of L_{11}, since $F - GR^{-1}G'X = U_{11}'L_{11}(U_{11}')^{-1}$.

There is an excellent discussion of the issues involved in this method for solving the steady state equation in [6]. Reliable software exists for state dimension up to 100, and the requirement that L_{11} possess negative real part eigenvalues, not standard in Schur algorithms, can be achieved through a stable algorithm based on orthogonal transformations with starting point an arbitrary Schur decomposition.

6. Recursive determination of steady state solution.

The steady state $\bar{P}$ can also be found by a limiting process, [7]. Let K_0 be such that $F + GK_0'$ has all eigenvalues with negative real parts. Define P_i, K_i recursively by

$$P_i(F + GK_i') + (F + GK_i')'P_i = -K_iRK_i' - Q \tag{E15}$$

$$K_{i+1} = -P_iGR^{-1} \tag{E16}$$

Then $P_i \geq P_{i+1}$ and $\lim_{i \to \infty} P_i = \bar{P}$. Further, $\|P_{i+1} - \bar{P}\| \leq c \|P_i - \bar{P}\|^2$ that is, convergence is quadratic. Reference [6] advocates the use of the Schur algorithm to determine a "first estimate" P_0 of $\bar{P}$ and then the use of the above algorithm initialized with $K_0 = -P_0 GR^{-1}$.

7. Steady state Riccati equation solutions via discrete time transformations.

The steady state $\bar{P}$ can also be found by setting up a discrete-time linear-quadratic problem for which the limiting solution of the discrete-time Riccati equation is also $\bar{P}$. Given F, G, $Q = DD'$ and R, define

$$A = (I + F)(I - F)^{-1}$$

$$B = \frac{1}{\sqrt{2}}(A + I)G = \sqrt{2}(I - F)^{-1}G$$

$$C = R + \tfrac{1}{2}B'QB$$

$$D = \frac{1}{2\sqrt{2}}(I + A')Q(I + A)G$$

$$= \sqrt{2}(I - F')^{-1}Q(I - F)^{-1}G$$

$$E = 2(I - F')^{-1}Q(I - F)^{-1}$$

$$= \tfrac{1}{2}(I + A')Q(I + A)$$

Then with

$$\Phi_{i+1} = A'\Phi_i A - [A'\Phi_i B + D][C + B'\Phi_i B]^{-1}[A'\Phi_i B + D]' + E \qquad \text{(E17)}$$

there holds $\lim_{i \to \infty} \Phi_i = \bar{P}$; see [8], where the calculation is done in a trivially different coordinate basis.

8. Steady state Riccati solution from spectral factorization.

Suppose that the characteristic polynomial of M is factorized as $p(s)p(-s)$, where $p(s)$ has all stable eigenvalues. Then $\bar{P}$ is uniquely defined by

$$p(M)\begin{bmatrix} I \\ \bar{P} \end{bmatrix} = 0 \qquad \text{(E18)}$$

This observation, from [9], follows easily from (E6):

$$W^{-1}p(M)W = \begin{bmatrix} p(\Lambda_1) & 0 \\ 0 & p(\Lambda_2) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & p(\Lambda_2) \end{bmatrix}$$

whence

$$p(M)\begin{bmatrix} W_{11} \\ W_{21} \end{bmatrix} = 0$$

9. Other methods.

Several other methods for solving the steady state Riccati equation exist, especially in discrete time, where their development has been driven by Kalman filtering problems. These include doubling algorithms,

Chandrasekhar algorithms, square root algorithms, "information filter" algorithms, and matrix sign function algorithms. For an introduction to these ideas, see [10] and [11].

REFERENCES

[1] R. E. Kalman, "Contributions to the Theory of Optimal Control," *Bol. Soc. Matem, Max.* (1960), pp. 102–119.

[2] R. E. Kalman and T. S. Englar, "A User's Manual for ASP C," *RIAS Rep. NASA*, 1965, Baltimore, Md.

[3] W. T. Reid, "A Matrix Differential Equation of the Riccati Type," *Am. J. Math.*, Vol. 68 (1946), pp. 237–246.

[4] J. J. Levin, "On the Matrix Riccati Equation," *Proc. Am. Math. Soc.*, Vol. 10 (1959), pp. 519–524.

[5] F. L. Lewis, *Optimal Control*. New York: John Wiley and Sons, 1986.

[6] W. F. Arnold, III and A. J. Laub, "Generalized Eigenproblem Algorithms and Software for Algebraic Riccati Equations," *Proc. IEEE*, Vol. 72, No. 12 (December 1984), pp. 1746–1754.

[7] D. L. Kleinman, "On an Iterative Technique for Riccati Equation Computation," *IEEE Trans. Auto. Control*, Vol. AC-13, No. 1 (February 1968), pp. 114–115.

[8] K. L. Hitz and B. D. O. Anderson, "Iterative Method of Computing the Limiting Solution of the Matrix Riccati Differential Equation," *Proc. IEEE*, Vol. 119, No. 9 (1972), pp. 1402–1406.

[9] R. S. Bucy and P. D. Joseph, *Filtering for Stochastic Processes with Applications to Guidance*. New York: Interscience Publishing, 1968.

[10] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Englewood Cliffs, N.J.: Prentice Hall, Inc., 1979.

[11] B. D. O. Anderson, "Second-order Convergent Algorithms for the Steady-State Riccati Equation," *Int. J. Control*, Vol. 28, No. 2 (1978), pp. 295–306.

Author Index

Subject Index