METHODS FOR SUBOPTIMAL DESIGN OF NONLINEAR CONTROL SYSTEMS

HIDEKIYO ITAKURA

DEPARTMENT OF ELECTRICAL ENGINEERING
KYOTO UNIVERSITY

OCTOBER, 1974

METHODS FOR SUBOPTIMAL DESIGN OF NONLINEAR CONTROL SYSTEMS

HIDEKIYO ITAKURA

DEPARTMENT OF ELECTRICAL ENGINEERING

KYOTO UNIVERSITY

KYOTO, JAPAN

OCTOBER, 1974

PREFACE

This paper is devoted to the study of some classes of optimization problems in nonlinear control systems. The systems under consideration are described by nonlinear differential equations containing small parameters. Nonlinear functions in the equations are assumed analytic. Suboptimal feedback controls for the systems are designed so as to optimize a performance index of integral type. The approximation method of approaches presented in the paper is based on a small parameter method, or a perturbation method. The method proposed is applied to some types of linear systems.

The text consists of five chapters. The first chapter is introductory one in which a perturbation method is outlined and the contents of the text are summarized. Chapter 2 describes a method for suboptimal design of nonlinear feedback systems and fully discusses various features in the design. The chapter that follows extends the method to various control problems containing small parameters, mainly to a large-scale problem and a problem of periodic systems. Chapter 4 is concerned with a suboptimal feedback control for nonlinear systems under noise disturbances. The last chapter deals with nonlinear differential game problems and presents approximation approaches to various types of solution of the games. Various typical examples attached in each chapter illustrate several calculational features of the proposed method. Five appendices concerning supplementary remarks are annexed to the text.

ACKNOWLEDGMENTS

The author owes a lasting debt of his sincere gratitude to Dr. Chihiro Hayashi, Professor of Kyoto University, who has given him constant and generous guidance and encouragement in promoting this work. In the preparation of the present paper, the author has been given many aids and useful advices by Dr. Yoshikazu Nishikawa, Professor of Kyoto University. Especially, the author is indebted to Professor Nishikawa for suggesting the field of research of the present thesis and stimulating interest in the research. Acknowledgment must also be made to Dr. Nobuo Sannomiya, Lecturer of Kyoto University, for his frequent, stimulating, and invaluable discussions of all kinds. The author is likewise grateful to the members and the graduate students of Professor Hayashi's and Professor Nishikawa's Laboratories for their excellent cooperation.

The Data Processing Center of Kyoto University has made computer time available to the author. The author would like to acknowledge the kind considerations of the staffs of the Data Processing Center. Finally, the author is glad to thank the Sakkokai Foundation for a fund of research-aid in promoting this investigation.

GLOSSARY OF SYMBOLS

Throughout this text except for in the appendices and the sections of examples, unless otherwise specified, the following symbol conventions are generally adhered to the rules: Scalars are denoted by lower case Greek letters, vectors by lower case Roman, and matrices by upper case Roman or Greek. Contrary to these rules, some exceptions are permitted due to notational usage: times t and T, performance index J, Hamiltonian H, dimensions of vectors n, m, etc., index integers i, j, k, ℓ, etc., number N, period L, mean value vector μ, covariance matrix σ, and so on. With no exception, vectors are in column form.

Following is a list of principal symbols appearing in this text.

$A(t)$: system matrix

$\tilde{A}(x, t)$: state-dependent system matrix in instantaneous linearization

$A_e(t)$: system matrix for an evader

$A_i(t)$: system matrix of the i-th subsystem

$A_p(t)$: system matrix for a pursuer

$A^{(0)}$: constant system matrix

$A^{(1)}(t)$: periodically-varying system matrix

$B^{(t)}$: control matrix of a system

$B_e(t)$: strategic control matrix for an evader

$B_i(t)$: control matrix of the i-th subsystem; strategic control matrix for Player i

$B_p(t)$: strategic control matrix for a pursuer

$B^{(0)}$: constant control matrix

$B^{(1)}(t)$: periodically-varying control matrix

cov: covariance operator for a noise process

$C(t)$: output matrix of a system; disturbance matrix of a noisy system

det: determinant operator for a matrix

$D(t)$: degradation matrix in a suboptimal performance index

$D(x, t)$: state-dependent control matrix in a system with controls appearing linearly

$D_{ij}(t)$: coupling matrix in a linear large-scale system

$\partial\phi/\partial x$: gradient vector whose i-th component is $\partial\phi/\partial x_i$ for a scalar ϕ and a vector x

$\partial v/\partial x$: matrix whose (i, j)-element is $\partial v_i/\partial x_j$ for vectors v and x

$\partial^2\phi/\partial x^2$: Hessian matrix whose (i, j)-element is $\partial^2\phi/\partial x_i \partial x_j$ for a scalar ϕ and a vector x

$\partial v(a)/\partial x$: value of $\partial v/\partial x$ at $x = a$

$\mathcal{E}$: expectation operator for a noise process

$f(x, t)$: nonlinear function in a system

$f_e(x, t)$: nonlinear function in an evader's system

$f_i(x, t)$: nonlinear coupling function in a large-scale system

$f_p(x, t)$: nonlinear function in a pursuer's system

F: weighting matrix in J related to final states

$\tilde{F}$: weighting matrix in J related to the difference between evader's and pursuer's final states

F_i: weighting matrix in J_i related to the i-th final substates; weighting matrix in Player i's J_i related to final states

H: Hamiltonian of a problem

H_i: Hamiltonian for Player i

i: index integer

$I^{(k)}$: $k \times k$-identity matrix

j: index integer

J: performance index

J^*: minimum value of J

J_i: performance index of the i-th subproblem or of Player i

J_s: value of J resulting from a suboptimal control

k: index integer

ℓ: index integer

L: period of time; number of points with a boundary condition

m: dimension of a control vector

m_e: dimension of an evader's strategic control vector

m_i: dimension of a control vector of the i-th subsystem; dimension of Player i's strategic control vector

m_p: dimension of a pursuer's strategic control vector

$\tilde{M}$: repeated Kronecker sum of a matrix M or M'

$M^{(i)}$: i-times repeated Kronecker sum of a matrix M or M'

M_{ij}: (i, j)-element of a matrix M; (i, j)-submatrix in a composite matrix M; notation of a matrix M related to Players i and j

n: dimension of a system

n_i: dimension of the i-th subsystem

N: number of subsystems, frequency components, or players

p(x, t): costate vector

$p_i(x, t)$: costate vector of Player i

P(t): optimal feedback gain matrix for a linear-quadratic problem

$\tilde{P}(x, t)$: state-dependent feedback gain matrix in instantaneous linearization

$P_i(t)$: Player i's optimal feedback gain matrix for a linear-quadratic game problem

$P_s(t)$: suboptimal feedback gain matrix

P_0: optimal constant feedback gain matrix for a time-invariant linear-quadratic problem

q: degree of a polynomial

Q(t): weighting matrix in J related to states

$Q_i(t)$: weighting matrix in J_i related to the i-th substates; weighting matrix in Player i's J_i related to states

r: index integer

R(t): weighting matrix in J related to controls

$R_e(t)$: weighting matrix in J related to evader's strategic controls

$R_i(t)$: weighting matrix in J_i related to the i-th subcontrols; weighting matrix in J_i related to Player i's strategic controls

$R_{ij}(t)$: weighting matrix in Player i's J_i related to Player j's strategic controls

$R_p(t)$: weighting matrix in J related to pursuer's strategic controls

s: dimension of an output vector; dimension of a noise vector

S(t): resultant system matrix for a linear-quadratic problem with the optimal control

$S_s(t)$: resultant system matrix with a suboptimal control

S_0: resultant constant system matrix for a time-invariant linear-quadratic problem with the optimal control

t: time

tr: trace operator for a matrix

T: final time

u: control vector of a system

u*: optimal control

u_e: strategic control vector of an evader

u_i: control vector of the i-th subsystem; strategic control vector of Player i

u_i^*: minimax strategy of Player i

$\bar{u}_i$: Nash equilibrium strategy of Player i

$\hat{u}_i$: Pareto optimal strategy of Player i

u_p: strategic control vector of a pursuer

u_s: suboptimal control

v*: optimal or exact value of a quantity v

v_i: i-th component of a vector v; notation of a quantity v for the i-th subproblem or for Player i

$v^{(i)}$: i-th-order coefficient of a quantity v developed into a power series

v_s: suboptimal or approximate value of a quantity v

v_0: time-independent form of a function v(t) or v(x, t) as $t \to \infty$ or $t \to -\infty$

v^0: initial value of a quantity v(t) at $t = \tau$

V(t): suboptimal performance matrix

w: noise vector

x: state vector of a system

x_e: state vector of an evader

x_i: state vector of the i-th subsystem

x_p: state vector of a pursuer

y: output vector of a system

z: Wiener process vector

δ: infinitesimal operator for a quantity

Δ: domain of a space

ε: small scalar parameter

μ: mean value of a noise process

σ: covariance matrix of a noise process

τ: initial time

$\Phi(t)$: fundamental matrix of a linear system

$\Psi(t)$: fundamental matrix of a linear system

ω: angular frequency

0: zero scalar; zero vector; zero matrix

$\dot{}$: differentiation with respect to time

': transposition of a vector or a matrix

>: positive definiteness of a matrix

$\geq$: positive semidefiniteness of a matrix

[,]: interval of time

$\triangleq$: definition of a quantity

$\| \ \|$: norm of a vector or a matrix

$\otimes$: Kronecker product of matrices

$\oplus$: Kronecker sum of matrices

CONTENTS

CHAPTER 1

INTRODUCTION

1.1 Perturbation Method

There have been numerous studies on the optimization problems for linear dynamical systems: linear regulator problem [37, 40],* linear stochastic control problem [24], linear differential game problem [33, 90], control problem of linear distributed parameter systems [29], linear filtering problem [38], state observer problem for linear systems [54, 55], and so on. Each of these linear problems has been solved in an elegant closed form. However, many problems faced today by engineers may involve nonlinearities in system dynamics. Relatively few works have been reported on nonlinear optimization problems. This is because in general the determination of the exact solutions is difficult or impossible for nonlinear problems, except for special classes of problems [16, 85, 89, 93].

There are some major trends in modern developments of nonlinear analysis: the purely numerical methods using a digital computer [23, 79], the methods of linearization with application of linear theories [11, 21, 27, 46, 78, 85, 86, 89], the methods of parameter optimization with solutions of fixed configuration [18, 20, 52, 53, 84-86], the analytical methods of successive approximations, and combinations of some of these. The analytical methods based on successive approximations, proposed in various forms by many investigators [1, 5, 25-27, 30, 39, 56, 81, 96], are most systematic. Most of the methods have given only formal procedures of successive approximations, but have kept away

---

* Numbers in brackets indicate references at the end of the text.

from related natures in successive calculations: uniqueness, solvability, stability, continuity, and so on.

Foremost among the successive approximation techniques is the small parameter method, or the method of perturbation in terms of a small parameter. The use of the perturbation method was, in the earlier period, limited to astrodynamical calculations. But the important contributions of H. Poincaré and later mathematicians or engineers have broadened applicability of the method to include a more general field of nonlinear problems [8, 13, 62], especially a field of nonlinear oscillations [9, 31]. An initial-value problem of differential equations is of primary interest in case of nonlinear oscillation problems; while a two-point boundary-value problem has to be solved in most case of optimization problems, to which the method may also be applicable.

This text presents approximate approaches to nonlinear dynamical optimization problems with use of the perturbation method, and furthermore fully discusses various features in successive calculations. All the problems treated in the text are reduced to the task of solving equations of the form*

$$g(x, v, \dot{x}, \dot{v}, t; \varepsilon) = 0 \tag{1.1}$$

where x, v, and g are n-, n_v-, and $(n + n_v)$-vectors, respectively, and t is time; ε is a scalar parameter. A dot over a quantity denotes differentiation with respect to t. The parameter ε, ideally a nondimensional number, may naturally appear in Eq. (1. 1), or may artificially be introduced for convenience. The magnitude of ε should

---

* Here and throughout the text, the symbol 0 is used, without notice, as zero scalar, zero vector, and zero matrix, but should be understood.

be small in some sense. The boundary conditions for Eq. (1.1) are given by

$$x(\tau) = x^0, \qquad v(T) = v^T \tag{1.2}$$

for we restrict our attention to the optimization problems of so-called regulator type. Since we wish to obtain controls of feedback type, the solution to Eq. (1.1) satisfying (1.2) may be found in the form as

$$v = h(x, t; \varepsilon) \tag{1.3}$$

In application of the perturbation method, we develop the desired quantity $h(x, t; \varepsilon)$ in powers of the small parameter ε multiplied by coefficients which are functions of x and t:

$$h(x, t; \varepsilon) = \sum_{k=0}^{\infty} \varepsilon^k v^{(k)}(x, t) \tag{1.4}$$

The solution (1.4) is usually so determined as, if $\varepsilon \to 0$, the generating solution

$$h(x, t; 0) = v^{(0)}(x, t) \tag{1.5}$$

is the exact solution to the unperturbed equation

$$g(x, v^{(0)}, \dot{x}, \dot{v}^{(0)}, t; 0) = 0 \tag{1.6}$$

We assume the solution of Eq. (1.6) possible; and we then proceed to determine the coefficients $v^{(k)}$ of the development (1.4) one by one, by solving a sequence of equations

$$g^{(k)}(x, v^{(k)}, t) = 0 \qquad (k = 1, 2, \ldots) \qquad (1.7)$$

This constitutes the principal subject of the text.

The above description, however, may be somewhat perfunctory. If the parameter ε is sufficiently small, the series of Eq. (1.4) converges fast enough so that merely the first two or three terms give good accuracy. When ε is large, the method ceases to be applicable, and for such cases analytical methods are largely unexplored. However, it is worth noticing that, from a practical point of view, the method may still be applicable with satisfactory results even for not very small values of the parameter.

1.2 Description of Contents

The text consists of five chapters, including the present introductory chapter. In the subsequent two chapters, methods for suboptimal design of nonlinear feedback systems are developed, with extensions to various control problems containing small parameters. The chapter that follows is concerned with a suboptimal feedback control for nonlinear systems with noise disturbances. The last chapter deals with nonlinear differential game problems and describes approximation approaches to various types of solution of the games. Complementary remarks are provided in five appendices to the text.

Chapter 2 presents a systematic procedure to construst a suboptimal feedback control for nonlinear systems in which a parameter ε is associated with nonlinearities. The nonlinearities are assumed analytic in states. For simplicity a performance index is considered to be quadratic in states and controls. For a sufficiently small value of ε, a desired feedback function may be expanded and sought in a power series in ε. The generating solution is the unperturbed solution, which is usually obtained by solving a Riccati type equation. Correction terms for improving the feedback function are yielded by successive solutions of a sequence of linear partial differential equations. The method of separation of variables permits to reduce the partial differential equation to a linear ordinary differential equation. By using the notion of the Kronecker sum of matrices, the ordinary differential equation is written in a simple conventional form, so that it can readily be solved either analytically or numerically by a computer routine. When the system is time-invariant and a control duration is infinite, and if the unperturbed system is completely controllable and completely observable, the differential equation is further reduced to an algebraic equation, whose solution is again straightforward.

Chapter 3 is concerned with extensions of the method of Chapter 2 to a wide variety of control problems containing small parameters. The problem of nonlinear systems with controls appearing linearly is a more generalized one of the problem of Chapter 2. The system considered is linear in controls, but contains product terms of states and controls.

Another straightforward application of the method is the synthesis of a suboptimal control for a large-scale system. If a system is composed of weakly-coupled subsystems, the parameter ε could also be associated with the couplings of the subsystems. The idea of suboptimal design is discussed for N-coupled systems.

A particular problem of a large-scale system is such that the overall system is linear. In this case, the exact calculation of the optimal control is theoretically possible; however, the high dimensionality of the system makes the conventional method impractical. It is then better to split a large-scale design problem into a set of simple subsystem problems. The approach which completely neglects the couplings may result in unsatisfactory system performance. Chapter 3 contains a discussion on a compromise between calculational simplicity of decoupled design and performance degradation.

In Chapter 3 an approximate construction of the optimal state regulator is also proposed for a periodically-varying linear system. The construction of an optimal state regulator for a linear time-varying system involves the solution of a matrix Riccati equation. The closed form solution of the Riccati equation is impossible even when a control duration is sufficiently large. Hence, the aim is to determine a linear feedback control which is relatively easy to implement, but results in suboptimal performance index. The system considered contains a periodic term which is taken to be perturbation to the time-invariant linear-system equation.

After the deterministic optimal control has been studied in Chapters 2 and 3, Chapter 4 treats the problem of stochastic optimal control. The synthesis of optimal feedback control is considered in the ensemble average sense assuming the presence of Gaussian noise. The problem is to design the optimal control for a nonlinear system disturbed by Gaussian white noise. A perfect observation of the state of the system is assumed. A system with a feedback controller well responds to the disturbances, even if the system is designed with neglect of noises. However, provided that statistical characteristics of noises are known in advance, a better control is expected to obtain, by taking the characteristics into account. By the reasoning similar to the deterministic case, it is difficult to solve analytically the problem of nonlinear noisy systems. Also in Chapter 4 the small parameter method is applied to the problem.

Chapter 5 deals with nonlinear differential game problems of nonzero-sum as well as of zero-sum. Preliminarily, important differences are illustrated between zero-sum and nonzero-sum games, referring to finite games formulated as rectangular games. A zero-sum two-person differential game is introduced whose state dynamics is governed by a slightly nonlinear differential equation. For simplicity, a cost function is considered to be quadratic in states and strategic controls. A minimax solution of the game is obtained in power-series form in ε. The zero-order term gives a well known Riccati equation, and higher-order terms yield a series of linear differential equations. This procedure is an application of the method of Chapter 2. In succession a nonzero-sum N-person differential game with slightly nonlinear dynamics is treated. Each player has his own cost of the quadratic type. The method similar to that for the zero-sum case is applied to approximate determination of the Nash equilibrium solution, the minimax solution, and the Pareto optimal solution.

Various typical examples attached in each of Chapter 2 through 5 illustrate several features of the proposed methods. Some of these examples are of physical nature, while others are artificial ones. The values of the performance index are calculated for suboptimal controls of the various orders, analytically and/or numerically using a digital computer.

As has been mentioned earlier, five appendices are annexed to the text. Appendix I is concerned with the definitions and related properties of the Kronecker product and sum of matrices, which are used in Chapter 2 to rewrite complicated differential equations into a simple conventional form. Appendix II derives dynamical equations for some control problems of space vehicles, which are used in the examples of the text. The approximate equations of the various orders are derived for two kinds of problems: the problem of attitude control of an orbiting satellite and the problem of orbit transfer between neighboring elliptic orbits. Appendices III and IV summarize the procedures for suboptimal solutions of the problems with a nonlinear output equation and with a nonquadratic performance index, respectively. Appendix V examines the transient variation of Nash equilibrium solution for a nonzero-sum differential game, associated with a example treated in Chapter 5.

CHAPTER 2

SUBOPTIMAL CONTROL OF NONLINEAR SYSTEMS

2.1 Introduction

There have been numerous studies on the optimal feedback control of linear dynamical systems or on the optimal design of linear regulators. On the other hand, relatively few works have been reported on nonlinear regulator problems, because of the difficulty in determining the exact optimal feedback function.

E. G. Al'brekht [1] has shown a sufficient condition for obtaining the optimal feedback control of a nonlinear analytic system, and developed formally a recursive procedure to construct a suboptimal control in a power series with respect to states. D. L. Lukes [56] has extended the work of Al'brekht and relaxed the analyticity condition to twice continuous differentiability. However, a closed form solution for the recursive procedure has not been demonstrated in their papers. J. D. Pearson [78] and W. L. Garrard, N. H. McClamroch, and L. G. Clark [27] have developed several approximation methods. Their methods are essentially based on the idea utilizing a technique similar to that for the linear systems.

In this chapter, a systematic procedure is proposed to construct a suboptimal state regulator for a class of nonlinear systems in which a parameter ε is associated with nonlinearities [66-68, 71, 76]. The ε-parameter, called the perturbation parameter, is either inherent in the system or introduced for convenience. This procedure extends the method developed by J. F. Baldwin and J. H. S. Williams [5]. They have tried to expand the solution of the Hamilton-Jacobi equation in a power series with respect to the parameter ε, but have not deeply discussed various features of successive calculations.

The problem considered is stated in Section 2.2. The nonlinearities are assumed analytic in the state. For simplicity a performance index is considered to be quadratic in the state and control. Section 2.3 discusses a necessary condition for the optimal solution, and Section 2.4 summarizes the solution to the linear regulator problem.

Section 2.5 outlines the method of instantaneous linearization proposed by J. D. Pearson [78]. This method is one of the representative approximation approaches to nonlinear problems. By this idea the calculation seems to be much simplified; however, the best possible solution can not necessarily be obtained, except for a one-dimensional problem [89], because of some arbitrariness in the system description. In short, the method is not necessarily effective to treat the system of high dimension.

Section 2.6 clarifies that, for a properly small value of ε, there exists a unique optimal solution neighboring that of the unperturbed linear problem. For the proof, an existence theorem for multi-point boundary-value problems is used. For a sufficiently small value of ε, a desired feedback function may be expanded and sought in a power series in ε. The detail of power-series expansion is discussed and the equations to determine the power series is derived in Section 2.7. The generating solution is the unperturbed solution, which is usually obtained by solving a Riccati type equation. Correction terms for improving the feedback function are yielded by successive solutions of a sequence of linear partial differential equations. In Section 2.8 the method of separation of variables permits to reduce the partial differential equation to a linear ordinary differential equation. Section 2.9 proves a lemma regarding an extended Liapunov equation. In Section 2.10, the lemma, together with the notion of the Kronecker sum of matrices, enables us to write the ordinary differential equation derived in Section 2.8 in a simple conventional form. Then the equation can readily be solved either analytically or numerically by a computer routine. Section 2.11 comments on the analogy between the stability condition of the feedback functions and the local controllability of nonlinear systems.

In practical calculation we have inevitably to truncate the power series, yielding a suboptimal control. The quality of the suboptimal control is examined in Section 2.12.

Several features of the method are illustrated by five typical examples in Section 2.13.

2.2 Problem Statement

We consider dynamical systems whose state is governed by the nonlinear differential equation

$$\dot{x} = A(t)x + \varepsilon f(x, t) + B(t)u \tag{2.1}$$

defined on the interval $[\tau, T]$ with

$$x(\tau) = x^0 \tag{2.2}$$

The output of the system is given by

$$y = C(t)x \tag{2.3}$$

In the above equations,

- x: n-state vector
- y: s-output vector
- u: m-control vector

f(x, t): n-vector function analytic in x and continuous in time t satisfying $f(0, t) = 0$ for any t

$A(t)$: $n \times n$-matrix continuous in t

$B(t)$: $n \times m$-matrix continuous in t

$C(t)$: $s \times n$-matrix continuous in t

ε: small scalar parameter

$\dot{}$: differentiation with respect to t

The nonlinearity is taken to be a perturbation to the system. The parameter ε, the perturbation parameter, is either inherent in the system or introduced for convenience.

Our problem is to find a feedback control function $u(x, t)$ for which the performance index

$$J = \frac{1}{2}y'(T)Fy(T) + \frac{1}{2}\int_{\tau}^{T}[y'Q(t)y + u'R(t)u]dt \tag{2.4}$$

is minimized. In Eq. (2.4)

F: $s \times s$-positive semidefinite symmetric matrix

$Q(t)$: $s \times s$-positive semidefinite symmetric matrix continuous in t

$R(t)$: $m \times m$-positive definite symmetric matrix continuous in t

$'$: transposition of a vector or a matrix

We assume that the control u is not constrained in magnitude and the final time T is specified.

Let us now see in what way the performance index given by Eq. (2.4) represents a reasonable mathematical translation of the physical requirements and specifications. First, consider the term $y'Q(t)y/2$. Clearly, since $Q \geqq 0$,* it is nonnegative for all y, so that it weights

---

* For a symmetric matrix M, the expressions $M > 0$ and $M \geqq 0$ mean that M is positive definite and M positive semidefinite, respectively.

large outputs much more heavily than small outputs. Second, we consider the term $u'R(t)u/2$. This term weights the cost of the control and penalizes the system much more severely for large controls than for small controls. The requirement $R > 0$ instead of $R \geqq 0$ is, as we shall see in Section 2.3, a condition for the existence of a finite control. The integral of the term quadratic in control is often referred to as the control energy. Hence our problem may be called a minimum-energy problem.

Finally, we turn our attention to the term $y'(T)Fy(T)/2$. It is referred to as the terminal cost, and its purpose is to guarantee the small output at the final time T. Accordingly this term should be included if the value of $y(T)$ is particularly important; if this is not the case, then we can set $F = 0$ and rely upon the rest term of J to guarantee that $y(T)$ is not excessively large.

We wish to emphasize that whether or not the performance index of the form (2.4) should be used for a given problem is up to the designer. Nonetheless, it turns out that this performance index has two very desirable properties: First, it is mathematically tractable, and second, as we shall see in Section 2.4, it yields the exact optimal feedback control for linear systems.

The physical motivation for the use of the performance index (2.4) has been described above. A self-evident but useful fact is left: If $Q = F = 0$, then the optimal control must be $u(t) = 0$. To exclude this trivial case, we shall assume throughout the text that Q and F are not both the zero matrices, although we shall allow either $Q = 0$ or $F = 0$.

2.3 Necessary Condition for the Optimality

Our aim in this section is to develop a necessary condition for the optimal solution of the problem of Section 2.2.

Let us assume that an optimal control exists for any initial state. We can use the minimum principle of L. S. Pontryagin [80] to obtain the necessary condition for the optimal control.* The Hamiltonian H of the system (2.1) with the output of Eq. (2.3) and the cost of Eq. (2.4) is given by

$$H = \frac{1}{2}x'C'QCx + \frac{1}{2}u'Ru + p'(Ax + \varepsilon f + Bu) \tag{2.5}$$

The n-costate vector $p(x, t)$ is the solution to the vector differential equation

$$\dot{p} = -\frac{\partial H}{\partial x} = -C'QCx - A'p - \varepsilon\left(\frac{\partial f}{\partial x}\right)'p \tag{2.6}$$

satisfying the boundary condition

$$p(x, T) = C'(T)FC(T)x(T) \tag{2.7}$$

The $n \times n$-matrix $\partial f/\partial x$ is defined in such a way that the (i, j)-element of $\partial f/\partial x$ is $\partial f_i/\partial x_j$.**

Due to the minimum principle, a necessary condition for the optimality is that the Hamiltonian be minimum with respect to u [80]. Hence, along the optimal trajectory, we must have

$$\frac{\partial H}{\partial u} = Ru^* + B'p = 0 \tag{2.8}$$

---

* Owing to no magnitude constraint on the control variables, we may also use the dynamic programming of R. Bellman [6] or the classical variational approach.

** The i-th component of a vector v is denoted by v_i.

or equivalently

$$u^*(x, t) = -R^{-1}(t)B'(t)p(x, t) \tag{2.9}$$

The assumption that $R(t) > 0$ guarantees the existence of $R^{-1}(t)$ for all $t \in [\tau, T]$.

The condition $\partial H/\partial u = 0$ yields only an extremum of H with respect to u. If the $m \times m$-matrix $\partial^2 H/\partial u^2$ is positive definite,* then the extremum of H is a minimum with respect to u. Since we have, from Eq. (2.8),

$$\frac{\partial^2 H}{\partial u^2} = R \tag{2.10}$$

the control u^* given by Eq. (2.9) does indeed minimize H and is optimal, if $p(x, t)$ is obtained.

The system resulting from use of the optimal control $u = u^*$ is

$$\dot{x} = A(t)x + \varepsilon f(x, t) - E(t)p \tag{2.11}$$

where E is the $n \times n$-symmetric matrix defined by**

$$E(t) \triangleq B(t)R^{-1}(t)B'(t) \tag{2.12}$$

The problem is now reduced to that of solving the two-point boundary-value problem: Eqs. (2.11), (2.2), (2.6), and (2.7).

---

* For a scalar function $\phi(x)$ analytic in x, $\partial^2\phi/\partial x^2$ is the symmetric matrix, usually called the Hessian matrix, whose (i, j)-element is $\partial^2\phi/\partial x_i \partial x_j$.

** The symbol $\triangleq$ denotes definition of a quantity.

2.4 Optimal Control for Linear Systems

In this section we briefly discuss the solution to the linear regulator problem.

It is well known that, for the unperturbed linear system, i.e., the system (2.1) with $\varepsilon = 0$, the solution $p(x, t)$ of Eq. (2.6) takes the form [37]

$$p(x, t) = P(t)x \tag{2.13}$$

where P, the $n \times n$-symmetric matrix, is the solution of the matrix Riccati equation

$$\dot{P} = - PA - A'P + PEP - C'QC \tag{2.14}$$

satisfying the boundary condition

$$P(T) = C'(T)FC(T) \tag{2.15}$$

It is noted that $P(t)$ is positive definite for all t $(< T)$. By combining Eq. (2.13) with (2.9) the optimal control is given by the linear form in state:*

$$u^*(x, t) = - R^{-1}(t)B'(t)P(t)x \tag{2.16}$$

The optimal trajectory is obtained by solving the linear homogeneous system

---

* Note that the optimal control is a function of the state x rather than of the output y. This fact may appear disappointing to us, since we might have expected the optimal control to be a function of the output. However, if the system is completely observable, we can compute the state from knowledge of the output. This is the problem of constructing an observer [54, 55].

$$\dot{x} = S(t)x, \qquad x(\tau) = x^0 \tag{2.17}$$

where S is the $n \times n$-matrix:

$$S(t) \triangleq A(t) - E(t)P(t) \tag{2.18}$$

The minimum value of J, J^*, is

$$J^* = \frac{1}{2}x'(\tau)P(\tau)x(\tau) \tag{2.19}$$

Now we consider the time-invariant systems with $T = \infty$. Let the matrices $A(t)$, $B(t)$, $C(t)$, $R(t)$, and $Q(t)$ be all constant matrices and let $F = 0$. Under the additional conditions that the system (2.1) and (2.3) with $\varepsilon = 0$ is completely controllable and completely observable* and that Q is positive definite, the optimal control is free of explicit dependence on t [3, p. 771]:

$$u^*(x) = -R^{-1}B'P_0x \tag{2.20}$$

where P_0 is the $n \times n$-constant positive definite matrix satisfying the quadratic matrix algebraic equation

$$-P_0A - A'P_0 + P_0EP_0 - C'QC = 0 \tag{2.21}$$

Due to the time-invariance, the initial time τ is let zero with no loss of generality. Thus the optimally controlled system is also time-invariant:

---

* For the definitions of controllability and observability, refer to any standard textbooks on control theories [e.g., 3, p. 200].

$$\dot{x} = S_0 x, \qquad x(0) = x^0 \tag{2.22}$$

where

$$S_0 \triangleq A - EP_0 \tag{2.23}$$

The minimum cost J^* is given by

$$J^* = \frac{1}{2} x'(0) P_0 x(0) \tag{2.24}$$

Moreover, the matrix S_0 is stable,* and consequently the optimal system (2.22) is asymptotically stable.

A terminal cost with $T = \infty$ does not make much engineering sense, since an engineer is always interested in the response of a system in a finite time. Nevertheless, we let $T \rightarrow \infty$ for the following reasons: First, we wish to guarantee that the output stays near zero after an initial transient interval and, second, we wish to avoid the specification of a large final time T.

2.5 Approximation Approaches Based on Linearization

The exact determination of the optimal control for nonlinear systems is generally impossible, because of the difficulty in solving the two-point boundary-value problem of Eqs. (2.11), (2.2), (2.6), and (2.7). One of approximation approaches to the nonlinear regulator problem is the method of instantaneous linearization first proposed by J. D. Pearson [78]. Several investigations on this method have been reported [11, 27,

---

* A constant matrix M is said to be stable as $t \rightarrow \infty$, or simply stable, if all the eigenvalues of M have negative real parts.

85, 86, 89]. The method is essentially based on the idea utilizing a technique similar to that for the linear systems. According to the method, the nonlinear equation (2.1) is rewritten into a linearized form such as

$$\dot{x} = \tilde{A}(x, t)x + B(t)u \tag{2.25}$$

where $\tilde{A}$ is an $n \times n$-state-dependent matrix. $\tilde{A}$ must satisfy

$$\tilde{A}(x, t)x = A(t)x + \varepsilon f(x, t) \tag{2.26}$$

Note that, when $n > 1$, the matrix $\tilde{A}$ can not uniquely be determined by Eq. (2.26). Pearson has proposed a suboptimal control

$$u_s = - R^{-1}(t)B'(t)\tilde{P}(x, t)x \tag{2.27}$$

where $\tilde{P}$ is the positive definite solution of the equation:

$$\dot{\tilde{P}} = - \tilde{P}\tilde{A} - \tilde{A}'\tilde{P} + \tilde{P}E\tilde{P} - C'QC \tag{2.28}$$

satisfying

$$\tilde{P}(x, T) = C'(T)FC(T) \tag{2.29}$$

This idea much simplifies the calculation formally; however, since Eq. (2.28) contains the state x, we again fall into the two-point boundary-value problem of Eqs. (2.25), (2.28), (2.2), and (2.29) with (2.27). For the time-invariant problem with an infinite control duration, Eq. (2.28) is reduced to the algebraic equation. In this case, we require to obtain the positive definite solution of the algebraic equation (2.28) with $\dot{\tilde{P}} = 0$ analytically, if possible. We can do it for a one-dimensional problem [89], but not for high-dimensional ones. In addition, the best possible solution can not necessarily be obtained, except for a one-dimensional problem, because of the arbitrariness in choosing the matrix $\tilde{A}$. N. Sannomiya [85] has given a necessary condition for Pearson's control to be optimal.

In short, the method of instantaneous linearization is not necessarily effective for systems of high dimension.

2.6 Uniqueness of the Solution

A systematic procedure to construct a suboptimal control will be proposed in the succeeding sections. Before that, we see the fact that, for a sufficiently small value of ε, there exists the unique optimal solution neighboring that of the unperturbed linear problem. To do this, an existence theorem of M. Urabe [95] for multi-point boundary-value problems is useful.

We consider the differential equation

$$\dot{v} = g(v,\ t) \tag{2.30}$$

with the L-point boundary condition

$$\sum_{i=1}^{L} G_i v(t_i) = h \qquad (t_1 < t_2 < \cdots < t_L) \tag{2.31}$$

where v is the k-vector and $g(v,\ t)$, the k-vector function, is continuously differentiable with respect to v in a domain Δ, Δ being a given domain of the tv-space intercepted by the two hyperplanes $t = t_1$ and $t = t_L$; G_i are $k \times k$-matrices and h is a k-vector. Assume that Eq. (2.30) has an approximate solution $v = v_s(t)$ lying in Δ and satisfying approximately the boundary condition (2.31).

The following lemma summarizes the fact concerning the existence of the solution of Eq. (2.30) with (2.31).

Lemma 2.1

Suppose that, for the approximate solution $v_s(t)$, there exist a $k \times k$-matrix $M(t)$ continuous in $t \in [t_1, t_L]$, a positive constant ρ, and a nonnegative constant γ (< 1) such that the $k \times k$-matrix

$$U \triangleq \sum_{i=1}^{L} G_i \Phi(t_i) \tag{2.32}$$

is nonsingular and such that*

$$\left\| \frac{\partial g}{\partial v} - M(t) \right\| \leq \frac{\gamma}{\alpha_1} \qquad \text{for any} \quad (t, v) \in \Delta_\rho \tag{2.33}$$

$$\frac{\alpha_1 \lambda + \alpha_2 \nu}{1 - \gamma} \leq \rho \tag{2.34}$$

where

$$\Delta_\rho \triangleq \left\{ (t, v) \mid \|v - v_s(t)\| \leq \rho, \quad t \in [t_1, t_L] \right\} \subset \Delta \tag{2.35}$$

Here, $\Phi(t)$ is the fundamental matrix of the linear homogeneous equation

$$\dot{a} = M(t)a \tag{2.36}$$

with the initial condition $\Phi(t_1) = I^{(k)}$, $I^{(k)}$ being the $k \times k$-identity matrix; α_1 is a positive constant such that

---

* The symbol $\|v\|$ denotes a norm of a vector v and $\|M\|$ a norm of a matrix M compatible with the vector norm previously defined, i.e., such that $\|Mv\| \leq \|M\|\|v\|$ holds for any M and v [22, p. 106].

$$\sqrt{2\int_{t_1}^{t_L}\|Z(t,\ \xi)\|^2 d\xi} \leq \alpha_1 \qquad \text{for any} \quad t \in [t_1,\ t_L] \tag{2.37}$$

where $Z(t,\ \xi)$ is the $k \times k$-piecewise continuous matrix defined by

$$Z(t,\ \xi) \triangleq \begin{cases} \Phi(t)[I^{(k)} - W]\Phi^{-1}(\xi) & \text{for} \quad \xi < t \\ -\ \Phi(t)W\Phi^{-1}(\xi) & \text{for} \quad \xi \geqq t \end{cases},$$

$$W \triangleq U^{-1} \sum_{i=\ell+1}^{L} G_i\Phi(t_i)$$

$$\text{for} \quad t_\ell \leq \xi < t_{\ell+1} \qquad (\ell = 1,\ 2,\ \ldots,\ L - 1) \tag{2.38}$$

The nonnegative constants α_2, λ, and ν are such that

$$\|\Phi(t)U^{-1}\| \leq \alpha_2 \qquad \text{for any} \quad t \in [t_1,\ t_L] \tag{2.39}$$

$$\|\dot{v}_s(t) - g[v_s(t),\ t]\| \leq \lambda \qquad \text{for any} \quad t \in [t_1,\ t_L] \tag{2.40}$$

$$\left\|\sum_{i=1}^{L} G_i v_s(t_i) - h\right\| \leq \nu \tag{2.41}$$

respectively. Then Eq. (2.30) has one and only one solution $v = v^*(t)$ satisfying the boundary condition (2.31) in Δ_ρ, and furthermore, for this solution, it holds that

$$\|v^*(t) - v_s(t)\| \leq \frac{\alpha_1\lambda + \alpha_2\nu}{1 - \gamma} \tag{2.42}$$

To apply the lemma to our problem, we combine Eqs. (2.11) and (2.6) to have the 2n-dimensional system:

$$\dot{v} = g(v,\ t),$$

$$v \triangleq \begin{bmatrix} x \\ p \end{bmatrix}, \qquad g(v,\ t) \triangleq \begin{bmatrix} A(t) & -E(t) \\ -C'(t)Q(t)C(t) & -A'(t) \end{bmatrix} v + \varepsilon \begin{bmatrix} f(x,\ t) \\ -(\frac{\partial f}{\partial x})'p \end{bmatrix} \tag{2.43}$$

The boundary conditions (2.2) and (2.7) are written as

$$G_1 v(\tau) + G_2 v(T) = h,$$

$$G_1 \triangleq \begin{bmatrix} I^{(n)} & 0 \\ 0 & 0 \end{bmatrix}, \qquad G_2 \triangleq \begin{bmatrix} 0 & 0 \\ -C'(T)FC(T) & I^{(n)} \end{bmatrix}, \qquad h \triangleq \begin{bmatrix} x^0 \\ 0 \end{bmatrix} \tag{2.44}$$

Application of Lemma 2.1 requires an approximate solution of Eqs. (2.43) with (2.44). Fortunately, the smallness of the parameter ε suggests use of the solution for the unperturbed system:

$$v_s(t) = \begin{bmatrix} x_s(t) \\ p_s(t) \end{bmatrix}, \qquad x_s(t) \triangleq \Psi(t)x^0, \qquad p_s(t) \triangleq P(t)x_s(t) \tag{2.45}$$

as an approximate solution. In Eqs. (2.45) $\Psi(t)$ is the fundamental matrix of Eqs. (2.17) satisfying $\Psi(\tau) = I^{(n)}$. Note that $v = v_s(t)$ satisfies the boundary condition (2.44) exactly.

<u>Theorem 2.1</u>

There exists a positive number ε_m such that, for any ε being $|\varepsilon| \leq \varepsilon_m$, Eqs. (2.43) have one and only one solution $v = v^*(t)$ satisfying the boundary condition (2.44) in the domain Δ_ρ, where

$$\Delta_\rho \triangleq \{(t,\ v) \mid \|v - v_s(t)\| \leq \rho,\quad t \in [\tau,\ T]\} \tag{2.46}$$

Furthermore the unique solution $v = v^*(t)$ satisfies

$$\|v^*(t) - v_s(t)\| \leq \frac{|\varepsilon|}{\varepsilon_m}\rho \tag{2.47}$$

Proof

Let us take the canonical equation for the unperturbed linear system as Eq. (2.36) in Lemma 2.1 so that

$$M(t) = \begin{bmatrix} A(t) & -E(t) \\ -C'(t)Q(t)C(t) & -A'(t) \end{bmatrix} \tag{2.48}$$

Thereupon we partition the $2n \times 2n$-fundamental matrix $\Phi(t)$ of the system (2.36) into four $n \times n$-submatrices:

$$\begin{gathered} \Phi(t) \triangleq \begin{bmatrix} \Phi_{11}(t) & \Phi_{12}(t) \\ \Phi_{21}(t) & \Phi_{22}(t) \end{bmatrix}, \\ \Phi_{ii}(\tau) = I^{(n)}, \qquad \Phi_{ij}(\tau) = 0 \qquad (i,\ j = 1,\ 2;\ i \neq j) \end{gathered} \tag{2.49}$$

Therefore, by Eq. (2.32), we have

$$\begin{gathered} U = \begin{bmatrix} I^{(n)} & 0 \\ \Phi_{21}(T) - C'(T)FC(T)\Phi_{11}(T) & Y \end{bmatrix}, \\ Y \triangleq \Phi_{22}(T) - C'(T)FC(T)\Phi_{12}(T) \end{gathered} \tag{2.50}$$

Due to the invertibility of the matrix Y [3, p. 759], U is nonsingular. To check the condition (2.33), let us calculate the value of α_1 satisfying inequality (2.37). From Eqs. (2.38), we have

$$\int \|\Phi(t)[I^{(2n)} - W]\Phi^{-1}(\xi)\| \qquad \text{for} \quad \tau \leq \xi < t \leq T$$

$$\|Z(t,\ \xi)\| = \begin{cases} \\ \|\Phi(t) W \Phi^{-1}(\xi)\| & \text{for} \quad \tau \leq t \leq \xi < T \end{cases},$$

$$W = U^{-1} G_2 \Phi(T) = \begin{bmatrix} 0 & 0 \\ Y^{-1}[\Phi_{21}(T) - C'(T)FC(T)\Phi_{11}(T)] & I^{(n)} \end{bmatrix} \tag{2.51}$$

Due to the smoothness of each quantity in Eqs. (2.51), there exists a nonnegative number η_1 such that

$$\|Z(t,\ \xi)\| \leq \frac{\eta_1}{\sqrt{2}} \qquad \text{for any} \quad t,\ \xi \in [\tau,\ T] \tag{2.52}$$

Hence we have

$$\sqrt{2\int_{\tau}^{T} \|Z(t,\ \xi)\|^2 d\xi} \leq (T - \tau)\eta_1 \tag{2.53}$$

which allows us to suppose that

$$\alpha_1 = (T - \tau)\eta_1 \tag{2.54}$$

By similar reasoning, there exists η_2 ($\geqq 0$) such that for M of Eq. (2.48)

$$\left\|\frac{\partial g}{\partial v} - M(t)\right\| = |\varepsilon| \left\| \begin{matrix} \frac{\partial f}{\partial x} & 0 \\ -\frac{\partial^2 (p'f)}{\partial x^2} & -(\frac{\partial f}{\partial x})' \end{matrix} \right\| \leq |\varepsilon|\eta_2 \quad \text{for any} \quad (t,\ v) \in \Delta_\rho \tag{2.55}$$

Thus, we see that the inequality (2.33) is true if

$$|\varepsilon| \leq \frac{\gamma}{(T - \tau)\eta_1\eta_2} \tag{2.56}$$

Lastly, let us check the condition (2.34). For the approximate solution (2.45), there exists a nonnegative number η_3 such that*

$$\|\dot{v}_s(t) - g[v_s(t), t]\| = |\varepsilon| \left\| \begin{matrix} f[\Psi(t)x^0, t] \\ -\left\{\dfrac{\partial f[\Psi(t)x^0, t]}{\partial x}\right\}' P(t)\Psi(t)x^0 \end{matrix} \right\| \leq |\varepsilon|\eta_3 \tag{2.57}$$

We may therefore suppose that

$$\lambda = |\varepsilon|\eta_3 \tag{2.58}$$

Since $v_s(t)$ satisfies the boundary condition (2.44) exactly, the constant ν in the inequality (2.41) may vanish. Thus the condition (2.34) is satisfied if

$$|\varepsilon| \leq \frac{(1 - \gamma)\rho}{(T - \tau)\eta_1\eta_3} \tag{2.59}$$

From the inequalities (2.56) and (2.59), we have finally

$$|\varepsilon| \leq \varepsilon_m \tag{2.60}$$

where

$$\varepsilon_m \triangleq \frac{1}{(T - \tau)(\eta_2 + \eta_3/\rho)\eta_1} \tag{2.61}$$

Let ρ be an arbitrary positive number. Then for any ε satisfying the inequality (2.60), the inequalities (2.56) and (2.59) hold, if we take

---

* For a function $g(x)$, the conventional expression $\partial g(a)/\partial x$ indicates the value of $\partial g/\partial x$ at $x = a$.

γ ($\geqq 0$) as

$$\gamma = \frac{1}{1 + \eta_3/\rho\eta_2} < 1 \tag{2.62}$$

This proves that, for any ε such that $|\varepsilon| \leq \varepsilon_m$, all the conditions of Lemma 2.1 are satisfied by Eqs. (2.43) and their approximate solution (2.45). Substituting Eqs. (2.54), (2.58), and (2.62) into the inequality (2.42) yields (2.47).

Q.E.D.

The positive number ε_m given by Eq. (2.61) is a bound of ε within which Eqs. (2.43) are assured to have the unique solution. The quantity $|\varepsilon|\rho/\varepsilon_m$ in the inequality (2.47) gives an error bound of the zero-order approximate solution $v_s(t)$.

2.7 Recurrence Equations for a Suboptimal Control

For the nonlinear system (2.1) with a sufficiently small value of ε, it may be reasonable to assume that the optimal feedback control be analytic in ε for all $t \in [\tau, T]$. Therefore we try to find a suboptimal control in a power-series form in ε.

We develop the costate vector $p(x, t)$ satisfying Eq. (2.6) into a power series with respect to ε:

$$p(x, t) = P(t)x + \sum_{k=1}^{\infty} \varepsilon^k p^{(k)}(x, t) \tag{2.63}$$

Differentiation of Eq. (2.63) with respect to t and use of Eq. (2.11) gives

$$\dot{p} = \dot{P}x + \sum_{k=1}^{\infty} \varepsilon^k \frac{\partial p^{(k)}}{\partial t} + [P + \sum_{k=1}^{\infty} \varepsilon^k \frac{\partial p^{(k)}}{\partial x}](Ax + \varepsilon f - Ep) \tag{2.64}$$

Substituting Eqs. (2.63) and (2.64) into (2.6) and equating the coefficients of the like powers of ε separately yields a sequence of equations:

$$\frac{\partial p^{(k)}}{\partial t} + \frac{\partial p^{(k)}}{\partial x} Sx + S'p^{(k)} = h^{(k)} \qquad (k = 1, 2, \ldots) \tag{2.65}$$

with the boundary conditions

$$p^{(k)}(x, T) = 0 \qquad \text{for any } x \tag{2.66}$$

The functions $h^{(k)}(x, t)$ are the n-vectors defined by

$$\begin{aligned} h^{(1)}(x, t) &\triangleq - (\frac{\partial f}{\partial x})'Px - Pf, \\ h^{(k)}(x, t) &\triangleq - (\frac{\partial f}{\partial x})'p^{(k-1)} - \frac{\partial p^{(k-1)}}{\partial x} f + \sum_{i=1}^{k-1} \frac{\partial p^{(i)}}{\partial x} Ep^{(k-i)} \\ &\qquad\qquad (k = 2, 3, \ldots) \end{aligned} \tag{2.67}$$

and S is the matrix defined by Eq. (2.18). The zero-order equation is identical with Eq. (2.14).

Successive solutions of Eqs. (2.65) with increasing k may determine the functions $p^{(k)}(x, t)$. Though Eqs. (2.65) are linear in $p^{(k)}$, their solutions are difficult given a general form of the perturbation $f(x, t)$. However, if $f(x, t)$ is a polynomial in x, $p^{(k)}(x, t)$ are also polynomials in x and their coefficients are exactly determined under appropriate conditions, as will be seen in the subsequent sections.

2.8 Solutions of the Recurrence Equations

In order to discuss the solutions of the recurrence equations (2.65), we consider the vector-matrix equation

$$\frac{\partial}{\partial t}g(x, t) + [\frac{\partial}{\partial x}g(x, t)]S(t)x + S'(t)g(x, t) = \frac{\partial}{\partial x}\theta(x, t) \tag{2.68}$$

with the boundary condition

$$g(x, T) = 0 \qquad \text{for any } x \tag{2.69}$$

$g(x, t)$ is an n-vector and $\theta(x, t)$ is a scalar. The function $\theta(x, t)$ is assumed to be continuous in t. The problem is to determine the function g given a polynomial θ.

The function θ is assumed to be, for the time being, a homogeneous polynomial of the degree, say, $r + 1$ in x, and is written as

$$\theta(x, t) = \frac{1}{r+1} \sum_{j_1, \ldots, j_{r+1}=1}^{n} \Theta_{j_1 j_2 \cdots j_{r+1}}(t) x_{j_1} x_{j_2} \cdots x_{j_{r+1}} \tag{2.70}$$

The coefficients $\Theta_{j_1 j_2 \cdots j_{r+1}}(t)$ are $(r + 1)$-index quantities, each index $j_1, j_2, \ldots, j_{r+1}$ running from 1 to n. Θ are assumed to be, without loss of generality, symmetric with respect to any pair of indices.* Hence we have $(n + r)!/(n - 1)!(r + 1)!$ different coefficients. In what follows, we use the summation convention that a twice repeated index is to be interpreted as a summation over that index from 1 to n.** By using this convention Eq. (2.70) is simply written as

$$\theta(x, t) = \frac{1}{r+1}\Theta_{j_1 j_2 \cdots j_{r+1}}(t) x_{j_1} x_{j_2} \cdots x_{j_{r+1}} \tag{2.71}$$

The i-th component of $\partial\theta/\partial x$, the gradient of θ, is given by

$$\frac{\partial\theta}{\partial x_i} = \Theta_{ij_1 \cdots j_r} x_{j_1} \cdots x_{j_r} \tag{2.72}$$

From Eq. (2.68) it easily follows that the i-th component of g is a homogeneous polynomial of the degree r in x. Hence one may write the i-th component of g as

$$g_i(x, t) = G^i_{j_1 j_2 \cdots j_r}(t) x_{j_1} x_{j_2} \cdots x_{j_r} \tag{2.73}$$

where the coefficients $G^i_{j_1 \cdots j_r}$ are symmetric with respect to the lower indices and satisfy

$$G^i_{j_1 j_2 \cdots j_r}(T) = 0 \tag{2.74}$$

Substituting Eqs. (2.72) and (2.73) into (2.68) leads to

---

* r-index quantities $M_{j_1 j_2 \cdots j_r}$ are said to be symmetric with respect to any pair of indices, or simply with respect to indices, if the values of M are invariant with any interchange of indices, i.e., if

$$M_{j_1 \cdots j_{i-1} j_i j_{i+1} \cdots j_{k-1} j_k j_{k+1} \cdots j_r} = M_{j_1 \cdots j_{i-1} j_k j_{i+1} \cdots j_{k-1} j_i j_{k+1} \cdots j_r}$$

holds for any integers i and k $(1 \le i \le r,\ 1 \le k \le r)$.

** This convention is not quite wild and is often used in tensor analyses, called Einstein's summation convention [51, p. 259].

$$(\dot{G}^i_{j_1 \cdots j_r} + rS_{kj_1} G^i_{kj_2 \cdots j_r} + S_{ki} G^k_{j_1 \cdots j_r} - \Theta_{ij_1 \cdots j_r}) x_{j_1} \cdots x_{j_r} = 0$$

$$(i = 1, 2, \ldots, n) \qquad (2.75)$$

where S_{ij} is the (i, j)-element of the matrix S. Equations (2.75) hold for arbitrary values of x if and only if the coefficients of $x_{j_1} x_{j_2} \cdots x_{j_r}$ vanish for any combinations of the values of indices. Hence we have $n(n + r - 1)!/(n - 1)!r!$ simultaneous equations for $G^i_{j_1 \cdots j_r}$:

$$\dot{G}^i_{j_1 \cdots j_r} = - S_{kj_1} G^i_{kj_2 \cdots j_r} - S_{kj_2} G^i_{j_1 k j_3 \cdots j_r} - \cdots$$

$$- S_{kj_r} G^i_{j_1 \cdots j_{r-1} k} - S_{ki} G^k_{j_1 \cdots j_r} + \Theta_{ij_1 \cdots j_r}$$

$$(i, j_1, \ldots, j_r = 1, 2, \ldots, n) \qquad (2.76)$$

with the terminal conditions (2.74).

Now the following lemma can be established.

Lemma 2.2

The matrix $\partial g/\partial x$ given by Eq. (2.68) is symmetric, and consequently there exists a scalar function $\phi(x, t)$ such that

$$g = \frac{\partial \phi}{\partial x} \qquad \text{for any} \quad t \in [\tau, T] \qquad (2.77)$$

Thus, Eq. (2.68) is equivalent to

$$\frac{\partial g}{\partial t} + \frac{\partial}{\partial x}(x'S'g) = \frac{\partial \theta}{\partial x} \qquad (2.78)$$

The coefficients of the polynomial g are determined by solving Eqs.

(2.76).

Proof

Since the coefficients Θ are symmetric with respect to all the indices, the coefficients G are also symmetric with respect to all the indices, including the upper and the lower ones. This is because Eqs. (2.76) are invariant with interchange of the upper index and any lower indices of G. It implies that $\partial g_i/\partial x_j = \partial g_j/\partial x_i$ for any pair of i and j. It is known [61, p. 14] that, when $\partial g/\partial x$ is symmetric, there exists a scalar function ϕ satisfying Eq. (2.77).

Q.E.D.

In the above discussion, the function θ is assumed to be a homogeneous polynomial in x. If θ is composed of terms of different degrees, the calculations should be done separately for different degree terms. Due to the linearity of Eq. (2.68), the total solution is obtained by simply summing up the resulting g polynomials.

By virtue of Lemma 2.2, we can show an important theorem to determine the polynomial $p^{(k)}(x, t)$. This theorem exhibits an important feature of the solution.

Theorem 2.2

The matrix $\partial p^{(k)}/\partial x$ given by Eq. (2.65) is symmetric for every $k \geq 1$. The coefficients of the polynomial $p^{(k)}$ are determined by solving Eqs. (2.76) separately for different degrees in x.

Proof

The function $h^{(1)}$ in Eqs. (2.67) is integrable with respect to x:

$$h^{(1)} = \frac{\partial}{\partial x}(- x'Pf) \tag{2.79}$$

Equation (2.79) is of the right-side form of Eq. (2.68). Due to Lemma 2.2, $\partial p^{(1)}/\partial x$ is symmetric. Then $h^{(2)}$ in Eqs. (2.67) is integrable with respect to x, and is rewritten into

$$h^{(2)} = \frac{\partial}{\partial x}[- f' p^{(1)} + \frac{1}{2} p^{(1)\prime} E p^{(1)}] \tag{2.80}$$

This is again of the right-side form of Eq. (2.68), and consequently $\partial p^{(2)}/\partial x$ is symmetric. Proceeding similarly, $h^{(k)}$ is rewritten into

$$h^{(k)} = \frac{\partial}{\partial x}[- f' p^{(k-1)} + \frac{1}{2} \sum_{i=1}^{k-1} p^{(i)\prime} E p^{(k-i)}] \tag{2.81}$$

and $\partial p^{(k)}/\partial x$ is known to be symmetric for every $k \geq 3$.

Q.E.D.

2.9 Extended Liapunov Equation

Here we make a lemma ready for examining, in the next section, the stability of the suboptimal control function. The following lemma regarding an extended Liapunov equation gives a backing to rewrite Eqs. (2.76) into a conventional vector-matrix form.

<u>Lemma 2.3</u>

Consider the n^r simultaneous algebraic equations for n^r quantities $X_{j_1 j_2 \cdots j_r}$:

$$M_{kj_1} X_{kj_2 \cdots j_r} + M_{kj_2} X_{j_1 k j_3 \cdots j_r} + \cdots + M_{kj_r} X_{j_1 \cdots j_{r-1} k}$$
$$= Y_{j_1 j_2 \cdots j_r} \qquad (j_1, j_2, \ldots, j_r = 1, 2, \ldots, n) \tag{2.82}$$

where M_{ij} and $Y_{j_1 j_2 \cdots j_r}$ are given quantities. All the indices run from 1 to n. Equations (2.82) can be rewritten into the vector-matrix form:

$$M^{(r)} v^{(r)} = b^{(r)} \tag{2.83}$$

where $M^{(r)}$ is the $n^r \times n^r$-matrix defined by

$$M^{(r)} \triangleq \underbrace{M' \oplus M' \oplus \ldots \oplus M'}_{r} \tag{2.84}$$

M is the $n \times n$-matrix such that the (i, j)-element is M_{ij}; $v^{(r)}$ and $b^{(r)}$ are the n^r-vectors of which the ℓ-th components are $X_{j_1 \cdots j_r}$ and $Y_{j_1 \cdots j_r}$, respectively, ℓ being

$$\ell \triangleq \sum_{k=1}^{r} (j_k - 1)n^{r-k} + 1 \tag{2.85}$$

The symbol $\oplus$ denotes the Kronecker sum of matrices.*

Proof

Let us proceed inductively. First, for $r = 2$, Eqs. (2.82) reduce to

$$M_{kj_1} X_{kj_2} + M_{kj_2} X_{j_1 k} = Y_{j_1 j_2} \tag{2.86}$$

By introducing the matrices X and Y whose (i, j)-elements are

---

* For definitions and related properties of the Kronecker product and sum, refer to Appendix I.

X_{ij} and Y_{ij}, respectively, Eqs. (2.86) are rewritten into

$$M'X + XM = Y \tag{2.87}$$

In terms of $v^{(2)}$ and $b^{(2)}$ as defined in the lemma, Eq. (2.87) is equivalent to, from Appendix I,

$$(M' \oplus M')v^{(2)} = b^{(2)} \tag{2.88}$$

Secondly we show that the lemma is true for $r = \ell + 1$ if it is true for $r = \ell$. For $r = \ell + 1$, Eqs. (2.82) are written as

$$\begin{aligned}(M_{kj_1} X_{kj_2 \cdots j_{\ell+1}} &+ M_{kj_2} X_{j_1 k \cdots j_{\ell+1}} + \cdots \\ &+ M_{kj_\ell} X_{j_1 \cdots j_{\ell-1} k j_{\ell+1}}) + M_{kj_{\ell+1}} X_{j_1 \cdots j_\ell k} = Y_{j_1 \cdots j_{\ell+1}}\end{aligned} \tag{2.89}$$

The terms in the parenthesis are identical, for a fixed value of $j_{\ell+1}$, with those on the left side of Eqs. (2.82). Hence, with use of $v^{(\ell+1)}$ and $b^{(\ell+1)}$ Eqs. (2.89) are rewritten into

$$[I^{(n)} \otimes M^{(\ell)} + M' \otimes I^{(n^\ell)}]v^{(\ell+1)} = b^{(\ell+1)} \tag{2.90}$$

where the symbol $\otimes$ denotes the Kronecker product of matrices and $I^{(k)}$ is the $k \times k$-identity matrix. Due to the definition of the Kronecker sum, Eq. (2.90) is equivalent to

$$M^{(\ell+1)} v^{(\ell+1)} = b^{(\ell+1)} \tag{2.91}$$

Q.E.D.

Let λ_i $(i = 1, 2, \ldots, n)$ be the eigenvalues of M' or equivalently of M. Since, from Appendix I, the n^r eigenvalues of $M^{(r)}$ are given by $\lambda_{j_1} + \lambda_{j_2} + \ldots + \lambda_{j_r}$ $(j_1, j_2, \ldots, j_r = 1, 2, \ldots, n)$, the following lemma concerning the uniqueness of the solution of Eqs. (2.82) is obtained. It will be used to establish Theorem 2.4 in the next section.

Lemma 2.4

If and only if any sum of r eigenvalues of M, $\lambda_{j_1} + \ldots + \lambda_{j_r}$ $(j_1, \ldots, j_r = 1, \ldots, n)$, is not zero, $X_{j_1 \ldots j_r}$ of Eqs. (2.82) are uniquely determined. In particular, if M is a stable matrix, $X_{j_1 \ldots j_r}$ are unique.

Furthermore, we can readily see that, if $Y_{j_1 \ldots j_r}$ are symmetric with respect to the indices, $X_{j_1 \ldots j_r}$ are also symmetric. Equations (2.82) are linear equations to determine the n-index quantities X given the two-index quantities M and the n-index quantities Y. For a special case of $n = 2$, so called the Liapunov equation [50], i.e., Eq. (2.87) is obtained. Hence Eqs. (2.82) are considered an extended version of the Liapunov equation. In fact, they may afford a Liapunov function of the n-form for linear systems.

2.10 Continuity and Stability of the Feedback Function

In this section the stability of the function $p^{(k)}$ is examined. To do this, we turn our attention back to the linear differential equations (2.76) again.

By virtue of Lemma 2.3 in Section 2.9, a lemma for an extended Liapunov equation, Eqs. (2.76) and (2.74) are rewritten into the vector-matrix form:

$$\dot{v} = -\tilde{S}(t)v + b(t), \qquad v(T) = 0 \tag{2.92}$$

where $\tilde{S}$ is the $(r + 1)$-times repeated Kronecker sum of S':

$$\tilde{S}(t) \triangleq \underbrace{S'(t) \oplus S'(t) \oplus \ldots \oplus S'(t)}_{r+1} \tag{2.93}$$

and v and b are the n^{r+1}-vectors such that the ℓ-th components are $G^{j_0}_{j_1 \cdots j_r}$ and $\Theta_{j_0 j_1 \cdots j_r}$, respectively, ℓ being

$$\ell \triangleq \sum_{k=0}^{r} (j_k - 1)n^{r-k} + 1 \tag{2.94}$$

Since Eqs. (2.92) satisfy a Lipschitz condition for ordinary differential equations, the solution of Eqs. (2.92) is unique and continuous in t [12, p. 22]. Therefore, by comparing the degrees in x in both sides of Eqs. (2.65), the following theorem is established:

<u>Theorem 2.3</u>

If $f(x, t)$ is a polynomial of the degree q in x, the coefficients being continuous in t, $p^{(k)}(x, t)$ given by Eq. (2.65) is uniquely determined and is the polynomial of the degree $k(q - 1) + 1$ in x whose coefficients are continuous in t. In particular, $p^{(1)}(x, t)$ is the polynomial of the same degree as $f(x, t)$.

In Eqs. (2.92) and (2.93), if S and b are time-invariant, and if S is stable, $v(t)$ is uniformly asymptotically stable relative to v_0 as $t \to -\infty$, because $\tilde{S}$ is time-invariant and, due to Appendix I, is stable, where v_0 is the solution of the linear algebraic equation

$$\tilde{S}v_0 = b \tag{2.95}$$

Due to Appendix I, Eq. (2.95) is uniquely solvable under the given condition. Thus the following lemma is obtained:

Lemma 2.5

In Eq. (2.68), if S reduces to a stable constant matrix S_0 and $\theta(x, t)$ reduces to be free of explicit dependence on t as $t \to -\infty$, then the function $g(x, t)$ is uniformly asymptotically stable as $t \to -\infty$, relative to the polynomial $g_0(x)$ such that satisfies the equation:

$$x'S_0'g_0(x) = \theta_0(x), \qquad \theta_0(x) \triangleq \lim_{t \to -\infty} \theta(x, t) \tag{2.96}$$

The coefficients $G^i_{j_1 \cdots j_r}$ of the polynomial $g_0(x)$ are given by Eq. (2.95).

Further the following lemma is useful in Section 2.12.

Lemma 2.6

If the equation

$$\frac{\partial g}{\partial t} + \frac{\partial}{\partial x}(x'S'g) = 0 \tag{2.97}$$

holds for any value of x with the condition (2.69), and if the matrix $\partial g/\partial x$ is known to be symmetric, then g is the identically zero vector. For the time-invariant case, if

$$x'S_0'g_0 = 0 \qquad \text{or} \qquad g_0'S_0x = 0 \tag{2.98}$$

holds for any value of x with a stable S_0, and if $\partial g_0/\partial x$ is symmetric, g_0 is the identically zero vector.

The following stability theorem for the function $p^{(k)}(x, t)$ is a direct consequence of Lemma 2.5:

Theorem 2.4

Assume that the system of Eqs. (2.1) and (2.3) and the performance index (2.4) are time-invariant with $F = 0$ and $Q > 0$. If the unperturbed linear system is completely controllable and completely observable, then, for every $k \geq 1$, $p^{(k)}(x, t)$, the polynomial function in x, given by Eqs. (2.65) is uniformly asymptotically stable as $t \to -\infty$, relative to the polynomial $p_0^{(k)}(x)$ such that satisfies the equation:

$$x'S_0'p_0^{(k)} = \theta_0^{(k)} \qquad (k = 1, 2, \ldots) \tag{2.99}$$

where $\theta_0^{(k)}(x)$ are the scalar functions defined by

$$\begin{aligned} \theta_0^{(1)}(x) &\triangleq -x'P_0 f, \\ \theta_0^{(k)}(x) &\triangleq -f'p_0^{(k-1)} + \frac{1}{2}\sum_{i=1}^{k-1} p_0^{(i)}{}'Ep_0^{(k-i)} \qquad (k = 2, 3, \ldots) \end{aligned} \tag{2.100}$$

S_0 is the constant matrix defined by Eq. (2.23). The polynomial function $p_0^{(k)}(x)$, satisfying Eq. (2.99), exists and is unique for every $k \geq 1$.

Now we have completed the procedure for calculating a suboptimal control. The solution of Eqs. (2.65) or (2.99) up to the order ℓ in ε gives the suboptimal control of the ℓ-th order. It is noted that, when a time-invariant problem with $T = \infty$ is considered, the complete controllability and complete observability of the unperturbed system suffice, due to Theorem 2.4, to yield a suboptimal control of an arbitrary order.

If for any x and $t \in [\tau, T]$

$$x'Pf = 0 \tag{2.101}$$

holds identically, then, from Eq. (2.79), $h^{(1)} = 0$ and consequently $p^{(1)} = 0$ from Lemma 2.6. Further we readily see, by induction, that $p^{(k)} = 0$ for every $k \geqq 1$. Similarly we can conclude that, if for any x and $t \in [\tau, T]$

$$h^{(2)} = 0 \tag{2.102}$$

holds identically, $p^{(k)} = 0$ for every $k \geqq 2$. However, note that, even if $h^{(\ell)} = 0$ for a certain $\ell \geqq 3$, $p^{(k)}$ is not always zero for any $k \geqq \ell$. Thus we obtain the following theorem:

Theorem 2.5

If the identity (2.101) holds for any x and $t \in [\tau, T]$, the linear control (2.16) is the exactly optimal one; if the identity (2.102) holds for any x and $t \in [\tau, T]$, the control

$$u^* = -R^{-1}(t)B'(t)[P(t)x + \varepsilon p^{(1)}(x, t)] \tag{2.103}$$

is exactly optimal.

This theorem gives a sufficient condition that the optimal feedback control for a nonlinear problem is exactly determined.

2.11 Relationship between the Conditions for Controllability and Stability

This section comments on the analogy between the condition of Theorem 2.4 and the controllability of nonlinear systems. For simplicity we consider only a time-invariant system with perfect observations, i.e., with $C = I^{(n)}$ in Eq. (2.3).

R. E. Kalman has introduced the concept of local controllability for nonlinear control systems in Discussion on L. Markus and E. B. Lee's paper [59]: Consider the nonlinear system

$$\dot{x} = g(x, u), \qquad g(0, 0) = 0 \tag{2.104}$$

If there exists a neighborhood Δ_x of the orign $x = 0$ such that each point $x^0 \in \Delta_x$ can be transferred to the orign in a finite time interval, using an admissible control function, then the system (2.104) is said to be locally completely controllable.

A sufficient condition for local complete controllability is known [58, 59]:

<u>Lemma 2.7</u>

Assume that $g(x, u)$, $\partial g/\partial x$, and $\partial g/\partial u$ are continuous in x and $u \in \Delta_u$, where Δ_u is the control restraint containing the origin $u = 0$. If it holds that

$$\operatorname{rank}[\tilde{B} \;\; \tilde{A}\tilde{B} \;\; \ldots \;\; \tilde{A}^{n-1}\tilde{B}] = n \tag{2.105}$$

where $\tilde{A}$ and $\tilde{B}$ are the $n \times n$- and $n \times m$-constant matrices, respectively, defined by

$$\tilde{A} \triangleq \frac{\partial g(0, 0)}{\partial x}, \qquad \tilde{B} \triangleq \frac{\partial g(0, 0)}{\partial u} \tag{2.106}$$

then the system (2.104) is locally completely controllable.

We turn our attention to our present system (2.1) which is assumed time invariant. For this system, $\tilde{A}$ and $\tilde{B}$ of Eqs. (2.106) are

$$\tilde{A} = A + \varepsilon \frac{\partial f(0)}{\partial x}, \qquad \tilde{B} = B \tag{2.107}$$

It follows that, if $\partial f(0)/\partial x = 0$, the condition (2.105), for any ε, is equivalent to that of complete controllability of the unperturbed linear system. When the Maclaurin expansion of f contains no term linear in x, $\partial f(0)/\partial x = 0$ is true. Such f may be called purely nonlinear. We conclude that, if the system (2.1) is locally completely controllable and $f(x)$ is purely nonlinear, then Theorem 2.4 holds.

2.12 Performance Evaluation

A closed-form exact solution as stated in Theorem 2.5 is generally impossible. Impossibility of infinite calculation forces us truncation of the series, yielding a suboptimal control. In this section the quality of a suboptimal control is examined.

Assume that a suboptimal feedback control $u = u_s(x, t)$ is given by

$$u_s(x, t) = - R^{-1}(t)B'(t)p_s(x, t) \tag{2.108}$$

where $p_s(x, t)$ is a suboptimal feedback function satisfying Eqs. (2.65) up to a certain order. The value of the performance index $J = J_s$ resulting from the feedback control of Eq. (2.108) would be a function of τ and $x(\tau)$ and be written as

$$J_s = \frac{1}{2}y'(T)Fy(T) + \frac{1}{2}\int_{\tau}^{T} (y'Qy + u_s'Ru_s)dt = \phi_s[x(\tau), \tau] \tag{2.109}$$

Evidently $\phi_s(0, \tau) = 0$ if $p_s(0, t) = 0$ for any $t \in [\tau, T]$. Now our task is to examine the function $\phi_s[x(\tau), \tau]$. Differentiation of Eq. (2.109) with respect to τ and use of Eqs. (2.1), (2.3), and (2.108) gives:

$$-\frac{1}{2}x'C'QCx - \frac{1}{2}p_s'Ep_s = \frac{\partial\phi_s}{\partial\tau} + (Ax + \varepsilon f - Ep_s)'\frac{\partial\phi_s}{\partial x} \tag{2.110}$$

In Eq. (2.110) all the values are estimated at $t = \tau$.

We develop p_s and ϕ_s into power series in ε:

$$\begin{aligned} p_s(x, t) &= P(t)x + \sum_{k=1}^{\infty} \varepsilon^k p_s^{(k)}(x, t), \\ \phi_s[x(\tau), \tau] &= \frac{1}{2}x'(\tau)P(\tau)x(\tau) + \sum_{k=1}^{\infty} \varepsilon^k \phi_s^{(k)}[x(\tau), \tau] \end{aligned} \tag{2.111}$$

where the matrix P is the solution of Eq. (2.14), and $p_s^{(k)}$ and $\phi_s^{(k)}$ satisfy, respectively,

$$p_s^{(k)}(x, T) = 0, \qquad \phi_s^{(k)}(x, T) = 0 \qquad \text{for any } x \tag{2.112}$$

Substitution of Eqs. (2.111) into (2.110) results in a sequence of equations:

$$\frac{\partial\phi_s^{(k)}}{\partial\tau} + x'S'\frac{\partial\phi_s^{(k)}}{\partial x} = \psi^{(k)} \qquad (k = 1, 2, \ldots) \tag{2.113}$$

where

$$\psi^{(1)} \triangleq -x'Pf,$$

$$\psi^{(k)} \triangleq - f' \frac{\partial \phi_s^{(k-1)}}{\partial x} - \frac{1}{2} \sum_{i=1}^{k-1} p_s^{(i)\prime} E p_s^{(k-i)} + \sum_{i=1}^{k-1} p_s^{(k-i)\prime} E \frac{\partial \phi_s^{(i)}}{\partial x} \tag{2.114}$$

$$(k = 2,\ 3,\ \ldots)$$

Then the following theorem can be established:

<u>Theorem 2.6</u>

If $p_s(x, t)$ is the optimal one, i.e., if

$$p_s^{(k)} = p^{(k)} \qquad (k = 1,\ 2,\ \ldots) \tag{2.115}$$

where $p^{(k)}$ are the solutions of Eqs. (2.65), then $J_s = J^*$ is given by

$$J^* = \frac{1}{2} x'(\tau) P(\tau) x(\tau) + \sum_{k=1}^{\infty} \varepsilon^k \phi^{(k)} [x(\tau),\ \tau] \tag{2.116}$$

where $\phi^{(k)}[x(\tau),\ \tau]$ are the scalar functions satisfying

$$\frac{\partial \phi^{(k)}}{\partial x} = p^{(k)}, \qquad \phi^{(k)}(0,\ \tau) = 0 \qquad (k = 1,\ 2,\ \ldots) \tag{2.117}$$

<u>Proof</u>

The theorem can be proved by induction. First comparison of Eq. (2.113) with (2.65) for $k = 1$ gives

$$\frac{\partial}{\partial \tau} [\frac{\partial \phi_s^{(1)}}{\partial x} - p^{(1)}] + \frac{\partial}{\partial x} \left\{ x' S' [\frac{\partial \phi_s^{(1)}}{\partial x} - p^{(1)}] \right\} = 0 \tag{2.118}$$

By virtue of Lemma 2.6 in Section 2.10, the above equation, along with Eqs. (2.112), means that

$$\frac{\partial \phi_s^{(1)}}{\partial x} = p^{(1)} \tag{2.119}$$

Second we assume that

$$\frac{\partial \phi_s^{(k)}}{\partial x} = p^{(k)} \qquad (k = 1, 2, \ldots, \ell - 1) \tag{2.120}$$

holds. Then, by comparing Eq. (2.113) with (2.65) for $k = \ell$, we have

$$\frac{\partial}{\partial \tau}[\frac{\partial \phi_s^{(\ell)}}{\partial x} - p^{(\ell)}] + \frac{\partial}{\partial x}\left\{x'S'[\frac{\partial \phi_s^{(\ell)}}{\partial x} - p^{(\ell)}]\right\} = 0 \tag{2.121}$$

and consequently

$$\frac{\partial \phi_s^{(\ell)}}{\partial x} = p^{(\ell)} \tag{2.122}$$

Q.E.D.

Further the following theorem of interest is obtained:

Theorem 2.7

If $p_s(x, t)$ is optimal up to the order ℓ in ε, then J_s is equal to J^* up to the order $2\ell + 1$ in ε, that is,

$$\phi_s^{(k)} = \phi^{(k)} \qquad (k = 1, 2, \ldots, 2\ell + 1) \tag{2.123}$$

where $\phi^{(k)}$ are the functions defined in Eq. (2.117), and $\phi_s^{(2\ell+2)}$, the $(2\ell + 2)$-th term in the expansion of J_s, is given by

$$\frac{\partial}{\partial\tau}[\frac{\partial\phi_s^{(2\ell+2)}}{\partial x} - p^{(2\ell+2)}] + \frac{\partial}{\partial x}\left\{x'S'[\frac{\partial\phi_s^{(2\ell+2)}}{\partial x} - p^{(2\ell+2)}]\right\}$$

$$= -\frac{1}{2}\frac{\partial}{\partial x}\left\{[p_s^{(\ell+1)} - p^{(\ell+1)}]'E[p_s^{(\ell+1)} - p^{(\ell+1)}]\right\} \tag{2.124}$$

Proof

Again the theorem is proved by induction. First, corresponding to $\ell = 0$, assume that

$$p_s(x,\ t) = P(t)x \tag{2.125}$$

From Eqs. (2.65) and (2.113), it is readily observed that Eq. (2.118) holds. Due to Lemma 2.6, it implies, along with Eq. (2.117) with $k = 1$, that $\phi_s^{(1)} = \phi^{(1)}$. Second, assume that, if

$$p_s^{(i)} = p^{(i)} \qquad (i = 1,\ 2,\ \ldots,\ \ell - 1) \tag{2.126}$$

then

$$\phi_s^{(i)} = \phi^{(i)} \qquad (i = 1,\ 2,\ \ldots,\ 2\ell - 1) \tag{2.127}$$

Besides if $p_s^{(\ell)} = p^{(\ell)}$, comparison of Eq. (2.113) with (2.65) for $k = 2\ell$ leads to

$$\frac{\partial}{\partial\tau}[\frac{\partial\phi_s^{(2\ell)}}{\partial x} - p^{(2\ell)}] + \frac{\partial}{\partial x}\left\{x'S'[\frac{\partial\phi_s^{(2\ell)}}{\partial x} - p^{(2\ell)}]\right\} = 0 \tag{2.128}$$

Again it implies that $\phi_s^{(2\ell)} = \phi^{(2\ell)}$. Further, examination of Eqs. (2.113) and (2.65) for $k = 2\ell + 1$ leads to $\phi_s^{(2\ell+1)} = \phi^{(2\ell+1)}$. Comparison of Eq. (2.113) with (2.65) for $k = 2\ell + 2$ yields Eq. (2.124).

Q.E.D.

Theorem 2.7 is an extension of the theorem obtained by P. V. Kokotović and J. B. Cruz, Jr. [42] for linear systems. A similar result has been reported by R. A. Werner and J. B. Cruz, Jr. [96] regarding an approximate construction of the optimal feedback control for systems with unknown parameters; where the control is expanded in a Taylor series with respect to the unknown parameters. However the present proof is given in a much simpler way by use of induction.

It is noted that Theorems 2.6 and 2.7 for a time-invariant system with infinite T are also proven [66, 68, 71].

2.13 Examples

(a) One-Dimensional Problem

In order to examine the stabilization property of the suboptimal control obtained by the present method, let us consider a simple one-dimensional problem:

$$\dot{x} = ax + \varepsilon g(x)x + u, \qquad y = x \tag{2.129}$$

$$J = \frac{1}{2}\int_0^\infty (x^2 + u^2)dt \tag{2.130}$$

where a is a constant. The exactly optimal control of the problem can be obtained [89]:

$$u^* = -[a + \varepsilon g(x) + \sqrt{a^2 + 1 + 2\varepsilon a g(x) + \varepsilon^2 g(x)^2}\,]x \tag{2.131}$$

The control of Eq. (2.131) assures the global asymptotic stability of the system (2.129).

Let us apply our method to the present problem. The task is to obtain the solutions of Eqs. (2.99). Due to the one-dimensionality of the problem, this is possible with no explicit polynomial expression of the function $g(x)$. The power-series control obtained is given by

$$u^* = \sum_{k=0}^{\infty} \varepsilon^k u^{(k)} \tag{2.132}$$

where

$$\begin{aligned}
u^{(0)} &\triangleq -(a + \sqrt{a^2+1})x, \\
u^{(1)} &\triangleq -(1 + \frac{a}{\sqrt{a^2+1}})g(x)x, \\
u^{(2)} &\triangleq -\frac{1}{2\sqrt{(a^2+1)^3}}g(x)^2 x, \\
u^{(3)} &\triangleq \frac{a}{2\sqrt{(a^2+1)^5}}g(x)^3 x, \\
u^{(4)} &\triangleq -\frac{4a^2-1}{8\sqrt{(a^2+1)^7}}g(x)^4 x, \\
&\cdots\cdots
\end{aligned} \tag{2.133}$$

Equation (2.132) along with (2.133) agrees with the formal expansion of the exact solution (2.131) with respect to ε.

We proceed to examine stability of the system (2.129) with truncated power-series controls. Let u_ℓ be the ℓ-th order suboptimal control:

$$u_\ell = \sum_{k=0}^{\ell} \varepsilon^k u^{(k)} \tag{2.134}$$

The resultant system with $u = u_\ell$ is of the form

$$\dot{x} = - g_\ell(x;\ a)x \tag{2.135}$$

The function g_ℓ is a polynomial of $\varepsilon g(x)$ containing the parameter a. The global asymptotic stability of the system (2.135) is equivalent to the positive definiteness of g_ℓ. Since it is impossible to find whether g_ℓ is positive definite or not for arbitrary form of $g(x)$, we here confine $g(x)$ to the function with a definite sign. Table 2.1 shows the lowest approximation to the optimal control assuring the global asymptotic stability of Eq. (2.135). When $\varepsilon g(x) \geqq 0$ and $a \leqq - 3.243$, the global

Table 2.1 Lowest Approximation to the Optimal Control Assuring Global Asymptotic Stability of the System (2.135)

Value of a	Value of function	
	$\varepsilon g(x) \geqq 0$	$\varepsilon g(x) \leqq 0$
$0 \leqq a$	u_1	u_0
$-\sqrt{2} < a < 0$	u_2	u_0
$-2.473 < a \leqq -\sqrt{2}$	u_3	u_0
$-3.243 < a \leqq -2.473$	u_4	u_0
$a \leqq -3.243$		u_0

stability can not be assured unless $u = u_\ell$ with $\ell \geqq 5$ is used.

The simple one-dimensional case is above-mentioned. For a high-dimensional problem, the stability examination is more difficult, because of impossibility of finding the sign definiteness of a multivariable function. Thus we have not always to be a sticker for global stability but may use a suboptimal control assuring stability in a finite region.

(b) Second-Order Problem Described by Duffing's Equation

We illustrate the calculation of a suboptimal control for the second-order system described by Duffing's equation [31, p. 20]:

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= \varepsilon x_1^3 + u \end{aligned} \tag{2.136}$$

with the output

$$y_i = x_i \qquad (i = 1,\ 2) \tag{2.137}$$

The performance index is

$$J = \frac{1}{2}\int_0^\infty (x_1^2 + x_2^2 + u^2)dt \tag{2.138}$$

Since the control duration is infinite and the linear system with $\varepsilon = 0$ is completely controllable and completely observable,* the determination of a suboptimal control is reduced to calculation of the constant

* For the conditions of explicit criteria of complete controllability and complete observability, refer to any standard textbooks on control theories [e.g., 3, p. 205].

matrix P_0 and the vector functions $p_0^{(k)}$ satisfying Eqs. (2.21) and (2.99), respectively.* The positive definite solution P_0 of Eq. (2.21) and the matrix S_0 defined by Eq. (2.23) are, respectively,

$$P_0 = \begin{bmatrix} \sqrt{3} & 1 \\ 1 & \sqrt{3} \end{bmatrix}, \qquad S_0 = \begin{bmatrix} 0 & 1 \\ -1 & -\sqrt{3} \end{bmatrix} \tag{2.139}$$

The polynomial $p_0^{(1)}$ is cubic in x_1 and x_2 due to Theorem 2.3. Further, owing to the symmetry of the matrix $\partial p_0^{(1)}/\partial x$, $p_0^{(1)}$ may be written, in component forms, as

$$\begin{aligned} p_{01}^{(1)} &= a_1x_1^3 + 3a_2x_1^2x_2 + 3a_3x_1x_2^2 + a_4x_2^3, \\ p_{02}^{(1)} &= a_2x_1^3 + 3a_3x_1^2x_2 + 3a_4x_1x_2^2 + a_5x_2^3 \end{aligned} \tag{2.140}$$

where a_i are constants to be determined. Substituting Eqs. (2.140) into (2.99) with $k = 1$ and using Eqs. (2.139) yields

$$\begin{aligned} &- a_2x_1^4 + (a_1 - \sqrt{3}a_2 - 3a_3)x_1^3x_2 + 3(a_2 - \sqrt{3}a_3 - a_4)x_1^2x_2^2 \\ &+ (3a_3 - 3\sqrt{3}a_4 - a_5)x_1x_2^3 + (a_4 - \sqrt{3}a_5)x_2^4 = - x_1^4 - \sqrt{3}x_1^3x_2 \end{aligned} \tag{2.141}$$

Equating the coefficients of the like powers of x_i in both sides separately gives the five simultaneous linear algebraic equations in a_i. The solution is

$$a_1 = \frac{10\sqrt{3}}{13}, \qquad a_2 = 1, \qquad a_3 = \frac{10}{13\sqrt{3}}, \qquad a_4 = \frac{3}{13}, \qquad a_5 = \frac{\sqrt{3}}{13} \tag{2.142}$$

---

* This statement will be omitted in the subsequent examples of this section, but always should be understood.

In the like manner, $p_0^{(2)}$ is determined from Eq. (2.99) with $k = 2$. Thus the suboptimal control of the second order is

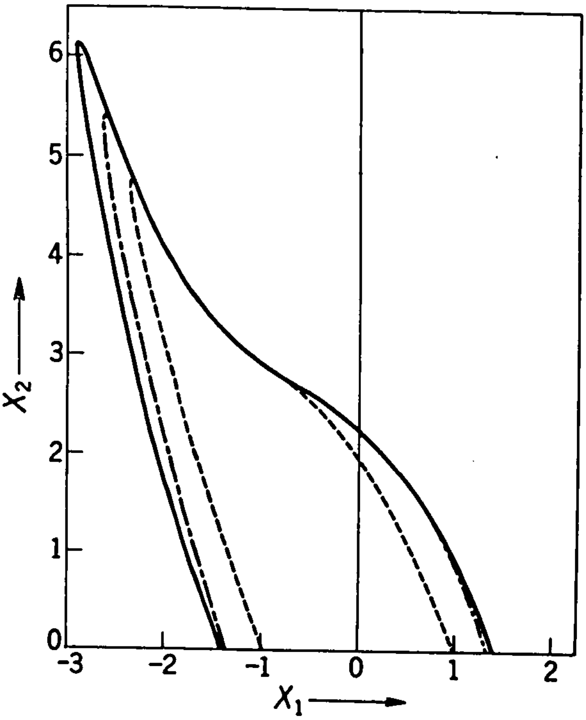
$$u_s = u^{(0)} + \varepsilon u^{(1)} + \varepsilon^2 u^{(2)} \tag{2.143}$$

where

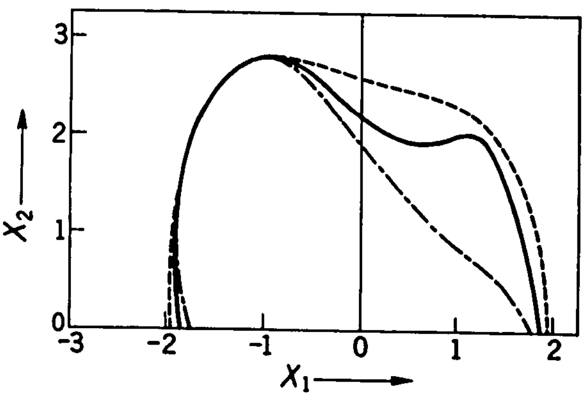
$$\begin{aligned}
u^{(0)} &\triangleq -x_1 - 1.7321x_2, \\
u^{(1)} &\triangleq -x_1^3 - 1.3323x_1^2x_2 - 0.6923x_1x_2^2 - 0.1332x_2^3, \\
u^{(2)} &\triangleq -0.5x_1^5 - 0.7392x_1^4x_2 - 0.3320x_1^3x_2^2 + 0.0190x_1^2x_2^3 \\
&\qquad + 0.0523x_1x_2^4 + 0.0112x_2^5
\end{aligned} \tag{2.144}$$

Now the stabilization property of the controller is examined. With $u_s = u^{(0)}$, the origin of Eqs. (2.136) is globally asymptotically stable with $\varepsilon < 0$, while it is only locally asymptotically stable with $\varepsilon > 0$. In fact, for $\varepsilon > 0$, there exist unstable singular points at $(\pm\sqrt{1/\varepsilon},\ 0)$, and trajectories passing through either of these points give the stability boundaries. With $u_s = u^{(0)} + \varepsilon u^{(1)}$, the origin becomes locally stable for $\varepsilon < 0$, so that the first-order approximation gives a better result than the zero-order one only in a limited region around the origin. On the other hand, with $\varepsilon > 0$, the first-order controller stabilizes the origin in the large. The second-order controller is of a locally stabilizing nature for both signs of ε.

Figure 2.1, showing the loci of the initial states for which $J = 5$, reflects the stabilization property mentioned above.* The curves are shown only on the upper-half plane, because they are symmetrical about the origin. Within the regions illustrated, the higher-order approxima-



(a) $\varepsilon = 1$



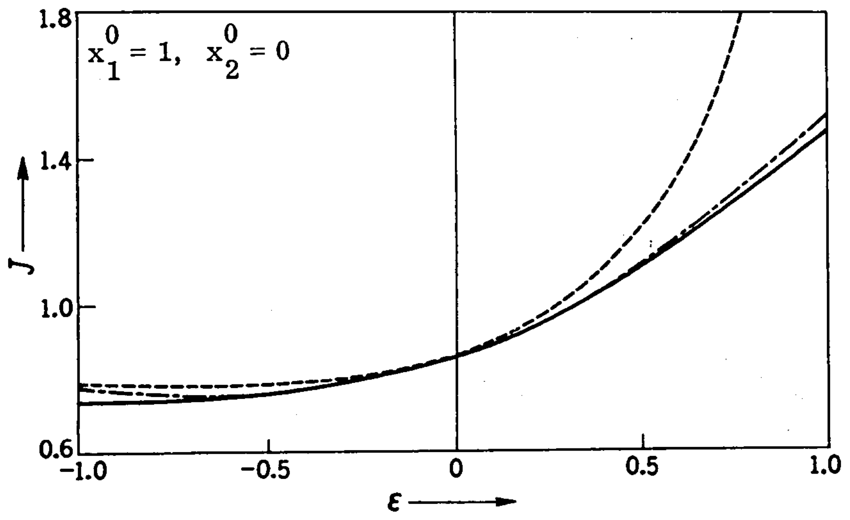
(b) $\varepsilon = -1$

———: J by $u_s = u^{(0)} + \varepsilon u^{(1)} + \varepsilon^2 u^{(2)}$

— - —: J by $u_s = u^{(0)} + \varepsilon u^{(1)}$

- - - -: J by $u_s = u^{(0)}$

Fig. 2.1. The loci of the initial states resulting J = 5.



———: J by $u_s = u^{(0)} + \varepsilon u^{(1)} + \varepsilon^2 u^{(2)}$

–––: J by $u_s = u^{(0)} + \varepsilon u^{(1)}$

-----: J by $u_s = u^{(0)}$

Fig. 2.2. Variation of the performance index J with ε for various approximations of the control.

tions improve the performance with $\varepsilon = 1$, but it is not necessarily with $\varepsilon = -1$.

Figure 2.2 illustrates examples of variation of J with ε for various approximations.

---

* This figure is obtained in the following way: We take various initial states $x_i(0)$ neighboring the origin, e.g., points on a circle of radius, say, 0.01, and we integral numerically Eqs. (2.136) and (2.138) in the negative time direction to obtain the states corresponding $J = -5$. The numerical integration using Runge-Kutta-Gill's process is carried out on a digital computer.

(c) Second-Order Problem Described by Van der Pol's Equation

We shall compare various suboptimal controls by the present method with those by a number of other known methods, using the system described by van der Pol's equation [31, p. 17]:

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= -x_1 + (1 - \varepsilon x_1^2)x_2 + u \end{aligned} \tag{2.145}$$

The performance index of Eq. (2.138) and the output of Eq. (2.137) are used.

The suboptimal control of the second order is determined by our method:

$$u_s = u^{(0)} + \varepsilon u^{(1)} + \varepsilon^2 u^{(2)} \tag{2.146}$$

where

$$\begin{aligned} u^{(0)} &\triangleq -0.4142x_1 - 2.6818x_2, \\ u^{(1)} &\triangleq 1.1162x_1^2x_2 + 0.5689x_1x_2^2 + 0.1128x_2^3, \\ u^{(2)} &\triangleq -0.1667x_1^4x_2 - 0.1505x_1^3x_2^2 - 0.0100x_1^2x_2^3 \\ &\qquad + 0.0291x_1x_2^4 + 0.0072x_2^5 \end{aligned} \tag{2.147}$$

There have been a number of suboptimal controls by various methods applied to this example. Let the typical suboptimal solutions be listed for the parameter $\varepsilon = 1$. This example has solved by J. D. Pearson [78], with use of the method of instantaneous linearization outlined in Section 2.5. His plausible assumption for the state-dependent matrix $\widetilde{A}$ in Eqs. (2.25) and (2.26) is

$$\widetilde{A} = \begin{bmatrix} 0 & 1 \\ -1 & 1 - x_1^2 \end{bmatrix} \tag{2.148}$$

and so his suboptimal control is

$$u_s = (1 - \sqrt{g^2 + 1})x_1 - (1 - x_1^2 + \sqrt{h^2 + 2g + 1 + 2\sqrt{g^2 + 1}})x_2 \tag{2.149}$$

$$g \triangleq -1, \qquad h \triangleq 1 - x_1^2 \tag{2.150}$$

Utilizing a minimax algorithm, N. Sannomiya [85] has shown that, for relatively large initial values of x_1 and x_2, the functions g and h defined by Eqs. (2.150) are the best in the forms of

$$g = -1 - cx_1x_2, \qquad h = 1 + (c - 1)x_1^2 \tag{2.151}$$

where c is a constant.

On the other hand, G. K. L. Kriechbaum and E. Noges [46] has proposed a least-squares approximation of nonlinear functions by a hyperplane through the origin and obtained the suboptimal control of the form of Eq. (2.149) with

$$g = -1 - \frac{2}{7}x_1x_2, \qquad h = 1 - \frac{5}{42}x_1^2 \tag{2.152}$$

W. L. Garrard, N. H. McClamroch, and L. G. Clark [27] and J. H. Burghart [11] have obtained a feedback gain matrix of power-series form for a linear model given in advance. Garrard et al. have solved approximately Hamilton-Jacobi equation to have the suboptimal control

$$u_s = -0.414x_1 - 2.682x_2 + 1.586x_1^2x_2 - 0.194x_1^4x_2 \qquad (2.153)$$

Burghart has found another form of control which contains the term x_1x_2. This seems a bit queer, since Eqs. (2.145) are symmetrical with respect to the origin $x_1 = 0$ and $x_2 = 0$.

Solutions based on a parameter optimization technique have been calculated by R. C. Durbeck [18] and A. G. Longmuir and E. V. Bohn [53]. Durbeck has assumed a reasonable form of function for the optimal performance index and had the suboptimal control

$$u_s = -0.080x_1 - 1.422x_2 + 0.219x_2^3 \qquad (2.154)$$

Longmuir et al. have had two possible controls under the assumption of control function with an appropriate form.

Table 2.2 shows how well the various controls perform for some initial conditions $x_1(0)$ and $x_2(0)$. The results by our method are added to this table. The performance obtained by the second-order control of our method is somewhat better than the others for the initial conditions considered. Other numerical examples for this problem are found in the reference [67].

(d) Rotational Motion Problem of a Rigid Body

Consider the third-order system governed by:

$$\begin{aligned} \dot{x}_1 &= \varepsilon a_1x_2x_3 + u_1, \\ \dot{x}_2 &= \varepsilon a_2x_3x_1 + u_2, \end{aligned} \qquad (2.155)$$

Table 2.2 Comparison of Values of the Performance Index for Various Approximations

Method	Initial condition $(x_1(0), x_2(0))$			
	(1, 1)	(2, 2)	(1, -1)	(2, -2)
Present method (second-order approximation)	2.560	6.614	2.002	3.920
Present method (first-order approximation)	2.573	∞	2.003	4.037
Present method (zero-order approximation)	2.699	8.108	2.034	5.107
Kriechbaum et al.	2.569	7.522	2.007	4.494
Garrard et al.	2.632	7.555	2.078	5.692
Durbeck	3.413	8.381	3.497	5.284
Pearson	2.616	6.796	2.072	4.492

$$\dot{x}_3 = \varepsilon a_3 x_1 x_2 + u_3$$

with the index of performance

$$J = \frac{1}{2}\int_0^{\infty} \sum_{i=1}^{3} \frac{1}{r_i}(q_i^2 x_i^2 + u_i^2)dt, \qquad (2.156)$$

$$q_i, \ r_i > 0 \qquad (i = 1, 2, 3)$$

The output is

$$y_i = x_i \qquad (i = 1, 2, 3) \qquad (2.157)$$

It is well known that, if

$$a_1 + a_2 + a_3 = 0 \tag{2.158}$$

Eqs. (2.155) describe the rotational motion of a rigid body.

First, the suboptimal control is calculated by the present procedure up to the third order:

$$u_{si} = u_i^{(0)} + \varepsilon u_i^{(1)} + \varepsilon^2 u_i^{(2)} + \varepsilon^3 u_i^{(3)} \qquad (i = 1,\ 2,\ 3) \tag{2.159}$$

where

$$\begin{aligned}
u_i^{(0)} &\triangleq - q_i x_i, \\
u_i^{(1)} &\triangleq c^{(1)} r_i x_j x_k, \\
u_i^{(2)} &\triangleq r_i x_i [c_k^{(2)} x_j^2 + c_j^{(2)} x_k^2], \\
u_i^{(3)} &\triangleq r_i x_j x_k [3c_i^{(3)} x_i^2 + c_j^{(3)} x_j^2 + c_k^{(3)} x_k^2], \\
c^{(1)} &\triangleq - \sum_{\ell=1}^{3} \frac{a_\ell q_\ell}{d r_\ell}, \\
c_i^{(2)} &\triangleq \frac{[a_i + c^{(1)} r_i/2] c^{(1)}}{d - q_i}, \\
c_i^{(3)} &\triangleq \frac{[a_j + c^{(1)} r_j] c_k^{(2)} + [a_k + c^{(1)} r_k] c_j^{(2)}}{d + 2q_i}
\end{aligned} \tag{2.160}$$

$$(i,\ j,\ k = 1,\ 2,\ 3;\ i \neq j \neq k \neq i),$$

$$d \triangleq q_1 + q_2 + q_3$$

It follows from Eqs. (2.160) that, if $c^{(1)} = 0$,

$$u_i^{(k)} = 0 \qquad (i = 1, 2, 3; k = 1, 2, \ldots) \tag{2.161}$$

This implies that $u_i^* = u_i^{(0)}$ is the exactly optimal control for this case. K. S. P. Kumar [47] has previously discussed a special case where $c^{(1)} = 0$, i.e.,

$$a_1 + a_2 + a_3 = 0, \qquad q_1 = q_2 = q_3 = 1, \qquad r_1 = r_2 = r_3 \tag{2.162}$$

which is one of norm-invariant problems treated in the next example.

Secondly, the method of instantaneous linearization by J. D. Pearson [78] is applied to Eqs. (2.155). For the system (2.155) the state-dependent matrix $\widetilde{A}$ in Eq. (2.25) is here assumed in the form

$$\widetilde{A} = \varepsilon \begin{bmatrix} 0 & a_1\nu_1 x_3 & a_1(1 - \nu_1)x_2 \\ a_2(1 - \nu_2)x_3 & 0 & a_2\nu_2 x_1 \\ a_3\nu_3 x_2 & a_3(1 - \nu_3)x_1 & 0 \end{bmatrix} \tag{2.163}$$

where ν_i are arbitrary constants. In order for Eq. (2.28) with $\widetilde{A}$ of Eq. (2.163) to have the exact solution with the parameters of Eqs. (2.162), it is required that

$$\begin{bmatrix} \nu_1 & 1 - \nu_2 & 0 \\ 1 - \nu_1 & 0 & \nu_3 \\ 0 & \nu_2 & 1 - \nu_3 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = 0 \tag{2.164}$$

Equation (2.164) holds for nontrivial values of the parameters a_i, if and only if arbitrary constants ν_i satisfy the following relation:

$$\nu_1\nu_2 + \nu_2\nu_3 + \nu_3\nu_1 + 1 = \nu_1 + \nu_2 + \nu_3 \tag{2.165}$$

It is noted that the form of $\tilde{A}$ should skillfully be chosen in order to obtain a good solution.

(e) Norm-Invariant Problem

In this example, we shall consider the control of a special class of systems: norm-invariant systems. The system (2.1) is called norm-invariant [4],* if A and f in Eq. (2.1) have the property that

$$x'[A(t)x + \varepsilon f(x, t)] = 0 \tag{2.166}$$

for any x and $t \in [\tau, T]$.

If the function $f(x, t)$ has no linear term in x, the condition (2.166) is equivalent to the conditions that A is skew-symmetric and that, for any x and $t \in [\tau, T]$,

$$x'f(x, t) = 0 \tag{2.167}$$

All the other matrices in Eq. (2.1) to (2.4) are assumed diagonal and constant:

$$B = C = I^{(n)}, \qquad F = gI^{(n)},$$

$$Q = \begin{bmatrix} q_1 & 0 & \cdots & 0 \\ 0 & q_2 & \cdots & 0 \\ & & \cdots\cdots & \\ 0 & 0 & \cdots & q_n \end{bmatrix}, \qquad R = \begin{bmatrix} r_1 & 0 & \cdots & 0 \\ 0 & r_2 & \cdots & 0 \\ & & \cdots\cdots & \\ 0 & 0 & \cdots & r_n \end{bmatrix}, \tag{2.168}$$

---

* The term "norm-invariant" is due to that, in the absence of the control, the Euclidian norm of the state vector remains constant.

$$q_i,\ r_i > 0 \quad (i = 1, 2, \ldots, n), \qquad g \geqq 0$$

If

$$q^2 \triangleq q_i, \quad r_i = 1 \quad (i = 1, 2, \ldots, n), \qquad q > 0 \tag{2.169}$$

then the solution of the Riccati equation (2.14) is given by

$$P(t) = q\frac{1 + h(t)}{1 - h(t)}I^{(n)}, \qquad h(t) \triangleq \frac{g - q}{g + q}e^{-2q(T-t)} \tag{2.170}$$

Due to Theorem 2.5 with Eqs. (2.167) and (2.170), the linear feedback control

$$u^* = -q\frac{1 + h}{1 - h}x \tag{2.171}$$

is the exactly optimal control for the present norm-invariant problem.

For the problem with infinite control duration, if

$$c^2 \triangleq q_i r_i \quad (i = 1, 2, \ldots, n), \qquad c > 0 \tag{2.172}$$

the solution of Eq. (2.21) is

$$P_0 = cI^{(n)} \tag{2.173}$$

and the optimal control is, in component form,

$$u_i^* = -\frac{c}{r_i}x_i \qquad (i = 1, 2, \ldots, n) \tag{2.174}$$

It is noted that neither the optimal control (2.171) nor (2.174) is related to A.

2.14 Concluding Remarks

A systematic procedure has been developed for the suboptimal design of a nonlinear state regulator. Theorem 2.1 guarantees the uniqueness of the optimal solution. When the nonlinearity is characterized by a polynomial function in the state, the definite way to determine a suboptimal feedback control of a power-series form is established by Theorems 2.2 and 2.3. The suboptimal control is yielded by solving a set of ordinary differential equations analytically or numerically. Since the procedure is completely systematic, it can efficiently be used in a computer-programmed computation.

The stability property of the suboptimal feedback function is discussed in Theorem 2.4, which is useful to construct the time-invariant control. When the system is time-invariant and the control duration is infinite, and if the unperturbed system is completely controllable and completely observable, the task is further reduced to solving an algebraic linear equation, whose solution is again straightforward. Theorem 2.5 gives a sufficient condition that the optimal control is exactly determined as a feedback form. The contents of Lemma 2.5 as well as of Lemma 2.3 could also be used to construct a desired polynomial function, e.g., a desired Liapunov function.

Although the present procedure of power-series expansion resembles, in some respects, that of E. G. Al'brekht [1] and D. L. Lukes [56], the introduction of the ε-parameter is useful to give the recursive formulas in a much neater fashion. Further, it facilitates the performance analysis to clarify the approximation property of the suboptimal policy qualitatively. However the quantitative estimation of the performance degradation is generally difficult. Theorems 2.6 and 2.7 reveal the fact that, if the feedback function is optimal up to the ℓ-th order in ε, the index of performance is minimized up to the $(2\ell + 1)$-th order. This is an extension of the theorem given by P. V. Kokotović and J. B. Cruz, Jr. [42] for linear systems, and is in accord with the theorem of R. A. Werner and J. B. Cruz, Jr. [96] for systems with unknown parameters.

The effectiveness of the method is illustrated by examining typical examples. Although the convergence of the series-form solution is not always assured for large nonlinearity, the skillful truncation of the series provides good approximations. Namely, the method may be useful, even if the nonlinearity is rather large. But, it is difficult in general to select the number of terms of the series which gives reasonable accuracy.

In short, the method developed provides a promising tool for a variety of problems, including linear or nonlinear, deterministic or stochastic systems optimization. Some of these problems will be subjects of the subsequent chapters.

CHAPTER 3

EXTENSIONS TO SOME CONTROL PROBLEMS CONTAINING SMALL PARAMETERS

3.1 Introduction

The method developed in Chapter 2 offers an efficient means for a suboptimal design of a variety of control systems. Among them are a problem of nonlinear systems with the control appearing linearly, a problem of large-scale control systems, a problem of linear control systems with periodic coefficients, a problem of control systems with nonlinear outputs, and a problem with a nonquadratic performance index. The first three problems are fully discussed in Sections 3.2 through 3.4 of this chapter. The last two problems are stated in Section 3.5, and their solutions are summarized in Appendices III and IV.

In Section 3.2 the problem of nonlinear systems with the control appearing linearly is treated, which is a more generalized one of the problem of Chapter 2. The system considered is linear in the control, but contains product terms of the state and control.

Another straightforward application of the method is the synthesis of a suboptimal control for a large-scale system. A special attention is recently paid to a class of large-scale problems by many researchers. One of the typical ideas is to use the concept of multilevel control or system decomposition [77, 97]. If a system is composed of weakly-coupled subsystems, generally containing nonlinearities, the parameter ε could also be associated with the couplings of the subsystems. In Section 3.3 the idea of suboptimal design with use of decomposition technique is discussed for N-coupled systems.

A particular problem in which all the subsystems and the couplings are linear has been discussed by P. V. Kokotović, W. R. Perkins, J. B. Cruz, Jr., and G. D'Ans [43] and the present author [74, 75], independently. In this case the exact calculation of the optimal control is theoretically possible; however, the high dimensionality of the system makes it impractical. It is then better to split a large-scale design problem into a set of simple subsystem problems. The approach which completely neglects the coupling may result in unsatisfactory system performance. Section 3.3 contains a discussion on a compromise between calculational simplicity and performance degradation.

As mentioned in Section 2.4, the construction of the optimal state regulator for a linear system involves the solution of a nonlinear matrix differential equation, Riccati equation, which lies at the heart of the problem. If the system is time-varying, the closed form solution of the Riccati equation is impossible even when the control duration is sufficiently large. For any time-varying problem, D. L. Kleinman and M. Athans [41] have proposed a suboptimal control consisting of an arbitrary prechosen set of linear independent scalar time functions. Section 3.4 deals with an approximate construction of the optimal state regulator for a periodically-varying linear system with quadratic performance index [69, 72, 73]. The aim is to determine a linear feedback control which is relatively easy to implement, yet one which results in suboptimal performance index. The system considered contains a periodic term which is taken to be perturbation to the time-invariant linear-system equation.

Typical examples of dynamical systems optimization are the attitude control and the trajectory optimization of a space vehicle. Dynamical equations describing the attitude or the trajectory of a satellite can be linearized under some reasonable assumptions. However, a more precise model is often required. Then we are faced inevitably with nonlinear and/or large-scale systems, and find the ordinary linear theories helpless. The above two kinds of space problems are adopted as examples attached in each section of this chapter. This is because one of the motivations of the present research is to solve these problems with nonlinearities and/or with high dimensionalities.

3.2 Problem of Nonlinear Systems with Controls Appearing Linearly

3.2.1 Problem Statement

In this section, an extended version of the nonlinear problem of Chapter 2 is treated. The problem of practical interest is that of a control system linear in the control variables, but containing products of state and control variables.

Consider nonlinear control systems governed by the equation*

$$\dot{x} = A(t)x + \varepsilon f(x, t) + [B(t) + \varepsilon D(x, t)]u \tag{3.1}$$

subject to the performance index

$$J = \frac{1}{2}x'(T)Fx(T) + \frac{1}{2}\int_{\tau}^{T}[x'Q(t)x + u'R(t)u]dt \tag{3.2}$$

where the $n \times m$-matrix function $D(x, t)$, continuous in t, is an analytic function in x; all the other quantities are the same as specified in Section 2.2. For simplicity, the output of the system is assumed to be the state itself.

---

* In particular, the system both with $f(x, t) = 0$ and with $D(x, t)$ linear in x is called the bilinear control system [82].

3.2.2 Construction of a Suboptimal Control

In the same manner as in Chapter 2, we can use the minimum principle to obtain the necessary condition for optimality. Also the arguments on the unique existence of solution can be held similarly. The Hamiltonian of the present problem is given by

$$H = \frac{1}{2}x'Qx + \frac{1}{2}u'Ru + p'[Ax + \varepsilon f + (B + \varepsilon D)u] \tag{3.3}$$

where $p(x, t)$ is the n-costate vector satisfying the equation

$$\dot{p} = - Qx - A'p - \varepsilon(\frac{\partial f}{\partial x})'p - \varepsilon\frac{\partial}{\partial x}(p'Du) \tag{3.4}$$

with the boundary condition

$$p(x, T) = Fx(T) \tag{3.5}$$

The optimal control, which has to minimize the Hamiltonian, is given by

$$u^*(x, t) = - R^{-1}(t)[B'(t) + \varepsilon D'(x, t)]p(x, t) \tag{3.6}$$

The problem is again reduced to solving the canonical equations (3.1), (3.4), and (3.5) with (3.6), and the small parameter method of the previous chapter is applicable, with slight modifications. Thus we develop the vector p into a power series:

$$p(x, t) = P(t)x + \sum_{k=1}^{\infty} \varepsilon^k p^{(k)}(x, t) \tag{3.7}$$

By substituting Eq. (3.7) into (3.4) with use of Eq. (3.6) and by equating the coefficients of the like powers of ε separately, we obtain a sequence of equations:

$$\dot{P} = - PA - A'P + PEP - Q \tag{3.8}$$

$$\frac{\partial p^{(k)}}{\partial t} + \frac{\partial p^{(k)}}{\partial x} Sx + S'p^{(k)} = h^{(k)} \qquad (k = 1, 2, \ldots) \tag{3.9}$$

together with the boundary conditions

$$P(T) = F \tag{3.10}$$

$$p^{(k)}(x, T) = 0 \qquad (k = 1, 2, \ldots) \tag{3.11}$$

E and S are the $n \times n$-matrices defined by Eqs. (2.12) and (2.18), respectively, and $h^{(k)}(x, t)$ are the n-vectors defined by

$$\begin{aligned}
h^{(1)}(x, t) &\triangleq -(\frac{\partial f}{\partial x})'Px - Pf + \frac{1}{2}\frac{\partial}{\partial x}(x'PGPx),\\
h^{(2)}(x, t) &\triangleq (\frac{\partial g}{\partial x})'p^{(1)} + \frac{\partial p^{(1)}}{\partial x} g + \frac{\partial p^{(1)}}{\partial x} Ep^{(1)} + \frac{1}{2}\frac{\partial}{\partial x}(x'PHPx),\\
h^{(3)}(x, t) &\triangleq (\frac{\partial g}{\partial x})'p^{(2)} + \frac{\partial p^{(2)}}{\partial x} g + [\frac{\partial}{\partial x}(HPx)]'p^{(1)} + \frac{\partial p^{(1)}}{\partial x} HPx + \frac{\partial p^{(1)}}{\partial x} Ep^{(2)}\\
&\quad + \frac{\partial p^{(2)}}{\partial x} Ep^{(1)} + \frac{1}{2}\frac{\partial}{\partial x}[p^{(1)}{}'Gp^{(1)}] + \left\{\frac{\partial p^{(1)}}{\partial x} - [\frac{\partial p^{(1)}}{\partial x}]'\right\}Gp^{(1)},\\
h^{(k)}(x, t) &\triangleq (\frac{\partial g}{\partial x})'p^{(k-1)} + \frac{\partial p^{(k-1)}}{\partial x} g + [\frac{\partial}{\partial x}(HPx)]'p^{(k-2)} + \frac{\partial p^{(k-2)}}{\partial x} HPx\\
&\quad + \sum_{i=1}^{k-1} \frac{\partial p^{(i)}}{\partial x} Ep^{(k-i)} + \frac{1}{2}\frac{\partial}{\partial x}[\sum_{i=1}^{k-2} p^{(i)}{}'Gp^{(k-1-i)}\\
&\quad + \sum_{i=1}^{k-3} p^{(i)}{}'Hp^{(k-2-i)}] + \sum_{i=1}^{k-2}\left\{\frac{\partial p^{(i)}}{\partial x} - [\frac{\partial p^{(i)}}{\partial x}]'\right\}Gp^{(k-1-i)}\\
&\quad + \sum_{i=1}^{k-3}\left\{\frac{\partial p^{(i)}}{\partial x} - [\frac{\partial p^{(i)}}{\partial x}]'\right\}Hp^{(k-2-i)} \qquad (k = 4, 5, \ldots)
\end{aligned} \tag{3.12}$$

The n-vector g and the $n \times n$-symmetric matrices G and H are defined by

$$
\begin{aligned}
g(x,\ t) &\triangleq G(x,\ t)P(t)x - f(x,\ t), \\
G(x,\ t) &\triangleq D(x,\ t)R^{-1}(t)B'(t) + B(t)R^{-1}(t)D'(x,\ t), \\
H(x,\ t) &\triangleq D(x,\ t)R^{-1}(t)D'(x,\ t)
\end{aligned}
\tag{3.13}
$$

Comparison of Eqs. (3.9) with (2.65) and (3.12) with (2.67) shows the following: Since the additional term $\varepsilon D(x,\ t)u$ exists in Eq. (3.1), there appear the terms containing $D(x,\ t)$ in Eqs. (3.12); while the left side of Eqs. (3.9) is kept unchanged. Then the method of Chapter 2 may still work effectively and the recursive calculation may proceed in much the same way.

We can readily see that the matrix $\partial p^{(k)}/\partial x$ is symmetric, in its turn that Theorem 2.2 of Chapter 2 is again true. In fact, the induction with respect to k enables us to show the integrability in x of $h^{(k)}(x,\ t)$ defined by Eqs. (3.12) and, as a consequence, the symmetry of $\partial p^{(k)}/\partial x$. The analogue of Theorem 2.3 is the following:

<u>Theorem 3.1</u>

If $f(x,\ t)$ and $D(x,\ t)$ are polynomials of the degree q_f and q_D, respectively, in x, the coefficients being continuous in t, $p^{(k)}(x,\ t)$ given by Eq. (3.9) is the polynomial of the degree $k(q - 1) + 1$ in x, whose coefficients are continuous in t, where

$$
q \triangleq \max\{q_f,\ q_D - 1\} \tag{3.14}
$$

When the time-invariant problem with the infinite control duration is under consideration, the following theorem, the version of Theorem 2.4, is useful:

<u>Theorem 3.2</u>

Suppose that Eqs. (3.1) and (3.2) are both time-invariant. If the unperturbed linear system, i.e., the system (3.1) with $\varepsilon = 0$, is completely controllable, then $p^{(k)}(x, t)$, the solution of Eq. (3.9), is uniformly asymptotically stable as $t \to -\infty$, relative to the polynomial $p_0^{(k)}(x)$ satisfying the equation

$$x'S_0'p_0^{(k)} = \theta_0^{(k)} \qquad (k = 1, 2, \ldots) \tag{3.15}$$

where S_0 is defined by Eq. (2.23) and $\theta_0^{(k)}$ are the scalar functions defined by

$$\begin{aligned}
\theta_0^{(1)}(x) &\triangleq x'P_0(\tfrac{1}{2}GP_0x - f), \\
\theta_0^{(2)}(x) &\triangleq g'p_0^{(1)} + \tfrac{1}{2}p_0^{(1)}{}'Ep_0^{(1)} + \tfrac{1}{2}x'P_0HP_0x, \\
\theta_0^{(3)}(x) &\triangleq g'p_0^{(2)} + x'P_0Hp_0^{(1)} + p_0^{(1)}{}'Ep_0^{(2)} + \tfrac{1}{2}p_0^{(1)}{}'Gp_0^{(1)}, \\
\theta_0^{(k)}(x) &\triangleq g'p_0^{(k-1)} + x'P_0Hp_0^{(k-2)} + \frac{1}{2}\sum_{i=1}^{k-1} p_0^{(i)}{}'Ep_0^{(k-i)} \\
&\quad + \frac{1}{2}\sum_{i=1}^{k-2} p_0^{(i)}{}'Gp_0^{(k-1-i)} + \frac{1}{2}\sum_{i=1}^{k-3} p_0^{(i)}{}'Hp_0^{(k-2-i)} \\
&\qquad (k = 4, 5, \ldots)
\end{aligned} \tag{3.16}$$

P_0 is given by Eq. (2.21).

3.2.3 Example

The following equations approximately describe, with consideration of the second-order nonlinearity, a three-axis attitude-control system of a space body whose orbit is circular:*

$$
\begin{aligned}
\dot{x}_1 &= x_2, \\
\dot{x}_2 &= cx_4 + \varepsilon x_4 x_6 + u_1 + \varepsilon x_3 u_3, \\
\dot{x}_3 &= x_4, \\
\dot{x}_4 &= -\, cx_2 - \varepsilon x_2 x_6 + u_2 - \varepsilon x_1 u_3, \\
\dot{x}_5 &= x_6, \\
\dot{x}_6 &= \varepsilon x_2 x_4 + \varepsilon x_3 x_4 + u_3 + \varepsilon x_1 u_2
\end{aligned}
\tag{3.17}
$$

The physical meanings of x_i and u_i are explained in Appendix II. The constant c is associated with orbiting angular velocity. The performance index is taken to be

$$
J = \frac{1}{2}\int_0^\infty \left(\sum_{i=1}^{6} x_i^2 + \sum_{i=1}^{3} u_i^2\right)dt \tag{3.18}
$$

The system (3.17) contains the product terms of x_i and u_j. The linear system with $\varepsilon = 0$ is completely controllable. Then the method of the present section is applicable. Due to the infinite control duration, a suboptimal control is obtained by calculating the polynomials $p_0^{(k)}$ satisfying Eqs. (3.15). When $c = 1$, the suboptimal control of the first order is given by

---

* See Eqs. (II.10) in Appendix II.

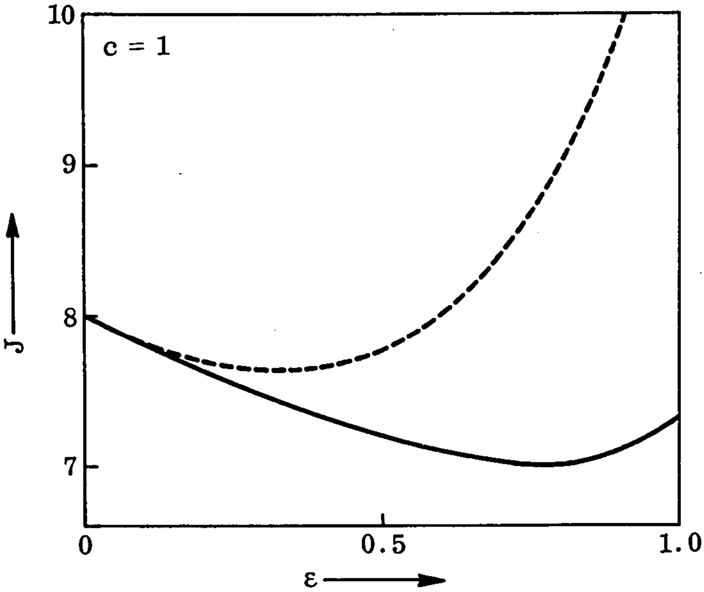
$$u_{si} = u_i^{(0)} + \varepsilon u_i^{(1)} \qquad (i = 1,\ 2,\ 3) \tag{3.19}$$

where

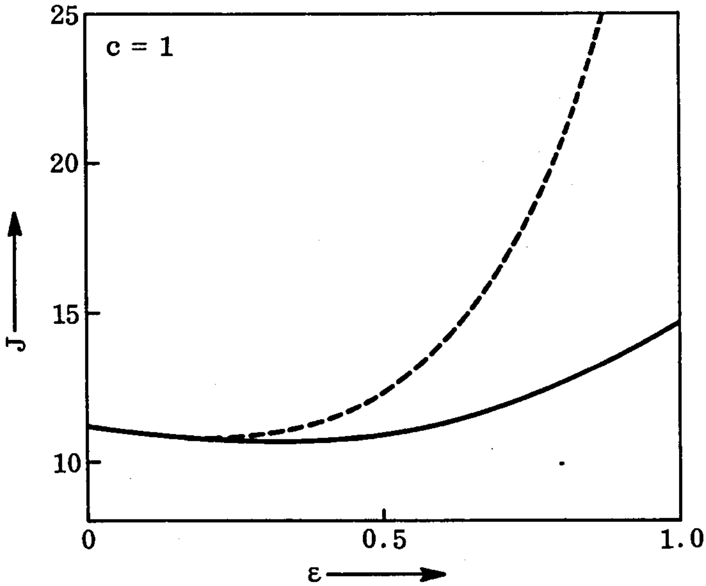
$$\begin{aligned}
u_1^{(0)} &\triangleq -0.8546x_1 - 1.6460x_2 + 0.5192x_3, \\
u_2^{(0)} &\triangleq -0.5192x_1 - 0.8546x_3 - 1.6460x_4, \\
u_3^{(0)} &\triangleq -x_5 - 1.7321x_6, \\
u_1^{(1)} &\triangleq -0.0763x_1x_5 + 0.1246x_1x_6 - 0.0293x_2x_5 + 0.0699x_2x_6 \\
&\quad + 0.5237x_3x_5 + 1.4021x_3x_6 - 0.0631x_4x_5 - 0.0657x_4x_6, \\
u_2^{(1)} &\triangleq 0.1622x_1x_5 - 0.0419x_1x_6 - 0.0631x_2x_5 - 0.0657x_2x_6 \\
&\quad - 0.2023x_3x_5 - 0.0403x_3x_6 - 0.1400x_4x_5 - 0.0700x_4x_6, \\
u_3^{(1)} &\triangleq -0.0191x_1^2 + 0.1246x_1x_2 + 0.3339x_1x_3 - 0.0419x_1x_4 \\
&\quad + 0.0350x_2^2 + 1.4021x_2x_3 - 0.0657x_2x_4 - 0.0744x_3^2 \\
&\quad - 0.0403x_3x_4 - 0.0350x_4^2
\end{aligned} \tag{3.20}$$

Since Eqs. (3.17) are derived with consideration of the second-order nonlinearity, the above suboptimal control with the same order nonlinearity may be reasonable.

Figure 3.1 compares the values of J obtained by the zero- and the first-order controls for various values of ε. The first-order control affords better results even for moderately large ε.



(a) $x_i^0 = \cancel{2}\,1$ $(i = 1, 2, \ldots, 6)$



(b) $x_1^0 = x_3^0 = x_5^0 = \cancel{1}\,2, \quad x_2^0 = x_4^0 = x_6^0 = 0$

——: J by $u_{si} = u_i^{(0)} + \varepsilon u_i^{(1)}$ -----: J by $u_{si} = u_i^{(0)}$

Fig. 3.1. Variation of the performance index J with ε.

3.3 Problem of Large-Scale Control Systems

3.3.1 Problem Statement

Another straightforward application of the method developed in Chapter 2 is synthesis of a suboptimal control for a large-scale system as composed of several subsystems of lower dimensions.

Consider the problem of optimizing the system

$$\dot{x}_i = A_i(t)x_i + \varepsilon f_i(x_1, x_2, \ldots, x_N, t) + B_i(t)u_i \qquad (i = 1, 2, \ldots, N) \tag{3.21}$$

with respect to the performance index

$$\begin{aligned} J &= \sum_{i=1}^{N} J_i, \\ J_i &\triangleq \frac{1}{2}x_i'(T)F_i x_i(T) + \frac{1}{2}\int_{\tau}^{T}[x_i'Q_i(t)x_i + u_i'R_i(t)u_i]dt \end{aligned} \tag{3.22}$$

where, for $i = 1, 2, \ldots, N$,

x_i: n_i-vector of the i-th substate

u_i: m_i-vector of the i-th subcontrol

$f_i(x_1, \ldots, x_N, t)$: n_i-vector function analytic in $x_1, x_2, \ldots, x_N$ and continuous in t satisfying $f_i(0, \ldots, 0, t) = 0$ for any t

$A_i(t)$: $n_i \times n_i$-matrix continuous in t

$B_i(t)$: $n_i \times m_i$-matrix continuous in t

F_i: $n_i \times n_i$-positive semidefinite symmetric matrix

$Q_i(t)$: $n_i \times n_i$-positive semidefinite symmetric matrix continuous in t

$R_i(t)$: $m_i \times m_i$-positive definite symmetric matrix continuous in t

ε: small scalar parameter

N: number of the subsystems

From the outset the performance index is separable in x_i and u_i. Thus, if the parameter ε is naught in Eqs. (3.21), the present problem is separated into the N independent problems each of which contains x_i, u_i, and J_i. These N independent problems are said subproblems, while the original problem is called the total problem. The n_i-vector function f_i is regarded as representing the coupling of the subsystems and ε regarded as the coupling parameter. The smallness of the coupling parameter means that the overall system is made up of weakly interacting subsystems.

Such a kind of large-scale control problems is recently paid special attention to by many researchers. By use of the concept of multilevel or hierarchical control, algorithms are proposed for computing the optimal solution to these problems [77, 97]. However, the solution obtained is mostly of no feedback form. In this section we try to treat the problems as a modification of those of Chapter 2.

3.3.2 Construction of a Suboptimal Control

If we define the n-vector x and m-vector u, respectively, by

$$x \triangleq \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}, \qquad u \triangleq \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_N \end{bmatrix} \tag{3.23}$$

where

$$n \triangleq \sum_{i=1}^{N} n_i, \qquad m \triangleq \sum_{i=1}^{N} m_i \tag{3.24}$$

then our problem is identical with the problem dealt with in Chapter 2. The optimal control is found to be

$$u_i^*(x_1, \ldots, x_N, t) = -R_i^{-1}(t)B_i'(t)[P_i(t)x_i + \sum_{k=1}^{\infty} \varepsilon^k p_i^{(k)}(x_1, \ldots, x_N, t)]$$

$$(i = 1, 2, \ldots, N) \tag{3.25}$$

where P_i and $p_i^{(k)}$ are the $n_i \times n_i$-matrix and the n_i-vector, respectively, satisfying

$$\dot{P}_i = -P_iA_i - A_i'P_i + P_iE_iP_i - Q_i \tag{3.26}$$

$$\frac{\partial p_i^{(k)}}{\partial t} + \frac{\partial}{\partial x_i} \sum_{j=1}^{N} x_j'S_j'p_j^{(k)} = \frac{\partial \theta^{(k)}}{\partial x_i} \qquad (k = 1, 2, \ldots) \tag{3.27}$$

under the boundary conditions

$$P_i(T) = F_i \tag{3.28}$$

$$p_i^{(k)}(0, \ldots, 0, T) = 0 \qquad (k = 1, 2, \ldots) \tag{3.29}$$

E_i, S_i, and $\theta^{(k)}$ are defined by, respectively,

$$E_i(t) \triangleq B_iR_i^{-1}B_i', \qquad S_i(t) \triangleq A_i - E_iP_i \tag{3.30}$$

$$\theta^{(1)}(x_1, \ldots, x_N, t) = -\sum_{j=1}^{N} x_j'P_jf_j,$$

$$\tag{3.31}$$

$$\theta^{(k)}(x_1, \ldots, x_N, t) = \sum_{j=1}^{N} [- f_j' p_j^{(k-1)} + \frac{1}{2} \sum_{\ell=1}^{k-1} p_j^{(\ell)\prime} E_j p_j^{(k-\ell)}]$$

$$(k = 2, 3, \ldots)$$

If the N subsystems are all completely controllable, the solution for the time-invariant problem with $T = \infty$, $F_i = 0$, and $Q_i > 0$ is given by

$$u_i^*(x_1, \ldots, x_N) = - R_i^{-1} B_i' [P_{0i} x_i + \sum_{k=1}^{\infty} \varepsilon^k p_{0i}^{(k)}(x_1, \ldots, x_N)]$$

$$(i = 1, 2, \ldots, N) \qquad (3.32)$$

P_{0i}, positive definite, and $p_{0i}^{(k)}$ are given by, respectively,

$$- P_{0i} A_i - A_i' P_{0i} + P_{0i} E_i P_{0i} - Q_i = 0 \qquad (3.33)$$

$$\sum_{j=1}^{N} x_j' S_{0j}' p_{0j}^{(k)} = \theta_0^{(k)} \qquad (k = 1, 2, \ldots) \qquad (3.34)$$

$$\begin{aligned} S_{0i} &\triangleq A_i - E_i P_{0i}, \\ \theta_0^{(k)}(x_1, \ldots, x_N) &\triangleq \lim_{t \to -\infty} \theta^{(k)}(x_1, \ldots, x_N, t) \end{aligned} \qquad (3.35)$$

Equations (3.26) or (3.33) are separated for each i. This means that the zero-order solution of the total problem is formed of the solutions obtained individually in each subproblem.

3.3.3 Weakly Coupled Linear Systems

A particular problem in which the overall system is linear has been discussed by P. V. Kokotović, W. R. Perkins, J. B. Cruz, Jr., and G. D'Ans [43] and the present author [74, 75], independently. In this case, the exact calculation of the optimal control is theoretically possible. However, the high dimensionality of the system makes it impractical. Intuition and experience may indicate how to split a large-scale design problem into a set of simple subsystem problems. The approach which completely neglects the couplings may result in unsatisfactory system performance. In this subsection, a compromise is sought between calculational simplicity and performance degradation.

Consider a weakly coupled linear system

$$\dot{x}_i = A_i(t)x_i + \varepsilon \sum_{j=1, j\neq i}^{N} D_{ij}(t)x_j + B_i(t)u_i \qquad (i = 1, 2, \ldots, N) \tag{3.36}$$

subject to the index of performance (3.22). All the quantities are specified in Section 3.3.1 with the addition that $D_{ij}(t)$ is the $n_i \times n_j$-matrix continuous in t.

The optimal feedback control for the overall problem is given by Eq. (2.16). Decomposing the feedback gain P of Eq. (2.16) into subsystem gains P_{ii} and coupling gains P_{ij} $(i, j = 1, 2, \ldots, N; i \neq j)$, we may write the optimal control as

$$u_i^*(x_1, \ldots, x_N, t) = -R_i^{-1}(t)B_i'(t) \sum_{j=1}^{N} P_{ij}(t)x_j \qquad (i = 1, 2, \ldots, N) \tag{3.37}$$

where P_{ij} are the $n_i \times n_j$-matrices satisfying the equations

$$\dot{P}_{ii} = -P_{ii}A_i - A_i'P_{ii} - \varepsilon \sum_{\ell=1, \ell\neq i}^{N} (P_{i\ell}D_{\ell i} + D_{\ell i}'P_{\ell i})$$

$$+\sum_{\ell=1}^{N} P_{i\ell}E_{\ell}P_{\ell i} - Q_i,$$

$$\dot{P}_{ij} = -P_{ij}A_j - A_i'P_{ij} - \varepsilon\Big(\sum_{\ell=1,\ell\neq j}^{N} P_{i\ell}D_{\ell j} + \sum_{\ell=1,\ell\neq i}^{N} D_{\ell i}'P_{\ell j}\Big) \tag{3.38}$$

$$+\sum_{\ell=1}^{N} P_{i\ell}E_{\ell}P_{\ell j} \qquad (i \neq j)$$

and the boundary conditions

$$P_{ii}(T) = F_i, \qquad P_{ij}(T) = 0 \quad (i \neq j) \tag{3.39}$$

The matrix equations (3.38) represent a system of n_p-coupled nonlinear scalar equations, where

$$n_p \triangleq \frac{1}{2}n(n + 1) \tag{3.40}$$

with n defined in Eqs. (3.24). When n is large, solving these equations presents extensive difficulties.

The problem considered here is to approximate the optimal feedback gain P_{ij} by a gain which is obtainable by much simpler calculations. To establish such an approximation, note that the calculations become simpler with $\varepsilon = 0$ and that the solutions $P_{ij}(t)$ of Eqs. (3.38) and (3.39) are analytical in ε for all $t \in [\tau, T]$. Therefore it is reasonable to expand P_{ij} in a power series with respect to ε:

$$P_{ij}(t) = \sum_{k=0}^{\infty} \varepsilon^k P_{ij}^{(k)}(t) \qquad (i, j = 1, 2, \ldots, N) \tag{3.41}$$

Substituting Eqs. (3.41) into (3.38) and equating the coefficients of the like powers of ε separately yields a sequence of equations:

$$\dot{P}_{ii}^{(0)} = -P_{ii}^{(0)}A_i - A_i'P_{ii}^{(0)} + P_{ii}^{(0)}E_iP_{ii}^{(0)} - Q_i,$$

$$P_{ij}^{(0)}(t) = 0 \qquad \text{for } t \in [\tau, T] \quad (i \neq j) \tag{3.42}$$

$$\dot{P}_{ij}^{(2k-1)} = -P_{ij}^{(2k-1)}S_j - S_i'P_{ij}^{(2k-1)} + H_{ij}^{(2k-1)} \quad (i \neq j),$$

$$P_{ii}^{(2k-1)}(t) = 0 \qquad \text{for } t \in [\tau, T] \qquad (k = 1, 2, \ldots) \tag{3.43}$$

$$\dot{P}_{ii}^{(2k)} = -P_{ii}^{(2k)}S_i - S_i'P_{ii}^{(2k)} + H_{ii}^{(2k)},$$

$$P_{ij}^{(2k)}(t) = 0 \qquad \text{for } t \in [\tau, T] \quad (i \neq j) \qquad (k = 1, 2, \ldots) \tag{3.44}$$

$$P_{ji}^{(k)} = P_{ij}^{(k)\prime} \qquad (k = 0, 1, 2, \ldots) \tag{3.45}$$

and the boundary conditions

$$P_{ii}^{(0)}(T) = F_i, \qquad P_{ij}^{(k)}(T) = 0 \quad (k = 1, 2, \ldots) \tag{3.46}$$

E_i are defined in Eqs. (3.30); $H_{ij}^{(k)}$ and S_i are the $n_i \times n_j$- and $n_i \times n_i$-matrices, respectively, defined by

$$H_{ij}^{(1)}(t) \triangleq -P_{ii}^{(0)}D_{ij} - D_{ji}'P_{jj}^{(0)},$$

$$H_{ii}^{(2)}(t) \triangleq \sum_{\ell=1, \ell\neq i}^{N} [-P_{i\ell}^{(1)}D_{\ell i} - D_{\ell i}'P_{\ell i}^{(1)} + P_{i\ell}^{(1)}E_\ell P_{\ell i}^{(1)}],$$

$$H_{ij}^{(2k+1)}(t) \triangleq -P_{ii}^{(2k)}D_{ij} - D_{ji}'P_{jj}^{(2k)}$$

$$+ \sum_{r=1}^{k} [P_{ii}^{(2r)} E_i P_{ij}^{(2k-2r+1)} + P_{ij}^{(2k-2r+1)} E_j P_{jj}^{(2r)}], \tag{3.47}$$

$$H_{ii}^{(2k+2)}(t) \triangleq \sum_{\ell=1, \ell \neq i}^{N} [- P_{i\ell}^{(2k+1)} D_{\ell i} - D'_{\ell i} P_{\ell i}^{(2k+1)}$$

$$+ \sum_{r=1}^{k} P_{i\ell}^{(2r-1)} E_\ell P_{\ell i}^{(2k-2r+3)}] + \sum_{r=1}^{k-1} P_{ii}^{(2r)} E_i P_{ii}^{(2k-2r+2)}$$

$$(k = 1, 2, \ldots)$$

$$S_i(t) \triangleq A_i - E_i P_{ii}^{(0)} \tag{3.48}$$

Thus the zero-order terms $P_{ii}^{(0)}$ of the series in Eqs. (3.41) are obtained by solving Eqs. (3.42) separately for each i, since they are not coupled. The first of Eqs. (3.42) represents a system of n_{pi} scalar equations, where

$$n_{pi} \triangleq \frac{1}{2} n_i (n_i + 1) \tag{3.49}$$

Provided that $n_1, n_2, \ldots, n_N$ are all equal, the total number of scalar equations in the first of Eqs. (3.42) is approximately one by N of the number of coupled scalar equations in Eqs. (3.38). This greatly reduces computation. However, if ε is not very small, the performance achieved by the zero-order series in Eqs. (3.41) may be far from optimum. The system may even be unstable.

Addition of higher-order terms will improve performance, while retaining the advantage of decoupled calculations. The procedure involves obtaining $P_{ij}^{(k)}$ successively from Eqs. (3.43) and (3.44) with increasing k. An examination of these equations shows that the homogeneous parts of the equations are the same for all k. Provided that $n_1, n_2, \ldots,$ n_N are equal, each of decoupled equations, Eqs. (3.43), represents a system of approximately $2n_p/N^2$ linear scalar equations and each of Eqs. (3.44) n_p/N^2. It is recalled that n_p is the number of coupled nonlinear scalar equations in Eqs. (3.38).

When the problem is time-invariant, $Q_i > 0$, $F_i = 0$, and $T = \infty$, $P_{ii}^{(0)}$ are the positive definite solutions of the nonlinear algebraic equations, Eqs. (3.42) with $\dot{P}_{ii}^{(0)} = 0$; $P_{ij}^{(k)}$ for $k \geqq 1$ are the solutions of the linear algebraic equations, Eqs. (3.43) and (3.44) with $\dot{P}_{ij}^{(k)} = 0$. If all the subsystems are completely controllable, it is guaranteed that there exist unique positive definite $P_{ii}^{(0)}$ and, due to direct use of Appendix I, exist unique $P_{ij}^{(k)}$ for $k \geqq 1$.

Theorems 2.6 and 2.7, which clarify the approximation property of a suboptimal policy in terms of powers of ε, are valid for the present linear problem. Theorem 2.7 for this problem is a particular case of the theorem obtained by P. V. Kokotović and J. B. Cruz, Jr. [42]. An explicit expression for the performance index resulting from a suboptimal control is also possible as follows. Let $u_{si}(x, t)$ be a suboptimal control such that

$$u_{si}(x, t) = -R_i^{-1}(t)B_i'(t)\sum_{j=1}^{N} P_{sij}(t)x_j \qquad (i = 1, 2, \ldots, N) \qquad (3.50)$$

where P_{sij} are suboptimal feedback gains (3.41) truncated after a certain order in ε. The value of the performance index $J = J_s$, Eqs. (3.22), associated with the control of Eqs. (3.50) may be written as

$$J_s = \frac{1}{2}x'(T)Fx(T) + \frac{1}{2}\int_{\tau}^{T}(x'Qx + u'Ru)dt = \frac{1}{2}x'(\tau)V(\tau)x(\tau) \qquad (3.51)$$

where

$$F \triangleq \begin{bmatrix} F_1 & 0 & \dots & 0 \\ 0 & F_2 & \dots & 0 \\ & \dots\dots & & \\ 0 & 0 & \dots & F_N \end{bmatrix}, \qquad Q \triangleq \begin{bmatrix} Q_1 & 0 & \dots & 0 \\ 0 & Q_2 & \dots & 0 \\ & \dots\dots & & \\ 0 & 0 & \dots & Q_N \end{bmatrix},$$

$$R \triangleq \begin{bmatrix} R_1 & 0 & \dots & 0 \\ 0 & R_2 & \dots & 0 \\ & \dots\dots & & \\ 0 & 0 & \dots & R_N \end{bmatrix} \tag{3.52}$$

and V is an $n \times n$-symmetric positive semidefinite matrix satisfying

$$V(T) = F \tag{3.53}$$

Let us derive equations for the suboptimal performance matrix V. Differentiation of Eq. (3.51) with respect to τ and use of Eqs. (3.36) and (3.50) yields

$$-\frac{1}{2}x'(Q + P_s E P_s)x = \frac{1}{2}x'(\dot{V} + VS_s + S_s'V)x \tag{3.54}$$

where P_s is the overall feedback gain matrix composed of the submatrices P_{sij} and S_s is the matrix defined by

$$S_s \triangleq A - EP_s,$$

$$A \triangleq \begin{bmatrix} A_1 & \varepsilon D_{12} & \dots & \varepsilon D_{1N} \\ \varepsilon D_{21} & A_2 & \dots & \varepsilon D_{2N} \\ & \dots\dots & & \\ \varepsilon D_{N1} & \varepsilon D_{N2} & \dots & A_N \end{bmatrix}, \qquad E \triangleq \begin{bmatrix} E_1 & 0 & \dots & 0 \\ 0 & E_2 & \dots & 0 \\ & \dots\dots & & \\ 0 & 0 & \dots & E_N \end{bmatrix} \tag{3.55}$$

Since Eq. (3.54) holds for arbitrary values of x, we have

$$\dot{V} = - VS_s - S_s' V - Q - P_s E P_s \tag{3.56}$$

Then the suboptimal performance matrix V must satisfy Eq. (3.56) with the boundary condition (3.53).

For the time-invariant problem with $Q_i > 0$, $F_i = 0$, and $T = \infty$, we assume that the system resulting from a suboptimal feedback gain P_s is stable, i.e., S_s is a stable matrix. Then the matrix V is given by the solution of the linear algebraic equation

$$VS_s + S_s' V = - Q - P_s E P_s \tag{3.57}$$

Due to Appendix I, stableness of S_s leads to uniqueness of the solution of Eq. (3.57). If unfortunately S_s is not stable, Eq. (3.57) turns into meaninglessness, because the value of the suboptimal performance grows infinite.

3.3.4 Examples

(a) Attitude Control Problem

A physical problem is examined to illustrate the application of the method above discussed. The following set of equations is a linearized model of the two-axis attitude-control system of an orbiting space body:*

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= \varepsilon x_4 + u_1, \\ \dot{x}_3 &= x_4, \end{aligned} \tag{3.58}$$

---

* See Eqs. (II.9) in Appendix II.

$$\dot{x}_4 = -\varepsilon x_2 + u_2$$

The physical meanings of x_i and u_i are explained in Appendix II. The performance index is assumed to be

$$J = \frac{1}{2}\int_0^{\infty} (x_1^2 + x_2^2 + x_3^2 + x_4^2 + u_1^2 + u_2^2)dt \tag{3.59}$$

The parameter ε is associated with the orbiting angular velocity. When $\varepsilon = 0$, i.e., the satellite is at rest, the system of Eqs. (3.58) is decoupled into two subsystems, each of them being a double integral system and being completely controllable. The method for weakly coupled linear systems can be applied then.

We shall illustrate the calculation for a suboptimal control. For the present problem, the value of various quantities in Eqs. (3.22) and (3.36) are specified by

$$\begin{aligned}
&N = 2, \qquad \tau = 0, \qquad T = \infty, \\
&Q_i = I^{(2)}, \qquad R_i = 1, \qquad F_i = 0, \\
&A_i = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad B_i = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad\qquad (i = 1,\ 2), \\
&D_{12} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \qquad D_{21} = \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix}
\end{aligned} \tag{3.60}$$

Equations (3.42) for the solutions of the decoupled subsystems are reduced to

$$-P_{ii}^{(0)}\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}P_{ii}^{(0)} + P_{ii}^{(0)}\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}P_{ii}^{(0)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad (i = 1,\ 2) \tag{3.61}$$

Equation (3.61) consists of three nonlinear scalar equations. The positive definite solution of the equation is

$$P_{11}^{(0)} = P_{22}^{(0)} = \begin{bmatrix} \sqrt{3} & 1 \\ 1 & \sqrt{3} \end{bmatrix} \tag{3.62}$$

The equations to determine $P_{ij}^{(1)}$ are Eqs. (3.43) with $k = 1$:

$$P_{12}^{(1)} \begin{bmatrix} 0 & 1 \\ -1 & -\sqrt{3} \end{bmatrix} + \begin{bmatrix} 0 & -1 \\ 1 & -\sqrt{3} \end{bmatrix} P_{12}^{(1)} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \tag{3.63}$$

The solution of Eq. (3.63) is uniquely obtained:

$$P_{12}^{(1)} = P_{21}^{(1)\prime} = \frac{1}{\sqrt{3}} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \tag{3.64}$$

By using $P_{ii}^{(0)}$ and $P_{ij}^{(1)}$, form a suboptimal gain P_{s1}:

$$P_{s1} = \frac{1}{\sqrt{3}} \begin{bmatrix} 3 & \sqrt{3} & 0 & \varepsilon \\ \sqrt{3} & 3 & -\varepsilon & 0 \\ 0 & -\varepsilon & 3 & \sqrt{3} \\ \varepsilon & 0 & \sqrt{3} & 3 \end{bmatrix} \tag{3.65}$$

This is the first-order approximation of the optimal gain P. The zero-order approximation of P, P_{s0}, is given by Eq. (3.65) with $\varepsilon = 0$. P is obtained from 10 coupled nonlinear equations, Eqs. (3.38), as

$$P = \begin{bmatrix} p_1 & p_2 & 0 & p_3 \\ p_2 & p_4 & -p_3 & 0 \\ 0 & -p_3 & p_1 & p_2 \\ p_3 & 0 & p_2 & p_4 \end{bmatrix} \tag{3.66}$$

where

$$p_1 \triangleq p_2 p_4 + \varepsilon p_3, \qquad p_2 \triangleq \frac{1}{2}(p_4^2 - 1), \qquad p_3 \triangleq \sqrt{1 - p_2^2} \tag{3.67}$$

and p_4 is the largest real root of the equation

$$p_4^6 + (\varepsilon^2 - 2)p_4^4 - (2\varepsilon^2 + 3)p_4^2 + \varepsilon^2 = 0 \tag{3.68}$$

Now, using Eq. (3.57), we proceed to examine the performance degradation resulting from suboptimal gains. We can readily see that S_s defined by Eqs. (3.55) is stable for any ε both when $P_s = P_{s0}$ and when $P_s = P_{s1}$. It follows that the value of the suboptimal performance is finite in either case. Thus the solutions $V = V_0$ and $V = V_1$ of Eq. (3.57) corresponding to $P_s = P_{s0}$ and $P_s = P_{s1}$, respectively, are uniquely given by

$$V_i = \begin{bmatrix} v_1^{(i)} & v_2^{(i)} & 0 & v_3^{(i)} \\ v_2^{(i)} & v_4^{(i)} & -v_3^{(i)} & 0 \\ 0 & -v_3^{(i)} & v_1^{(i)} & v_2^{(i)} \\ v_3^{(i)} & 0 & v_2^{(i)} & v_4^{(i)} \end{bmatrix} \qquad (i = 0,\ 1) \tag{3.69}$$

where

$$\begin{aligned} &v_1^{(0)} \triangleq \sqrt{3} + \frac{\varepsilon^2}{\sqrt{3}}, \qquad v_2^{(0)} \triangleq 1, \qquad v_3^{(0)} \triangleq \frac{\varepsilon}{\sqrt{3}}, \qquad v_4^{(0)} \triangleq \sqrt{3}, \\ &v_1^{(1)} \triangleq \frac{1}{\sqrt{3}d}(27 + 9\varepsilon^2 + \varepsilon^4), \qquad v_2^{(1)} \triangleq \frac{1}{2d}(18 + \varepsilon^2), \\ &v_3^{(1)} \triangleq \frac{1}{\sqrt{3}d}(9 + \varepsilon^2)\varepsilon, \qquad v_4^{(1)} \triangleq \frac{3\sqrt{3}}{2d}(6 + \varepsilon^2), \end{aligned} \tag{3.70}$$

$d \triangleq 9 + 2\varepsilon^2$

It is easily seen that the difference ΔV of V_0 and V_1:

$$\Delta V \triangleq V_0 - V_1 \tag{3.71}$$

is positive definite for any $\varepsilon \neq 0$. Therefore the first-order suboptimal gain is always better than the zero-order for any ε: J resulting from the first-order gain is less than that from the zero-order gain for arbitrary nonzero ε.

In Fig. 3.2, the suboptimal performance matrices V_0 and V_1 are compared with the optimal performance matrix P for various values of ε in $0 \leq \varepsilon \leq 2$. Observe that the values of the elements of V_1 remain within 10 % of those of P. The values of performance achieved by using the various gains P_{s0}, P_{s1}, and P can be read out of Fig. 3.2 for arbitrary initial conditions; for example, those for the initial condition $x_1(0) = x_2(0) = 1$ and $x_3(0) = x_4(0) = 0$ are given by the sums of values read out of the figures (a), (b), and (d). Other numerical examples are found in the references [74, 75].

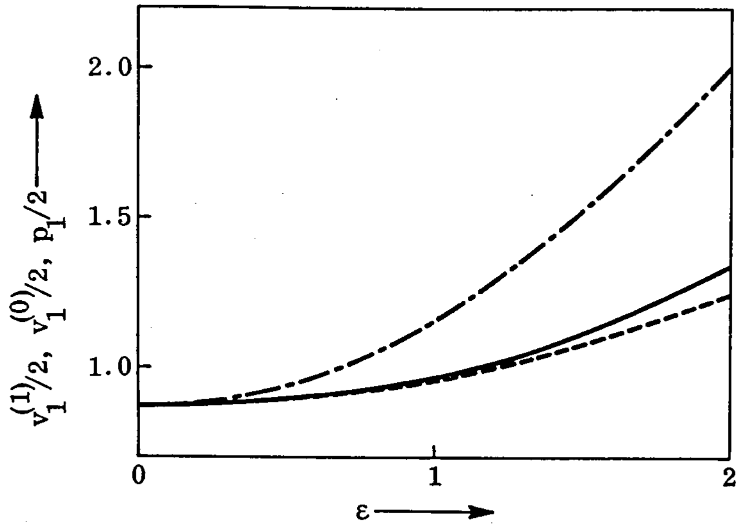
(b) Orbit Transfer Problem

A space problem is adopted again. The motion of a space vehicle which transfers between two neighboring circular orbits is approximately described by the equations containing the second-order nonlinearity:*

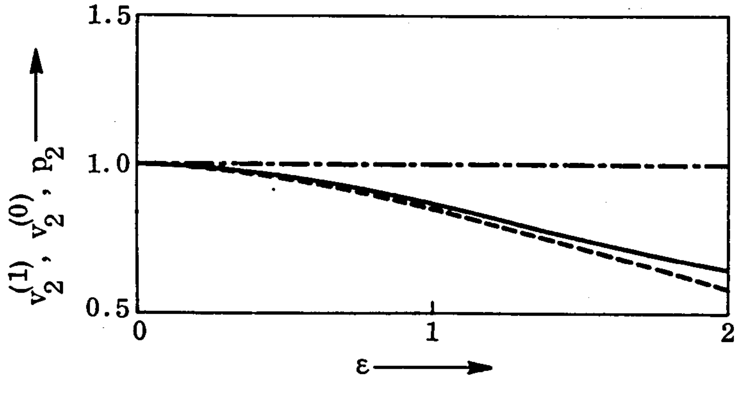
$$\dot{x}_1 = x_2,$$

---

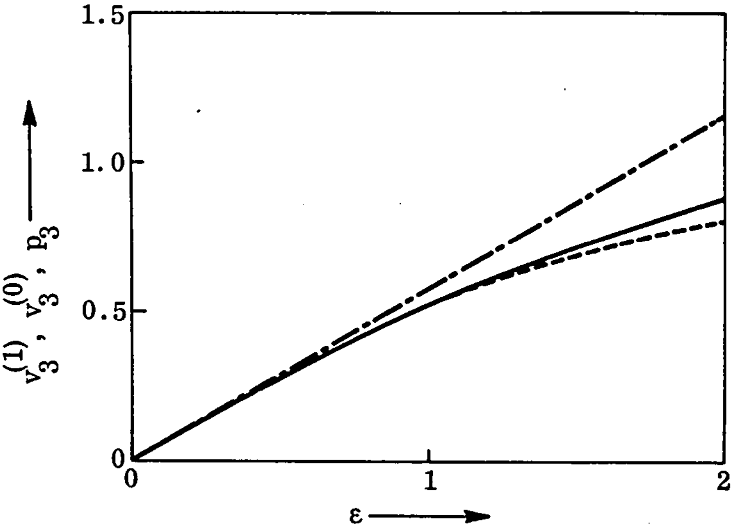
* See Eqs. (II.18) in Appendix II.



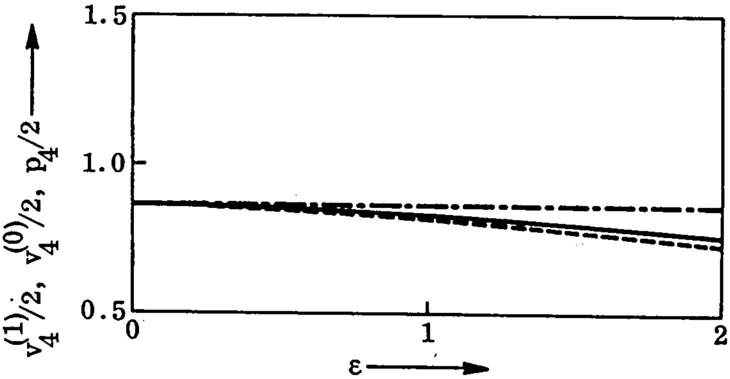
(a) $v_1^{(1)}$, $v_1^{(0)}$, and p_1 vs. ε



(b) $v_2^{(1)}$, $v_2^{(0)}$, and p_2 vs. ε



(c) $v_3^{(1)}$, $v_3^{(0)}$, and p_3 vs. ε



(d) $v_4^{(1)}$, $v_4^{(0)}$, and p_4 vs. ε

———: $v_i^{(1)}$ —·—·—: $v_i^{(0)}$ -----: p_i

Fig. 3.2. Variation of $v_i^{(1)}$, $v_i^{(0)}$, and p_i with ε.

$$
\begin{aligned}
\dot{x}_2 &= -\,2x_4 + 2\varepsilon x_1 x_3 + u_1, \\
\dot{x}_3 &= x_4, \\
\dot{x}_4 &= 2x_2 + 3x_3 + \varepsilon(x_1^2 - 2x_3^2 + x_5^2) + u_2, \\
\dot{x}_5 &= x_6, \\
\dot{x}_6 &= -\,x_5 + 2\varepsilon x_3 x_5 + u_3
\end{aligned}
\tag{3.72}
$$

The physical meanings of x_i and u_i are explained in Appendix II. The performance index is taken to be quadratic in accordance with consideration of the minimum fuel transfer of a low-thrust power-limited rocket:

$$
J = \frac{1}{2}\int_0^{\infty} \left(\sum_{i=1}^{6} x_i^2 + \sum_{i=1}^{3} u_i^2\right) dt \tag{3.73}
$$

The system (3.72) with $\varepsilon = 0$ is decoupled into two linear subsystems: the systems in x_1 to x_4 and in x_5 and x_6. Hence the method of Section 3.3.2 for large-scale systems with nonlinear couplings is applicable to Eqs. (3.72) with (3.73). Due to the infinite control duration, determination of a suboptimal control is reduced to computation of the constant matrices P_{0i} and the vector functions $p_{0i}^{(k)}$ given by Eqs. (3.33) and (3.34), respectively. Finally the suboptimal control of the first-order is as follows:

$$
u_{si} = u_i^{(0)} + \varepsilon u_i^{(1)} \qquad (i = 1,\ 2,\ 3) \tag{3.74}
$$

where

$$
u_1^{(0)} \triangleq -\,0.3200x_1 \quad -\,1.9697x_2 \quad -\,2.9960x_3 \quad -\,0.6732x_4,
$$

$$u_2^{(0)} \triangleq \ 0.9474x_1 \quad - 0.6732x_2 \quad - 4.0119x_3 \quad - 2.4244x_4,$$

$$u_3^{(0)} \triangleq - 0.4142x_5 \quad - 1.3522x_6,$$

$$\begin{aligned} u_1^{(1)} \triangleq & - 0.5631x_1^2 - 0.7243x_1x_2 - 2.2322x_1x_3 - 0.3097x_1x_4 \\ & - 0.1363x_2^2 - 0.1511x_2x_3 + 0.1413x_2x_4 + 1.6626x_3^2 \\ & + 1.2042x_3x_4 + 0.1773x_4^2 - 0.4285x_5^2 - 0.3314x_5x_6 \\ & - 0.1084x_6^2, \end{aligned} \tag{3.75}$$

$$\begin{aligned} u_2^{(1)} \triangleq & - 1.1902x_1^2 - 0.3097x_1x_2 + 1.1041x_1x_3 + 0.7724x_1x_4 \\ & + 0.0706x_2^2 + 1.2042x_2x_3 + 0.3546x_2x_4 + 3.0067x_3^2 \\ & + 0.9346x_3x_4 - 0.0028x_4^2 - 0.8079x_5^2 - 0.5849x_5x_6 \\ & - 0.1619x_6^2, \end{aligned}$$

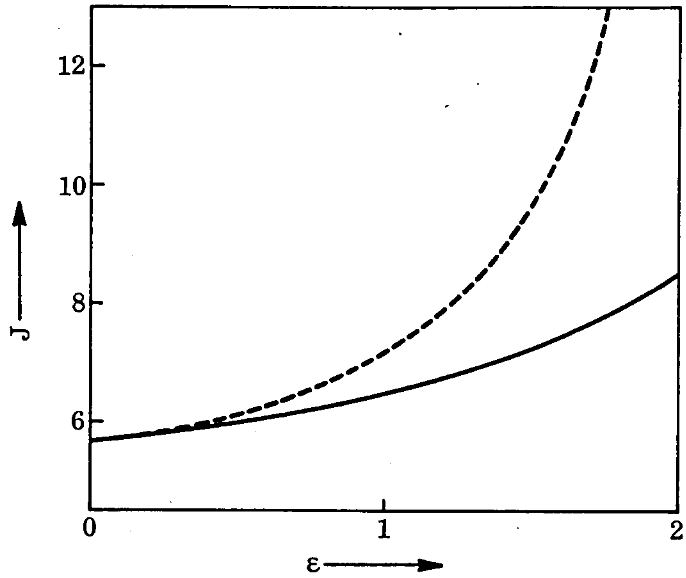
$$\begin{aligned} u_3^{(1)} \triangleq & \ 0.2257x_1x_5 + 0.0791x_1x_6 - 0.3314x_2x_5 - 0.2168x_2x_6 \\ & - 1.9368x_3x_5 - 1.0710x_3x_6 - 0.5849x_4x_5 - 0.3239x_4x_6 \end{aligned}$$

Figure 3.3 illustrates examples of variation of J with ε obtained by the zero- and first-order controls.

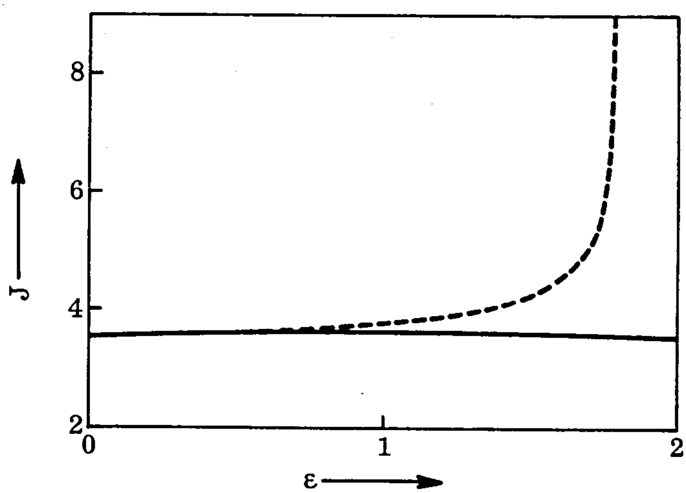
3.4 Problem of Linear Control Systems with Periodic Coefficients

3.4.1 Problem Statement

It has been mentioned in Section 2.4 that the optimal solution $u = u^*$ for the linear regulator problem is generated by the linear feedback control of the form (2.16). The elements of the feedback gain ma-



(a) $x_1^0 = x_3^0 = x_5^0 = 1, \quad x_2^0 = x_4^0 = x_6^0 = 0$



(b) $x_1^0 = x_3^0 = x_5^0 = 0, \quad x_2^0 = x_4^0 = x_6^0 = 1$

——: J by $u_{si} = u_i^{(0)} + \varepsilon u_i^{(1)}$ -----: J by $u_{si} = u_i^{(0)}$

Fig. 3.3. Variation of the performance index J with ε.

trix $P(t)$ are the solution of the matrix Riccati equation (2.14), which lies at the heart of the problem. If the system is time-varying, the closed form solution of $P(t)$ is impossible even when the control duration is sufficiently large. In spite of apparent mathematical simplicity of formulation, there may be a certain engineering shortcoming associated with the implementation of the time-varying feedback gains.

This section deals with an approximate construction of the optimal state regulator for a linear system with periodic coefficients. The goal is to determine a linear feedback control which is relatively easy to implement, with little performance degradation.

Consider dynamical systems governed by the linear time-varying equation

$$\dot{x} = [A^{(0)} + \varepsilon A^{(1)}(t)]x + [B^{(0)} + \varepsilon B^{(1)}(t)]u \tag{3.76}$$

where

x: n-state vector

u: m-control vector

$A^{(0)}$: n x n-constant matrix

$B^{(0)}$: n x m-constant matrix

$A^{(1)}(t)$: n x n-matrix continuous and periodic in t with period L

$B^{(1)}(t)$: n x m-matrix continuous and periodic in t with period L

ε: small scalar parameter

In the present discussion, the periodic terms of the system are taken to be perturbations to the time-invariant system equation.

The problem is to find a feedback control $u(x, t)$ for which the quadratic index of performance

$$J = \frac{1}{2} \lim_{T\to\infty} \int_{\tau}^{T} (x'Qx + u'Ru)dt \tag{3.77}$$

is minimized, where

Q: n × n-positive definite constant matrix

R: m × m-positive definite constant matrix

It is assumed throughout this section that the unperturbed time-invariant system, i.e., the system (3.76) with $\varepsilon = 0$, is completely controllable.

3.4.2 Construction of a Suboptimal Control and Related Characteristics

Due to the assumption of complete controllability, the optimal feedback control for the unperturbed time-invariant system is given by Eq. (2.20), that is,

$$u^*(x) = - R^{-1}B^{(0)\prime}P^{(0)}x \tag{3.78}$$

where $P^{(0)}$ is the n × n-symmetric positive definite matrix uniquely determined by the quadratic algebraic equation

$$- P^{(0)}A^{(0)} - A^{(0)\prime}P^{(0)} + P^{(0)}EP^{(0)} - Q = 0 \tag{3.79}$$

with

$$E \triangleq B^{(0)}R^{-1}B^{(0)\prime} \tag{3.80}$$

Similarly, the optimal control $u^*(x, t)$ for the periodic system of Eq. (3.76) is given by Eq. (2.16). By virtue of the nomenclature in this section, we write

$$u^*(x, t) = - R^{-1}[B^{(0)} + \varepsilon B^{(1)}(t)]'P(t)x \tag{3.81}$$

where $P(t)$ is the solution of the matrix Riccati equation

$$\dot{P} = - P[A^{(0)} + \varepsilon A^{(1)}] - [A^{(0)} + \varepsilon A^{(1)}]'P$$
$$+ P[B^{(0)} + \varepsilon B^{(1)}]R^{-1}[B^{(0)} + \varepsilon B^{(1)}]'P - Q \tag{3.82}$$

with the boundary condition

$$\lim_{T \to \infty} P(T) = 0 \tag{3.83}$$

Since the closed form solution of Eq. (3.82) is generally impossible to obtain, we have to rely upon a numerical calculation. On the other hand, we might expect that, under appropriate conditions, the steady-state solution of Eq. (3.82) as $t \to -\infty$ be periodic and use of the steady-state solution throughout the control be legitimate because of the infinite control duration. Our primal aim is then to have approximately the periodic solution in analytic form.

For a sufficiently small value of ε, it is reasonable to assume that the optimal feedback gain $P(t)$ given by Eq. (3.82) be analytic in ε for all $t \in [\tau, T)$. Therefore we try to find a suboptimal gain in a power-series form in ε. We develop the steady-state solution $P(t)$ of Eq. (3.82) into a power series:

$$P(t) = P^{(0)} + \sum_{k=1}^{\infty} \varepsilon^k P^{(k)}(t) \tag{3.84}$$

where $P^{(k)}(t)$ $(k = 1, 2, \ldots)$, symmetric, are $n \times n$-matrices to be determined. Substituting Eq. (3.84) into (3.82) and equating the coefficients of the like powers of ε separately yields a sequence of equations:

$$\dot{P}^{(k)} = - P^{(k)}S - S'P^{(k)} + H^{(k)}(t) \qquad (k = 1, 2, \ldots) \tag{3.85}$$

with the boundary conditions

$$\lim_{T\to\infty} P^{(k)}(T) = 0 \tag{3.86}$$

$H^{(k)}$ and S are the $n \times n$-matrices defined by

$$\begin{aligned}
H^{(1)}(t) &\triangleq - P^{(0)}A^{(1)}(t) - A^{(1)\prime}(t)P^{(0)} \\
&\quad + P^{(0)}[B^{(0)}R^{-1}B^{(1)\prime}(t) + B^{(1)}(t)R^{-1}B^{(0)\prime}]P^{(0)}, \\
H^{(k)}(t) &\triangleq - P^{(k-1)}(t)A^{(1)}(t) - A^{(1)\prime}(t)P^{(k-1)}(t) + \sum_{i=1}^{k-1} P^{(i)}(t)EP^{(k-i)}(t) \\
&\quad + \sum_{i=0}^{k-1} P^{(i)}(t)[B^{(0)}R^{-1}B^{(1)\prime}(t) + B^{(1)}(t)R^{-1}B^{(0)\prime}]P^{(k-i-1)}(t) \\
&\quad + \sum_{i=0}^{k-2} P^{(i)}(t)B^{(1)}(t)R^{-1}B^{(1)\prime}(t)P^{(k-i-2)}(t) \\
&\qquad (k = 2, 3, \ldots)
\end{aligned} \tag{3.87}$$

$$S \triangleq A^{(0)} - EP^{(0)} \tag{3.88}$$

The zero-order equation for $P^{(0)}$ is identical with Eq. (3.79). Successive solutions of Eqs. (3.85) with increasing k may determine the functions $P^{(k)}(t)$. Since Eqs. (3.85) are linear in $P^{(k)}$, their steady-state solutions are uniquely determined under appropriate conditions. The following theorem is established:

Theorem 3.3

If the unperturbed time-invariant system, i.e., the system (3.76) with $\varepsilon = 0$, is completely controllable, the steady-state solution, as $t \to -\infty$, of Eq. (3.85) is uniquely determined, and is continuous and periodic with period L.

Proof

The theorem can readily be proved inductively. From complete controllability of the unperturbed system, all the eigenvalues of the matrix S of Eq. (3.88) have negative real parts, as shown in Section 2.4. Therefore the steady-state solution $P^{(k)}(t)$, as $t \to -\infty$, of Eq. (3.85) for any $k \geq 1$ uniquely exists and is formally written as

$$P^{(k)}(t) = -e^{-S't} \lim_{T\to\infty} [\int_t^T e^{S'\xi} H^{(k)}(\xi) e^{S\xi} d\xi] e^{-St} \tag{3.89}$$

First, since $H^{(1)}$ of Eqs. (3.87) is continuous and periodic with period L, the evaluation of the quantity $P^{(1)}(t + L)$ shows that $P^{(1)}(t)$ is continuous and periodic with period L.* Second, assume that the steady-state solutions $P^{(k)}(t)$ with $k = 1, 2, \ldots, \ell$ are all continuous and periodic. Then it is readily observed that $H^{(\ell+1)}(t)$ is continuous and periodic and consequently the steady-state solution $P^{(\ell+1)}(t)$ is continuous and periodic.

Q.E.D.

By using these solutions $P^{(k)}(t)$, we may write the optimal feedback control $u^*(x, t)$ as

---

* We may write

$$P^{(1)}(t + L) = -e^{-S't} \lim_{T\to\infty} [\int_{t+L}^T e^{S'(\xi-L)} H^{(1)}(\xi) e^{S(\xi-L)} d\xi] e^{-St}$$

An appropriate variable transformation in the integral yields

$$P^{(1)}(t + L) = -e^{-S't} \lim_{T\to\infty} [\int_t^T e^{S'\xi} H^{(1)}(\xi + L) e^{S\xi} d\xi] e^{-St}$$

Since $H^{(1)}$ is periodic with period L, we have $P^{(1)}(t + L) = P^{(1)}(t)$.

$$u^*(x, t) = -K^{(0)}x - \sum_{k=1}^{\infty} \varepsilon^k K^{(k)}(t)x \tag{3.90}$$

where $K^{(0)}$, constant, and $K^{(k)}(t)$ $(k = 1, 2, \ldots)$, periodic in t with period L, are given by, respectively,

$$\begin{aligned} K^{(0)} &\triangleq R^{-1}B^{(0)\prime}P^{(0)}, \\ K^{(k)}(t) &\triangleq R^{-1}[B^{(0)\prime}P^{(k)}(t) + B^{(1)\prime}(t)P^{(k-1)}(t)] \qquad (k = 1, 2, \ldots) \end{aligned} \tag{3.91}$$

The linear periodic version of Theorem 2.5 in Chapter 2 is readily obtained as follows:

<u>Theorem 3.4</u>

Provided that $B^{(1)} = 0$. If $H^{(1)}$ defined in Eqs. (3.87) satisfies $H^{(1)}(t) = 0$ for any t, the constant feedback gain

$$P = P^{(0)} \tag{3.92}$$

is the exactly optimal one; if $H^{(2)}(t) = 0$ holds for any t, the feedback gain

$$P(t) = P^{(0)} + \varepsilon P^{(1)}(t) \tag{3.93}$$

is exactly optimal.

3.4.3 Systems with Periodic Coefficients of Finite Fourier Series

In this subsection we consider a particular system governed by an equation whose periodic coefficients consist of finite numbers of frequency components, i.e., the system described by

$$\dot{x} = [A^{(0)} + \varepsilon A^{(1)}(t)]x + B^{(0)}u,$$
$$A^{(1)}(t) = \sum_{\ell=1}^{N} (A_{c\ell} \cos 2\ell\omega t + A_{s\ell} \sin 2\ell\omega t) \tag{3.94}$$

where $A_{c\ell}$ and $A_{s\ell}$ are $n \times n$-constant matrices and ω is a positive constant. The index of performance is that given by Eq. (3.77). For this problem, the Riccati equation can be solved by an algebraic procedure as shown below.

Following the procedure of Section 3.4.2, we obtain, for the system (3.94), the optimal feedback control

$$u^*(x, t) = - R^{-1}B^{(0)\prime}[P^{(0)} + \sum_{k=1}^{\infty} \varepsilon^k P^{(k)}(t)]x \tag{3.95}$$

where $P^{(k)}(t)$ are the periodic solutions of Eqs. (3.85) to (3.87) with $B^{(1)} = 0$. The functions $H^{(k)}(t)$ defined by Eqs. (3.87) can now be written as

$$H^{(k)}(t) = \sum_{\ell=0}^{kN} [H_{c\ell}^{(k)} \cos 2\ell\omega t + H_{s\ell}^{(k)} \sin 2\ell\omega t], \qquad (k = 1, 2, \ldots) \tag{3.96}$$
$$H_{s0}^{(k)} = 0$$

where $H_{c\ell}^{(k)}$ and $H_{s\ell}^{(k)}$ are the $n \times n$-constant matrices. As a consequence the periodic solutions are given by

$$P^{(k)}(t) = \sum_{\ell=0}^{kN} [P_{c\ell}^{(k)} \cos 2\ell\omega t + P_{s\ell}^{(k)} \sin 2\ell\omega t] \qquad (k = 1, 2, \ldots) \tag{3.97}$$

where $P_{c\ell}^{(k)}$ and $P_{s\ell}^{(k)}$ are the $n \times n$-constant matrices satisfying the equation

$$P_{c\ell}^{(k)}S + S'P_{c\ell}^{(k)} + 2\ell\omega P_{s\ell}^{(k)} = H_{c\ell}^{(k)},$$

$$P_{s\ell}^{(k)}S + S'P_{s\ell}^{(k)} - 2\ell\omega P_{c\ell}^{(k)} = H_{s\ell}^{(k)}$$

$$(\ell = 0,\ 1,\ 2,\ \ldots,\ kN;\ k = 1,\ 2,\ \ldots) \qquad (3.98)$$

Equations (3.98) result from substituting Eqs. (3.96) and (3.97) into (3.85) and equating the coefficients of the like components of frequency separately.

Thus solving the linear algebraic equations (3.98) with increasing k determines $P^{(k)}(t)$. Due to Appendix I concerning the Kronecker product and sum, for fixed values of k and ℓ, Eqs. (3.98) are rewritten into the vector-matrix form:

$$\begin{bmatrix} S' \oplus S' & 2\ell\omega I^{(n^2)} \\ -2\ell\omega I^{(n^2)} & S' \oplus S' \end{bmatrix} \begin{bmatrix} p_c \\ p_s \end{bmatrix} = \begin{bmatrix} h_c \\ h_s \end{bmatrix} \qquad (3.99)$$

where $I^{(n^2)}$ is the $n^2 \times n^2$-identity matrix; p_c, p_s, h_c, and h_s are the n^2-vectors of which the q-th components are the (i, j)-elements of the matrices $P_{c\ell}^{(k)}$, $P_{s\ell}^{(k)}$, $H_{c\ell}^{(k)}$, and $H_{s\ell}^{(k)}$, respectively, q being $n(i - 1) + j$. Since the constant matrix S is stable, so is the Kronecker sum of S', $S' \oplus S'$. Then it can be readily verified that, for any ℓ and ω, the $2n^2 \times 2n^2$-coefficient matrix in Eq. (3.99) is nonsingular.* Hence Eq. (3.99), namely Eqs. (3.98), are uniquely solvable.

In the above discussion, the matrix $B^{(1)}$ is assumed identically zero, for simplicity. When this is not the case, the procedure in this subsection, however, is also applicable.

3.4.4 Performance Degradation Analysis

Similarly to the statements in Chapter 2, for a sufficiently small value of ε, the infinite power series of Eq. (3.84) is convergent to the optimal solution. However, impossibility of the infinite calculation forces us truncation of the series, yielding a suboptimal control. In this subsection the linear periodic versions of the performance evaluation theorems in Section 2.12 are summarized.

Let $u = u_s(x, t)$ be a suboptimal feedback control of the power-series form:

$$u_s(x, t) = - R^{-1}B^{(0)\prime}P^{(0)}x - \sum_{k=1}^{\infty} \varepsilon^k K_s^{(k)}(t)x \tag{3.100}$$

where $K_s^{(k)}(t)$ are those given by Eqs. (3.91) up to a certain order of k. The value of the performance index $J = J_s$ resulting from the feedback control of Eq. (3.100) would be written as

$$J_s = \frac{1}{2} \lim_{T\to\infty} \int_{\tau}^{T} (x'Qx + u_s' Ru_s)dt = \phi[x(\tau), \tau] \tag{3.101}$$

We develop ϕ into the power series in ε:

---

* When $\ell = 0$, the statement is clearly true. When $\ell > 0$, it is sufficient to show that the determinant

$$\det\begin{bmatrix} M & I^{(k)} \\ -I^{(k)} & M \end{bmatrix} = \det M \cdot \det(M + M^{-1})$$

is nonzero, M being a $k \times k$-stable matrix. In fact, if M is stable, we can readily see that the matrix $M + M^{-1}$ is stable also.

$$\phi[x(\tau), \ \tau] = \frac{1}{2}x'(\tau)P^{(0)}x(\tau) + \sum_{k=1}^{\infty} \varepsilon^k \phi^{(k)}[x(\tau), \ \tau] \tag{3.102}$$

where the matrix $P^{(0)}$ is the positive definite solution of Eq. (3.79). Evidently $\phi^{(k)}(0, \ \tau) = 0$ because $u_s(0, \ t) = 0$. Then the following theorems are the linear periodic versions of Theorems 2.6 and 2.7:

Theorem 3.5

If $u_s(x, \ t)$ is the optimal one, that is, if, for all $k \geq 1$, $K_s^{(k)}(t)$ are those given by Eqs. (3.91), then $J_s = J^*$ is given by

$$J^* = \frac{1}{2}x'(\tau)P^{(0)}x(\tau) + \frac{1}{2}\sum_{k=1}^{\infty} \varepsilon^k x'(\tau)P^{(k)}(\tau)x(\tau) \tag{3.103}$$

where $P^{(k)}$ are the solutions of Eqs. (3.85).

Theorem 3.6

If $u_s(x, \ t)$ is optimal up to the order ℓ in ε and truncated after that, that is, if $K_s^{(k)}$ is given by Eqs. (3.91) for $k = 1, \ 2, \ \ldots,$ ℓ and $K_s^{(k)} = 0$ for every $k \geq \ell + 1$, then J_s is equal to J^* up to the order $2\ell + 1$ in ε, i.e.,

$$\phi^{(k)} = \frac{1}{2}x'(\tau)P^{(k)}(\tau)x(\tau) \qquad (k = 1, \ 2, \ \ldots, \ 2\ell + 1) \tag{3.104}$$

and the $(2\ell + 2)$-th term in the expansion of J_s, $\phi^{(2\ell+2)}$, satisfies

$$\frac{\partial\phi^{(2\ell+2)}}{\partial\tau} + x'S'\frac{\partial\phi^{(2\ell+2)}}{\partial x}$$
$$= \frac{1}{2}x'\Big\{\dot{P}^{(2\ell+2)} + P^{(2\ell+2)}S + S'P^{(2\ell+2)}$$

$$- [P^{(\ell+1)}B^{(0)} + P^{(\ell)}B^{(1)}]R^{-1}[B^{(0)\prime}P^{(\ell+1)} + B^{(1)\prime}P^{(\ell)}]\}x \qquad (3.105)$$

The above theorems are proved, by using induction, in much the same way as in Section 2.12. Theorem 3.6 is a particular case of the theorem obtained by P. V. Kokotović and J. B. Cruz, Jr. [42].

Theorems 3.5 and 3.6 clarify the approximation property of the suboptimal policy in terms of powers of ε. An expression for the performance degradation, on behalf of a suboptimal performance matrix as obtained in Section 3.3.3, is now derived as follows. Let $u_s(x, t)$ be a suboptimal control such that

$$u_s(x, t) = - K_s(t)x \qquad (3.106)$$

The value of the performance index $J = J_s$ associated with the control of Eq. (3.106) is written as

$$J_s = \frac{1}{2} \lim_{T\to\infty} \int_{\tau}^{T} (x'Qx + u_s'Ru_s)dt = \frac{1}{2}x'(\tau)P(\tau)x(\tau) + \frac{1}{2}x'(\tau)D(\tau)x(\tau) \qquad (3.107)$$

where P is the exact solution of Eq. (3.82) and D is an $n \times n$-symmetric positive semidefinite matrix satisfying

$$\lim_{T\to\infty} D(T) = 0 \qquad (3.108)$$

That is to say, the first term on the right side of Eq. (3.107) is the optimal value of J; while the second term represents the degradation of J due to the approximation. Let us derive an equation which the degradation matrix D satisfies. Differentiation of Eq. (3.107) with respect to τ and use of Eqs. (3.76) and (3.106) leads to

$$-\frac{1}{2}x'(Q + K_s'RK_s)x = \frac{1}{2}x'[\dot{P} + \dot{D} + (P + D)S_s + S_s'(P + D)]x \tag{3.109}$$

where S_s is the matrix defined by

$$\begin{aligned} S_s(t) &\triangleq A(t) - B(t)K_s(t), \\ A(t) \triangleq A^{(0)} + \varepsilon A^{(1)}(t), &\qquad B(t) \triangleq B^{(0)} + \varepsilon B^{(1)}(t) \end{aligned} \tag{3.110}$$

Since Eq. (3.109) holds for arbitrary values of x, we have, upon use of Eq. (3.82),

$$\dot{D} = - DS_s - S_s'D - (R^{-1}B'P - K_s)'R(R^{-1}B'P - K_s) \tag{3.111}$$

3.4.5 Examples

(a) Orbit Transfer Problem

A physical problem is examined to illustrate the application of the present method. Suboptimal controls are compared with the optimal one.

The motion of a space vehicle which transfers between two neighboring low-eccentricity elliptic orbits is approximately described by*

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= - 2x_4 + \varepsilon(\cos t\cdot x_1 + 2 \sin t\cdot x_3 - 4 \cos t\cdot x_4) + u_1, \\ \dot{x}_3 &= x_4, \\ \dot{x}_4 &= 2x_2 + 3x_3 + \varepsilon(- 2 \sin t\cdot x_1 + 4 \cos t\cdot x_2 + 10 \cos t\cdot x_3) + u_2 \end{aligned} \tag{3.112}$$

* See Eqs. (II.16) and (II.17) in Appendix II.

and

$$\dot{x}_5 = x_6,$$
$$\dot{x}_6 = -x_5 - 3\varepsilon \cos t \cdot x_5 + u_3 \tag{3.113}$$

The physical meanings of the quantities x_i and u_i are explained in Appendix II. The parameter ε, appearing in the periodic terms, represents the eccentricity of the initial vehicle orbit. That is to say, periodic variation in the orbital velocity along the elliptic orbit produces those periodic terms. When $\varepsilon = 0$, the system is completely controllable. Hence the method of the foregoing subsections is applicable to the present problem. The system of Eqs. (3.112), the in-plane motion, is independent of the system of Eqs. (3.113), the out-of-plane motion. Accordingly we discuss the two motions separately.

First, the system of Eqs. (3.113) is treated and the calculation for a suboptimal control is illustrated. For a more general investigation, we introduce a variable frequency parameter ω to consider the system

$$\dot{z}_1 = z_2,$$
$$\dot{z}_2 = -z_1 - \tilde{\varepsilon} \cos 2\omega t \cdot z_1 + v \tag{3.114}$$

where v is the control variable. Equations (3.114) free of the control are the familiar equations of Mathieu type [31, p. 86]. The performance index is taken to be quadratic:

$$J = \frac{1}{2} \int_{T}^{\infty} (z_1^2 + z_2^2 + v^2) dt \tag{3.115}$$

The zero-order solution of the Riccati equation is calculated from Eq. (3.79):

$$P^{(0)} = \begin{bmatrix} \sqrt{4\sqrt{2} - 2} & \sqrt{2} - 1 \\ \sqrt{2} - 1 & \sqrt{2\sqrt{2} - 1} \end{bmatrix} \tag{3.116}$$

The first-order coefficient in the power series (3.84), $P^{(1)}$, may be written as

$$P^{(1)}(t) = \begin{bmatrix} c_1 & c_3 \\ c_3 & c_2 \end{bmatrix} \cos 2\omega t + \begin{bmatrix} s_1 & s_3 \\ s_3 & s_2 \end{bmatrix} \sin 2\omega t \tag{3.117}$$

where c_i and s_i are constants to be determined. Substituting Eq. (3.117) into (3.85) with $k = 1$, using Eq. (3.116), and equating the coefficients of $\cos 2\omega t$ and $\sin 2\omega t$ in both sides separately yields the six simultaneous linear algebraic equations for c_i and s_i. The solution of the equations is

$$\begin{aligned}
c_1 &= \frac{1}{d}\sqrt{2\sqrt{2} - 1}\,[-\,2\sqrt{2}\omega^4 - 3(2 - \sqrt{2})\omega^2 + 3(2 - \sqrt{2})], \\
c_2 &= \frac{1}{d}\sqrt{2\sqrt{2} - 1}\,[2\omega^4 - (2 + \sqrt{2})\omega^2 - (2 - \sqrt{2})], \\
c_3 &= \frac{1}{d}[-\,\omega^4 - (7 - \sqrt{2})\omega^2 - (5\sqrt{2} - 6)], \\
s_1 &= \frac{1}{d}[4(\sqrt{2} - 1)\omega^4 - (8\sqrt{2} - 9)\omega^2 - (11 - 4\sqrt{2})]\omega, \\
s_2 &= \frac{1}{d}[(4\sqrt{2} - 1)\omega^2 - (4\sqrt{2} - 5)]\omega, \\
s_3 &= \frac{1}{d}\sqrt{2\sqrt{2} - 1}\,[2\omega^4 + 3(\sqrt{2} - 1)\omega^2 - 3(\sqrt{2} - 1)]\omega, \\
d &\triangleq 4\omega^6 - (5 - 2\sqrt{2})\omega^4 + (9 - 2\sqrt{2})\omega^2 + 2(2\sqrt{2} - 1)
\end{aligned} \tag{3.118}$$

Then the first-order approximate solution of the Riccati equation (3.82) is

$$P_s(t) = P^{(0)} + \tilde{\varepsilon}P^{(1)}(t) \tag{3.119}$$

Using the elements of the above matrix P_s, we have the suboptimal control

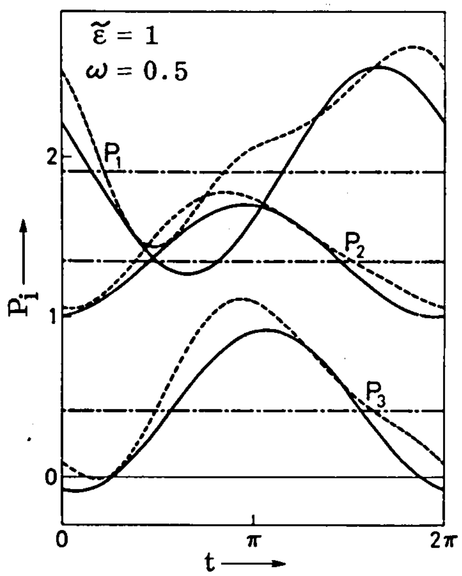
$$v_s = -P_3(t)z_1 - P_2(t)z_2, \qquad \begin{bmatrix} P_1 & P_3 \\ P_3 & P_2 \end{bmatrix} \triangleq P_s \tag{3.120}$$

Figure 3.4 shows, for $\tilde{\varepsilon} = 1$ and $\omega = 0.5$, the approximate solutions of Eq. (3.119) with (3.116) and (3.117) as well as the exact solution. The exact solution is obtained by a numerical integration of Eq. (3.82) using a digital computer. Figures 3.5 and 3.6 show examples of variation of J with the perturbation parameter $\tilde{\varepsilon}$ and with the initial time τ, respectively, for various approximations of the control. These figures are obtained, based upon the discussion from Eq. (3.106) to (3.111). The stabilization property of the controllers is illustrated in Fig. 3.7.* The solid and the chain lines show the stability boundaries of the system with the first-order and the zero-order controls, respectively, on the $(1/\omega, \tilde{\varepsilon})$ plane. The stability boundary of the system without control is also shown dotted. We see that the first-order approximation improves significantly the stabilizability of the control.

Secondly, the system of Eqs. (3.112) is considered. The performance index is assumed to be

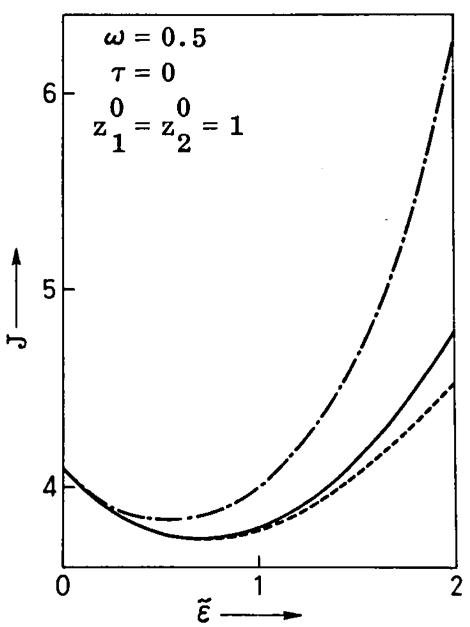
---

* Floquet's theory [31, p. 83] is used for the stability investigation, because the system considered is linear and periodic.



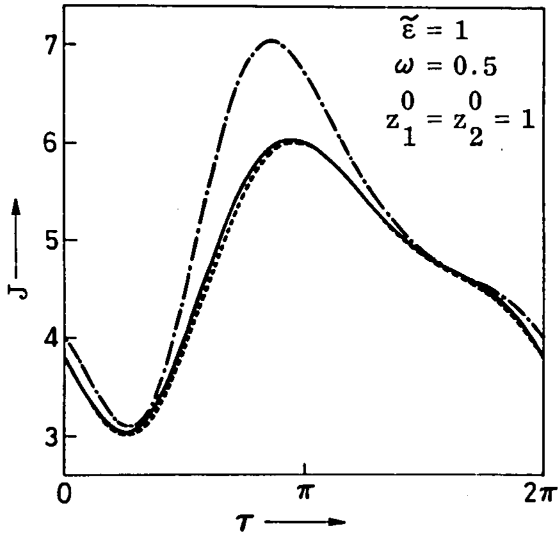
——: $P^{(0)} + \tilde{\varepsilon}P^{(1)}(t)$ —-—: $P^{(0)}$ -----: exact P(t)

Fig. 3.4. Approximate and exact solutions of the Riccati equation for the problem (3.114) and (3.115).



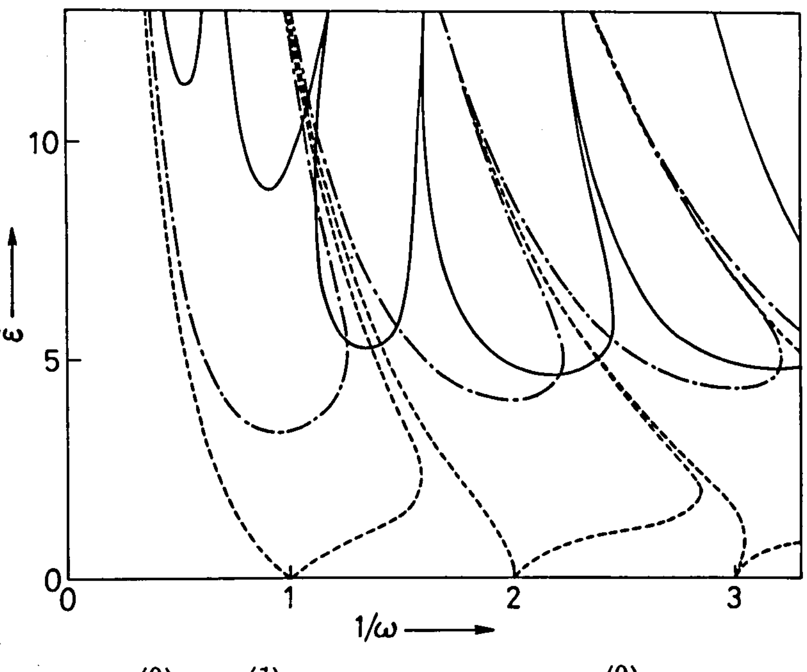
——: J by $P_s = P^{(0)} + \tilde{\varepsilon}P^{(1)}$ —-—: J by $P_s = P^{(0)}$ -----: J^*

Fig. 3.5. Variation of J of Eq. (3.115) with $\tilde{\varepsilon}$.



——: J by $P_s = P^{(0)} + \tilde{\varepsilon}P^{(1)}$ — — —: J by $P_s = P^{(0)}$ - - - - -: J*

Fig. 3.6. Variation of J of Eq. (3.115) with τ.



——: with $P_s = P^{(0)} + \tilde{\varepsilon}P^{(1)}$ — — —: with $P_s = P^{(0)}$ - - - - -: with v = 0

Fig. 3.7. Stability chart for Eqs. (3.114) with the controls of various approximations.

$$J = \frac{1}{2}\int_{\tau}^{\infty}(x_1^2 + x_2^2 + x_3^2 + x_4^2 + u_1^2 + u_2^2)dt \qquad (3.121)$$

After a modest amount of algebraic calculation, we have the approximate solution of the Riccati equation and the associated suboptimal controls:

$$u_{si} = u_i^{(0)} + \varepsilon u_i^{(1)} \qquad (i = 1,\ 2) \qquad (3.122)$$

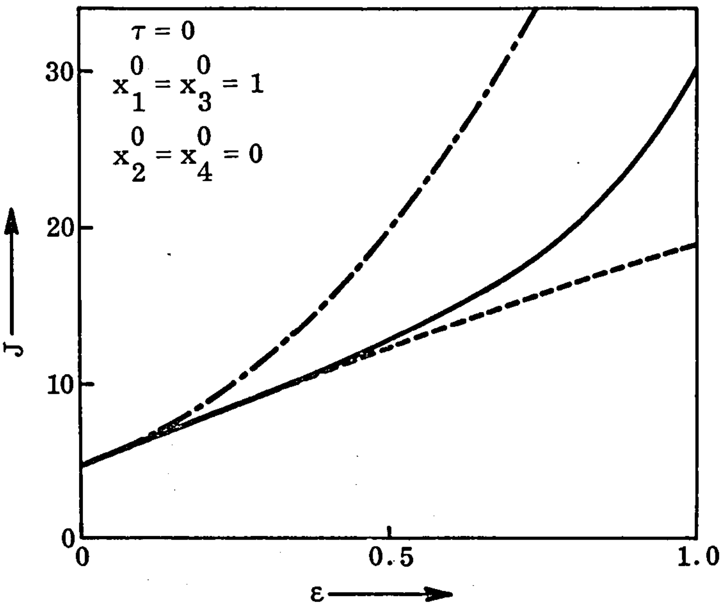
where

$$\begin{aligned}
u_1^{(0)} &\triangleq -0.3200x_1 - 1.9697x_2 - 2.9960x_3 - 0.6732x_4, \\
u_2^{(0)} &\triangleq \ \ 0.9474x_1 - 0.6732x_2 - 4.0119x_3 - 2.4244x_4, \\
u_1^{(1)} &\triangleq (1.3196x_1 - 1.2962x_2 - 9.3930x_3 - 1.3114x_4)\cos t \\
&\quad + (0.5722x_1 + 1.1201x_2 + 3.2442x_3 + 0.7430x_4)\sin t, \\
u_2^{(1)} &\triangleq (1.5454x_1 - 1.3114x_2 - 6.7712x_3 - 0.2291x_4)\cos t \\
&\quad + (1.6571x_1 + 0.7430x_2 + 1.9562x_3 + 0.0349x_4)\sin t
\end{aligned} \qquad (3.123)$$

In Figs. 3.8 and 3.9, examples of variations of J with ε and with τ, respectively, are shown for various approximations of the control. Other numerical examples for this problem are found in the references [69, 72, 73].

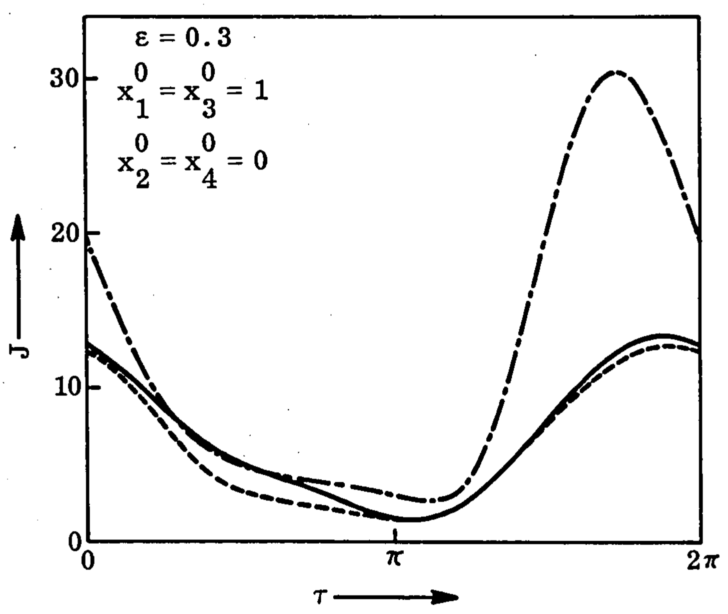
(b) Attitude Control Problem

The following set of equations is a linearized model of the two-axis attitude-control system of a space body whose orbit is a low-eccentricity ellipsis:*



——: J by $u_{si} = u_i^{(0)} + \varepsilon u_i^{(1)}$ —·—·—: J by $u_{si} = u_i^{(0)}$ -----: J*

Fig. 3.8. Variation of J of Eq. (3.121) with ε.



——: J by $u_{si} = u_i^{(0)} + \varepsilon u_i^{(1)}$ —·—·—: J by $u_{si} = u_i^{(0)}$ -----: J*

Fig. 3.9. Variation of J of Eq. (3.121) with τ.

$$
\begin{aligned}
\dot{x}_1 &= x_2, \\
\dot{x}_2 &= cx_4 + \varepsilon \cos ct \cdot x_4 + u_1, \\
\dot{x}_3 &= x_4, \\
\dot{x}_4 &= - cx_2 - \varepsilon \cos ct \cdot x_2 + u_2
\end{aligned}
\tag{3.124}
$$

The physical meanings of x_i and u_i are explained in Appendix II. The parameters c and ε are associated with the mean angular velocity and the eccentricity of the orbit, respectively. The performance index of Eq. (3.121) is used.

When $c = 1$, the suboptimal periodic control of the first order is given by

$$
u_{si} = u_i^{(0)} + \varepsilon u_i^{(1)} \qquad (i = 1, 2) \tag{3.125}
$$

where

$$
\begin{aligned}
u_1^{(0)} &\triangleq - 0.8546x_1 - 1.6460x_2 + 0.5192x_3, \\
u_2^{(0)} &\triangleq - 0.5192x_1 - 0.8546x_3 - 1.6460x_4, \\
u_1^{(1)} &\triangleq (0.3311x_1 + 0.1714x_2 + 0.2785x_3)\cos t \\
&\quad + (- 0.0757x_1 - 0.0981x_2 - 0.1541x_3)\sin t, \\
u_2^{(1)} &\triangleq (- 0.2785x_1 + 0.3311x_3 + 0.1714x_4)\cos t \\
&\quad + (0.1541x_1 - 0.0757x_3 - 0.0981x_4)\sin t
\end{aligned}
\tag{3.126}
$$

---

* See Eqs. (II.7) in Appendix II.

Figures 3.10 and 3.11 show examples of variations of J with ε and with τ, respectively, for various approximations of the control.

3.5 Other Problems of Nonlinear Control Systems

We may remain some of important classes of nonlinear control problems unconsidered; for instance, the problem with a nonlinear output equation, the problem with a nonquadratic performance index, the problem with combination of some problems treated so far, and so on.

In Chapter 2, the output of the system has been assumed linear in state x, as given by Eq. (2.3). However, an output nonlinear in x:

$$y = C(t)x + \varepsilon g(x, t) \tag{3.127}$$

may represent a reasonable mathematical translation of physical models.

In a special problem, it may be desirable to minimize the nonquadratic performance index such as

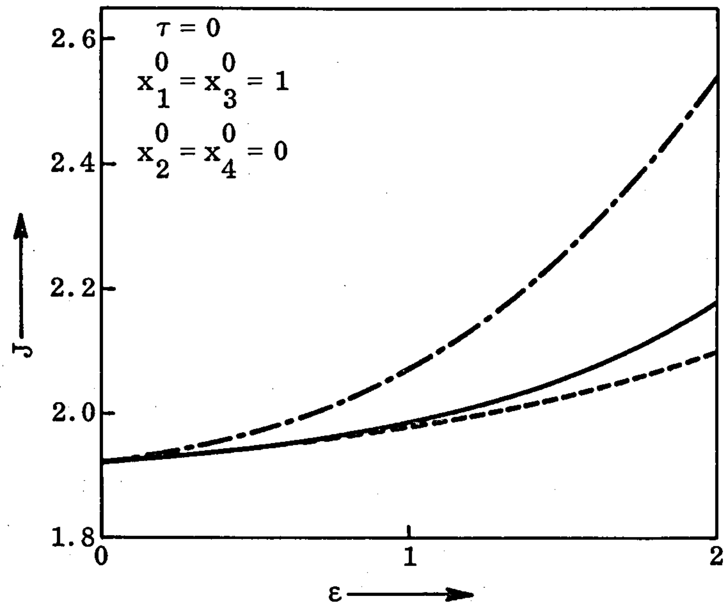
$$J = \frac{1}{2}y'(T)Fy(T) + \varepsilon\phi[y(T)] + \frac{1}{2}\int_{\tau}^{T}[y'Q(t)y + 2\varepsilon\psi(y, t) + u'R(t)u]dt \tag{3.128}$$

rather than the quadratic one, where $\phi(y)$ and $\psi(y, t)$ are scalar functions nonnegative for arbitrary values of their arguments.

The solutions for these two classes of problems are summarized in Appendices III and IV.

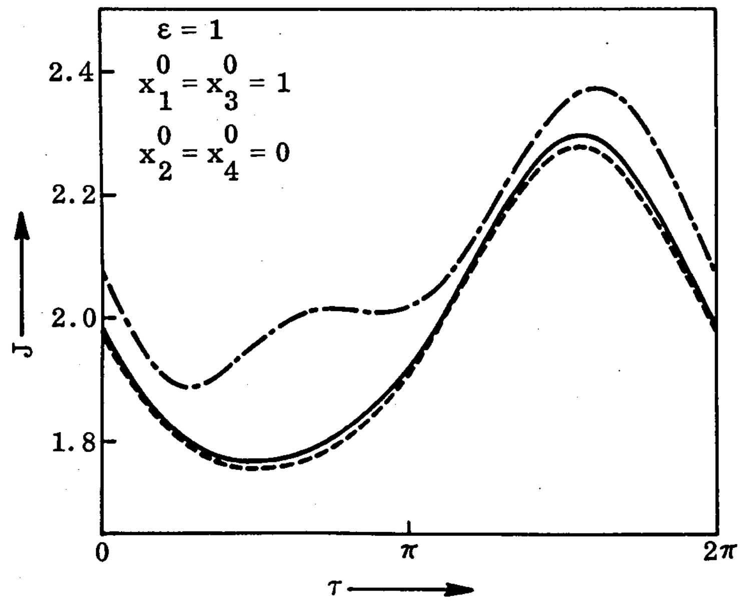
3.6 Concluding Remarks

Several straightforward applications of the method developed in Chapter 2 have been discussed. Section 3.2 has dealt with the problem of systems linear in the control, but containing product terms of the



——: J by $u_{si} = u_i^{(0)} + \varepsilon u_i^{(1)}$ —·—·—: J by $u_{si} = u_i^{(0)}$ - - - - -: J*

Fig. 3.10. Variation of J of Eq. (3.121) with ε.



——: J by $u_{si} = u_i^{(0)} + \varepsilon u_i^{(1)}$ —·—·—: J by $u_{si} = u_i^{(0)}$ - - - - -: J*

Fig. 3.11. Variation of J of Eq. (3.121) with τ.

state and control. Although these product terms are contained in the system equation, the method of Chapter 2 still works effectively. In fact, these terms lead to additional terms in the recurrence equations, but no substantial change occurs in solving the equations. The calculation of a suboptimal control proceeds in much the same way as in Chapter 2.

As shown in Section 3.3, an application to a composite system, whether it is linear or nonlinear, is also straightforward. In a particular problem with linear couplings, the method of Chapter 2 has enlarged the range of applications. The first step of the solution involves the solution of lower-order Riccati equation of each subsystem. In the subsequent steps one has only to solve linear equations. This fact much relaxes the complexity of computation. The various theorems in Chapter 2 apply to the composite problem. Especially, the complete controllability of all the subsystems is sufficient to guarantee the uniqueness of an improved solution.

The method for weakly coupled linear systems is applicable to any problem for which solution of a matrix Riccati type of equation is required; for instance, linear filtering problems with the Riccati equation for a covariance of estimation error [38], linear distributed control problems with the Riccati equation for an operator [29], and so on. An important problem remaining unconsidered for a large-scale system is how to decompose the system into subsystems [87]. A certain decomposition will result in less value of performance than others. Hence, it is significant to choose carefully the suitable scheme of decomposition before proceeding to the present method. Another problem of a large-scale system is the problem in which vanishing of a small parameter yields reducing the system order. This kind of problem has been discussed by P. V. Kokotović, P. Sannuti, and R. A. Yackel [44, 45] and

M. Aoki [2] with use of a singular perturbation technique and an aggregation technique, respectively.

The subject of Section 3.4 has been about an approximate construction of the optimal feedback control for a linear system with periodic coefficients. The procedure is based on a power-series expansion in a small parameter. When the control duration is sufficiently large, the solution of the Riccati equation is reduced to solving a nonlinear algebraic equation and obtaining steady-state solutions of a sequence of linear time-invariant differential equations. It has been shown that, if the unperturbed system is completely controllable, the steady-state solutions for those linear differential equations exist and are periodic. If, in particular, the periodic coefficients in the system equation are given in finite Fourier series, the steady-state solutions are obtained merely by solving linear algebraic equations. An approximation property of the suboptimal policy is clarified and the performance degradation suffered by the approximation is estimated.

For several control problems of a space-vehicle motion, the suboptimal controls of the various orders have been calculated by the present procedure. In the illustrations, the values of the performance index resulting from these controls have been compared for variation of the value of the small parameter ε. The approximate controls afford good results even for moderately large ε, and we have observed the effectiveness of the proposed methods.

CHAPTER 4

SUBOPTIMAL CONTROL OF NONLINEAR NOISY SYSTEMS

4.1 Introduction

Having studied deterministic optimal control in the foregoing chapters, we will now proceed to consideration of stochastic optimal control. In particular, we will consider the synthesis of feedback control which is optimal in the sense of ensemble average under the presence of random disturbances.

We restrict our attention to systems with Gaussian noise. We do this for the reasons that many technically important systems can be adequately represented by models disturbed by Gaussian noise and that the theory for systems with Gaussian noise can readily be applied in practical design problems.

The problem to be considered is that of designing the optimal control for a nonlinear system disturbed by Gaussian white noise, where the performance index is quadratic. A perfect observation of the state of the system is assumed. A feedback control system makes a great feature that, even if the system is designed with neglect of noises, the system well responds to the disturbances. However, provided that statistical characteristics of noises are known in advance, we may expect to obtain a better control by taking the characteristics into account.

J. J. Florentin [24] has originally investigated such a class of problem and obtained the closed form solution to a linear case. Subsequently many researchers have participated in a variety of stochastic optimal control problems. H. J. Kushner [49] has surveyed their works. By similar reasoning to a deterministic case, very few nonlinear noisy problems have been solved analytically so far. In this chapter the small parameter method developed in Chapter 2 is applied to the problem [70]. This is an extension of the work tried by Y. Sawaragi and T. Ono [88] on a special problem of one-dimensional dynamics.

In the next section the problem is stated, and related preliminaries are annexed on stochastic processes. Section 4.3 derives the partial differential equation satisfied by the optimal performance index, and in succession Section 4.4 summarizes the solution to a linear case. For a nonlinear case, due to an additional term associated with the noise, the equation is even harder to solve than that in a deterministic case. Section 4.5 attempts to obtain the solution in a power-series form with respect to a small parameter ε. The procedure is based on the idea for deterministic problems developed in Chapter 2. What differs from a deterministic case is pointed out. The control system with the noise of nonzero mean is discussed in Section 4.6.

Section 4.7 presents an alternative method to obtain a suboptimal feedback control of the same noisy system, where the exact determination of the optimal control is possible for the corresponding deterministic system. It is assumed that the deterministic control is given by the form of finite series in ε as in Theorem 2.5 of Chapter 2, and that the noise variance is adequately small. Thus, the solution is obtained in a power-series form with respect to the variance. In order to make comparison of the solutions by the methods of Sections 4.5 and 4.7, a one-dimensional problem is minutely examined in Section 4.8.

4.2 Problem Statement and Preliminaries

Let the control system be represented by the following nonlinear model:

$$\dot{x} = A(t)x + \varepsilon f(x, t) + B(t)u + C(t)w(t) \tag{4.1}$$

where

x: n-state vector

u: m-control vector

$f(x, t)$: n-vector function analytic in x and continuous in t satisfying $f(0, t) = 0$ for any t

$w(t)$: s-noise vector

$A(t)$: $n \times n$-matrix continuous in t

$B(t)$: $n \times m$-matrix continuous in t

$C(t)$: $n \times s$-matrix continuous in t

ε: small scalar parameter

The random disturbance $w(t)$ is assumed to be of the white Gaussian type. Physically the whiteness may be interpreted to be with very large covariance and very short correlation time relative to characteristic times of the system.

The resemblance of Eq. (4.1) to an ordinary differential equation is misleading, because w is not a function. To eliminate this confusing property, we usually write Eq. (4.1) in the form*

$$dx = A(t)x\,dt + \varepsilon f(x, t)\,dt + B(t)u\,dt + C(t)\,dz \tag{4.2}$$

where z is a Wiener process or classically a Brownian motion process.** Let

$$\delta z(t) \triangleq z(t + \delta t) - z(t) \tag{4.3}$$

---

* A precise interpretation of Eq. (4.2) is given in any standard textbooks on stochastic processes [e.g., 17, p. 273]. Equation (4.2) is also written as an integral equation by use of the so-called stochastic integral.

** The representation dz is, formally, regarded as $w(t)dt$. As for some relations between a white Gaussian noise and a Wiener process, see the reference [36, p. 85].

with an infinitesimal time increment δt. The quantity δz is assumed to have the following statistical moments:

$$\begin{aligned} \mathcal{E}[\delta z(t)] &= 0, \\ \mathcal{E}[\delta z(t)\,\delta z'(t)] &= \sigma\,\delta t \end{aligned} \tag{4.4}$$

where σ, symmetric positive definite, is an $s \times s$-covariance matrix and $\mathcal{E}$ is the expectation operator. Higher moments are of the order higher than δt, under the assumption of Gaussianess of $z(t)$.

Let the performance index for the system (4.2) be the ensemble average of a quadratic form like the one considered in Chapter 2:

$$J(x,\ \tau) = \mathcal{E}\left\{\frac{1}{2}x'(T)Fx(T) + \frac{1}{2}\int_{\tau}^{T}[x'Q(t)x + u'R(t)u]dt\right\} \tag{4.5}$$

where

F: $n \times n$-positive semidefinite matrix

Q(t): $n \times n$-positive semidefinite matrix continuous in t

R(t): $m \times m$-positive definite matrix continuous in t

Clearly, we wish to determine the control u to minimize this average value.

4.3 Necessary Condition for the Optimality

Let $\phi(x,\ t)$ be the optimal value of J starting from the state x at time t:

$$\phi(x,\ t) \triangleq \min_{u} J(x,\ t) \tag{4.6}$$

Assume that ϕ is differentiable with respect to t and twice differentiable with respect to x. Now, with use of the principle of optimality and with the added complication of including second-order terms in the Taylor series expansion, it is straightforward to show that

$$\phi(x,\ t) = \min \mathcal{E}\Big\{\phi(x,\ t) + \frac{\partial\phi}{\partial t}\delta t + (\frac{\partial\phi}{\partial x})'[(Ax + \varepsilon f + Bu)\delta t + C\,\delta z] + \frac{1}{2}\delta z' C' \frac{\partial^2\phi}{\partial x^2} C\,\delta z + \frac{1}{2}(x'Qx + u'Ru)\delta t\Big\} \tag{4.7}$$

where $\mathcal{E}$ represents the expectation over δz. For the above equation to have meaning, the system (4.2) must have the so-called Markovian property. It is known that the various assumptions in the problem statement of Section 4.2 certainly assure the satisfaction of the Markovian property [48].

In Eq. (4.7), upon taking expectations, cancelling the common ϕ, dividing by δt, and letting $\delta t \rightarrow 0$, we have the equation from which the optimal u is obtained. Due to the principle of optimality, the optimal control is given by

$$u^* = -\ R^{-1}B'\frac{\partial\phi}{\partial x} \tag{4.8}$$

We also obtain the equation for ϕ:

$$\frac{\partial\phi}{\partial t} + (Ax + \varepsilon f)'\frac{\partial\phi}{\partial x} - \frac{1}{2}(\frac{\partial\phi}{\partial x})' E \frac{\partial\phi}{\partial x} + \frac{1}{2}\mathrm{tr}(C'\frac{\partial^2\phi}{\partial x^2}C\sigma) + \frac{1}{2}x'Qx = 0 \tag{4.9}$$

where

$$E \triangleq BR^{-1}B' \tag{4.10}$$

and the symbol tr denotes the trace of a matrix. Due to the definition of ϕ by Eqs. (4.6) and (4.5), the boundary condition for Eq. (4.9) is

$$\phi(x, T) = \frac{1}{2}x'(T)Fx(T) \tag{4.11}$$

With $\sigma = 0$, i.e., with no process noise, Eq. (4.9) becomes the Hamilton-Jacobi partial differential equation for the deterministic case [3, p. 357].

4.4 Optimal Control for Linear Systems

This section summarizes the optimal solution to the special case in which Eq. (4.2) is linear. The solution of Eq. (4.9) under (4.11) with $\varepsilon = 0$ is [24]

$$\phi(x, t) = \frac{1}{2}x'P(t)x + \frac{1}{2}\psi(t) \tag{4.12}$$

where P is the $n \times n$-symmetric matrix and ψ is the scalar satisfying the equations, respectively,

$$\begin{aligned} \dot{P} &= - PA - A'P + PEP - Q, & P(T) &= F, \\ \dot{\psi} &= - \mathrm{tr}(C'PC\sigma), & \psi(T) &= 0 \end{aligned} \tag{4.13}$$

With $\sigma = 0$, we obtain $\psi(t) = 0$ for any $t \in [\tau, T]$; then Eq. (4.12) is identical with the optimal performance index of Section 2.4. Since $\psi(t) > 0$ for $t < T$ due to the positive definiteness of P and σ, the presence of the process noise increases the performance index, on the average. Clearly, the optimal control $u = u^*$ is, from Eq. (4.8), identical with the deterministic control of Section 2.4:

$$u^* = - R^{-1}(t)B'(t)P(t)x \tag{4.14}$$

4.5 Solutions of the Recurrence Equations for a Suboptimal Control

For the nonlinear case ($\varepsilon \neq 0$), Eq. (4.9) is even harder to solve than the Hamilton-Jacobi equation, because of the existence of the second-order term. Very few truly nonlinear examples have been solved so far.

We shall attempt to obtain the solution ϕ of Eq. (4.9) under the boundary condition (4.11) in a power-series form with respect to ε:

$$\phi(x, t) = \frac{1}{2}x'P(t)x + \frac{1}{2}\psi(t) + \sum_{k=1}^{\infty} \varepsilon^k \phi^{(k)}(x, t) \tag{4.15}$$

Note that the generating solution is Eq. (4.12). Partial differentiation of Eq. (4.15) with respect to t and x gives

$$\begin{aligned}
\frac{\partial\phi}{\partial t} &= \frac{1}{2}x'\dot{P}x + \frac{1}{2}\dot{\psi} + \sum_{k=1}^{\infty} \varepsilon^k \frac{\partial\phi^{(k)}}{\partial t}, \\
\frac{\partial\phi}{\partial x} &= Px + \sum_{k=1}^{\infty} \varepsilon^k \frac{\partial\phi^{(k)}}{\partial x}, \\
\frac{\partial^2\phi}{\partial x^2} &= P + \sum_{k=1}^{\infty} \varepsilon^k \frac{\partial^2\phi^{(k)}}{\partial x^2}
\end{aligned} \tag{4.16}$$

Substituting Eqs. (4.16) into (4.9) and equating the coefficients of the like powers of ε separately yields a sequence of equations:

$$\frac{\partial\phi^{(k)}}{\partial t} + x'S'\frac{\partial\phi^{(k)}}{\partial x} + \frac{1}{2}\mathrm{tr}[C'\frac{\partial^2\phi^{(k)}}{\partial x^2} C\sigma] = \theta^{(k)} \qquad (k = 1, 2, \ldots) \tag{4.17}$$

with the boundary conditions

$$\phi^{(k)}(x, T) = 0 \qquad (k = 1, 2, \ldots) \tag{4.18}$$

$\theta^{(k)}(x, t)$ are the scalar functions defined by

$$\begin{aligned} \theta^{(1)}(x, t) &\triangleq - x'Pf, \\ \theta^{(k)}(x, t) &\triangleq - f'\frac{\partial\phi^{(k-1)}}{\partial x} + \frac{1}{2}\sum_{i=1}^{k-1} [\frac{\partial\phi^{(i)}}{\partial x}]' E \frac{\partial\phi^{(k-i)}}{\partial x} \qquad (k = 2, 3, \ldots) \end{aligned} \tag{4.19}$$

and S is the matrix defined by

$$S \triangleq A - EP \tag{4.20}$$

The zero-order equations for P and ψ are identical with Eqs. (4.13).

Successive solutions of Eqs. (4.17) with increasing k may determine the functions $\phi^{(k)}(x, t)$. Equations (4.17) differ from the counterpart for the deterministic case, i.e., Eqs. (2.65), through the inclusion of the second derivatives. This fact may not make our circumstances difficult. That is, as in the deterministic problem of Chapter 2, if f is a polynomial in x, $\phi^{(k)}$ are also polynomials in x and their coefficients are exactly determined under appropriate conditions. The following theorem states the details.

Theorem 4.1

If $f(x, t)$ is a homogeneous polynomial of the degree q in x, the coefficients being continuous in t, then $\phi^{(k)}(x, t)$ given by Eq. (4.17) is the polynomial, but not homogeneous, of the degree $k(q - 1) + 2$ in x. When k is even or q is odd, the function $\phi^{(k)}$ contains the terms of the degrees $k(q - 1) + 2$, $k(q - 1)$, $\ldots$, 2, and 0; while, when k is odd and q is even, the terms of the degrees $k(q - 1) +$

2, $k(q - 1)$, ..., 3, and 1 appear.* The coefficients of the polynomial $\phi^{(k)}$ are continuous in t and are given by the equation of the form (2.92).

Define the vector

$$g_{r-1}^{(k)}(x, t) \triangleq \frac{\partial \phi_r^{(k)}}{\partial x} \qquad (r \geq 1) \tag{4.21}$$

where $\phi_r^{(k)}(x, t)$ is the homogeneous term of the degree r collected from the function $\phi^{(k)}(x, t)$. The following theorem, which is the version of Theorem 2.4 for the deterministic problem, shows asymptotic stability of the function $g_r^{(k)}(x, t)$. The theorem is useful for calculating a suboptimal control function of a time-invariant problem with infinite, or sufficiently long, control duration.

Theorem 4.2

If the problem (4.2) and (4.5) is time-invariant with $F = 0$ and $Q > 0$, and if the corresponding unperturbed, noiseless linear problem ($\varepsilon = \sigma = 0$) is completely controllable, then $g_r^{(k)}(x, t)$ for every $k \geq 1$ and $r \geq 0$ is uniformly asymptotically stable as $t \to -\infty$, relative to the polynomial $g_r^{(k)}(x)$ which satisfies the equation

$$x'S'g_r^{(k)} = \theta_{r+1}^{(k)} - \frac{1}{2}\mathrm{tr}[C'\frac{\partial g_{r+2}^{(k)}}{\partial x}C\sigma] \tag{4.22}$$

$\theta_r^{(k)}$ is the collection of the homogeneous terms of the degree r in

---

* Hence, if q is even, $g^{(k)} \triangleq \partial\phi^{(k)}/\partial x$, the k-th order coefficient in the control function, k being odd, contains the term independent of x.

$\theta^{(k)}$ defined by Eqs. (4.19). The polynomial function $g_r^{(k)}(x)$, satisfying Eq. (4.22), exists and is unique for every $k \geqq 1$ and $r \geqq 0$.

The proofs of the above two theorems are omitted here, because they are excuted in ways closely similar to those in Chapter 2. Equation (4.22) shows that, for a fixed value of k, $g_r^{(k)}$ is determined with decreasing r. The zero-degree term $\phi_0^{(k)}(t)$ in $\phi^{(k)}(x, t)$, which is not considered in Theorem 4.2, is uniformly asymptotic as $t \to -\infty$, relative to the quantity

$$\begin{aligned} \phi_0^{(1)}(t) &= \frac{1}{2}\mathrm{tr}[C'\frac{\partial g_1^{(1)}}{\partial x} C\sigma](T - t), \\ \phi_0^{(k)}(t) &= \left\{\frac{1}{2}\mathrm{tr}[C'\frac{\partial g_1^{(k)}}{\partial x} C\sigma] - \theta_0^{(k)}\right\}(T - t) \qquad (k = 2, 3, \ldots) \end{aligned} \tag{4.23}$$

in which the matrix $\partial g_1^{(k)}/\partial x$ is, of course, independent of x. The solution $\phi_0^{(k)}$, however, has no influence on the determination of a suboptimal control, due to Eq. (4.8).

It is an interesting problem whether the performance evaluation theorems, Theorems 2.6 and 2.7 in Chapter 2, hold for the present case or not. We can conclude, after some complicated manipulations, that the theorems are also valid for the present problem. We omit the theorems and the related explanations, however, only because those are cumbersome.

4.6 Systems with Noise of Nonzero Mean

Consider again the problem of optimizing the noisy system (4.2) with respect to the expected performance index (4.5). The noise process $z(t)$ is of the Wiener type like as in the preceding sections, but of the nonzero mean:*

$$\mathcal{E}[\delta z(t)] = \mu\,\delta t, \qquad \operatorname{cov}[\delta z(t)\,\delta z'(t)] = \sigma\,\delta t \tag{4.24}$$

where μ is an s-constant vector and σ is the matrix as defined in Section 4.2. The symbol cov denotes the covariance operator. Since from the first of Eqs. (4.24)

$$\begin{aligned} \operatorname{cov}[\delta z\,\delta z'] &= \mathcal{E}[(\delta z - \mu\,\delta t)(\delta z - \mu\,\delta t)'] \\ &= \mathcal{E}[\delta z\,\delta z'] - \mu\mu'\delta t^2 \end{aligned} \tag{4.25}$$

we have

$$\mathcal{E}[\delta z(t)\,\delta z'(t)] = \sigma\,\delta t \tag{4.26}$$

in the expression up to the first order of δt.

The optimal control $u = u^*$ is again given by Eq. (4.8). Using Eq. (4.26) and the first of Eqs. (4.24), we obtain

$$\frac{\partial \phi}{\partial t} + (Ax + \varepsilon f + C\mu)'\frac{\partial \phi}{\partial x} - \frac{1}{2}\left(\frac{\partial \phi}{\partial x}\right)' E \frac{\partial \phi}{\partial x} + \frac{1}{2}\operatorname{tr}\left(C'\frac{\partial^2 \phi}{\partial x^2}C\sigma\right) + \frac{1}{2}x'Qx = 0 \tag{4.27}$$

The boundary condition for $\phi(x, t)$ is given by Eq. (4.11). Note that the same equation as Eq. (4.27) can formally be derived for the system (4.2) with a constant term $C\mu$ and with a noise of zero mean.

The generating solution of Eq. (4.27) is as follows [24]:

---

* The term "Wiener process with nonzero mean" is not rigorous and should be precisely referred to a temporally homogeneous additive process.

$$\phi(x, t) = \frac{1}{2}x'P(t)x + d'(t)x + \frac{1}{2}\psi(t) \tag{4.28}$$

where d is the n-vector and P and ψ are of the same dimensions as in Section 4.4 satisfying, respectively,

$$\begin{aligned}
\dot{P} &= - PA - A'P + PEP - Q, & P(T) &= F, \\
\dot{d} &= - S'd - PC\mu, & d(T) &= 0, \\
\dot{\psi} &= - tr(C'PC\sigma) + d'Ed - 2d'C\mu, & \psi(T) &= 0
\end{aligned} \tag{4.29}$$

The matrix S is defined by Eq. (4.20).

Accordingly, we try to obtain the solution of Eq. (4.27) in a power-series form:

$$\phi(x, t) = \frac{1}{2}x'P(t)x + d'(t)x + \frac{1}{2}\psi(t) + \sum_{k=1}^{\infty} \varepsilon^k \phi^{(k)}(x, t) \tag{4.30}$$

The functions $\phi^{(k)}(x, t)$ satisfy a sequence of equations:

$$\frac{\partial\phi^{(k)}}{\partial t} + (Sx + C\mu)'\frac{\partial\phi^{(k)}}{\partial x} + \frac{1}{2}tr[C'\frac{\partial^2\phi^{(k)}}{\partial x^2} C\sigma] = \theta^{(k)} \qquad (k = 1, 2, \ldots) \tag{4.31}$$

where the right-hand sides $\theta^{(k)}(x, t)$ have the same form as Eqs. (4.19). The existence of the terms $\mu'C'\partial\phi^{(k)}/\partial x$ leads to modification of Theorem 4.1 with little different contents: If $f(x, t)$ is a homogeneous polynomial of the degree q in x, the function $\phi^{(k)}(x, t)$ is also the polynomial and contains the terms of all the degrees up to $k(q - 1) + 2$ inclusive. For the present case, Theorem 4.2 is still valid with the equation

$$x'S'g_r^{(k)} = \theta_{r+1}^{(k)} - \mu' C' g_{r+1}^{(k)} - \frac{1}{2}\mathrm{tr}[C' \frac{\partial g_{r+2}^{(k)}}{\partial x} C\sigma] \tag{4.32}$$

in place of Eq. (4.22).

4.7 Alternative Method for the Case in Which the Optimal Solution Is Known Exactly for an Associated Deterministic Problem

This section presents an alternative method to obtain a suboptimal feedback control of the stochastic system (4.1) or (4.2) with (4.5) for which the exact determination of the truely optimal nonstochastic control is possible.

We again consider the problem of Section 4.2. Assume that the optimal control $u = u_d^*$ is exactly determined for the corresponding deterministic problem and it is given by

$$\begin{aligned} u_d^* &= - R^{-1}(t)B'(t)p(x, t), \\ p(x, t) &= P(t)x + \sum_{k=1}^{\ell} \varepsilon^k p^{(k)}(x, t) \end{aligned} \tag{4.33}$$

Furthermore, we assume that the noise covariance is adequately small and is written as

$$\sigma = \tilde{\varepsilon}\tilde{\sigma} \tag{4.34}$$

where $\tilde{\varepsilon}$ is a small scalar parameter and $\tilde{\sigma}$ is the $s \times s$-symmetric positive definite matrix.

Due to Theorem 2.2 and Lemma 2.2, there exists a scalar polynomial function $\phi^{(0)}(x, t)$ satisfying

$$\frac{\partial \phi^{(0)}}{\partial x} = p, \qquad \phi^{(0)}(0,\ t) = 0 \qquad \text{for any} \quad t \in [\tau,\ T] \tag{4.35}$$

Therefore, we shall attempt to obtain the solution ϕ of Eq. (4.9) in a power-series form with respect to $\tilde{\varepsilon}$:

$$\phi(x,\ t) = \sum_{k=0}^{\infty} \tilde{\varepsilon}^k \phi^{(k)}(x,\ t) \tag{4.36}$$

having the generating solution $\phi^{(0)}$ defined by Eqs. (4.35). Substituting Eq. (4.36) into (4.9) and equating the coefficients of the like powers of $\tilde{\varepsilon}$ separately yields a sequence of equations:

$$\frac{\partial \phi^{(0)}}{\partial t} + (Ax + \varepsilon f)'p - \frac{1}{2}p'Ep + \frac{1}{2}x'Qx = 0 \tag{4.37}$$

$$\frac{\partial \phi^{(k)}}{\partial t} + (Ax + \varepsilon f - Ep)'\frac{\partial \phi^{(k)}}{\partial x} = \theta^{(k)} \qquad (k = 1,\ 2,\ \ldots) \tag{4.38}$$

with the boundary conditions

$$\phi^{(0)}(x,\ T) = \frac{1}{2}x'Fx \tag{4.39}$$

$$\phi^{(k)}(x,\ T) = 0 \qquad (k = 1,\ 2,\ \ldots) \tag{4.40}$$

$\theta^{(k)}$ are the scalar functions defined by

$$\theta^{(1)} \triangleq -\frac{1}{2}\mathrm{tr}(C'\frac{\partial p}{\partial x}C\tilde{\sigma}),$$

$$\theta^{(k)} \triangleq -\frac{1}{2}\mathrm{tr}[C'\frac{\partial^2 \phi^{(k-1)}}{\partial x^2}C\tilde{\sigma}] + \frac{1}{2}\sum_{i=1}^{k-1}[\frac{\partial \phi^{(i)}}{\partial x}]'E\frac{\partial \phi^{(k-i)}}{\partial x} \tag{4.41}$$

$$(k = 2,\ 3,\ \ldots)$$

Due to the fact that the function p is optimal for the corresponding deterministic problem, Eq. (4.37) with (4.39) is of itself satisfied.

Although successive solutions of Eqs. (4.38) with increasing k may determine the functions $\phi^{(k)}$, the closed form solutions of $\phi^{(k)}$ are in general impossible. In fact, if f is a polynomial of the degree q in x, then, due to Theorem 2.3, p is the polynomial of the degree $\ell(q-1)+1$ in x, and $\theta^{(1)}$ is of the degree $\ell(q-1)$; it follows that, provided $\phi^{(k)}$ is obtained in the polynomial form, $\phi^{(k)}$ should contain terms of infinitely high degree.

Nonetheless, two special cases are solvable: The first is that of the one-dimensional problem in which Eq. (4.38) is analytically solvable in $\partial\phi^{(k)}/\partial x$. This case is minutely discussed in the next section. The second is the case in which the corresponding deterministic problem has the optimal feedback control linear in x. In this case, the function p is given by

$$p = P(t)x \tag{4.42}$$

and $\theta^{(1)}$ defined in Eqs. (4.41):

$$\theta^{(1)} = -\frac{1}{2}\mathrm{tr}(C'PC\tilde{\sigma}) \tag{4.43}$$

is free from x. From Eq. (4.38) with (4.40) for $k = 1$, we have

$$\phi^{(1)} = \frac{1}{2}\int_t^T \mathrm{tr}(C'PC\tilde{\sigma})dt \tag{4.44}$$

Since $\phi^{(1)}$ is independent of x, we know successively

$$\phi^{(k)}(x,\ t) = 0 \qquad \text{for any } x \text{ and } t \in [\tau,\ T] \qquad (k = 2,\ 3,\ \ldots) \tag{4.45}$$

Thus we obtain the following theorem:

Theorem 4.3

If the linear feedback control (2.16) is optimal for the nonlinear deterministic problem, Eqs. (4.1) and (4.5) with $\sigma = 0$, then the same linear control is optimal for the nonlinear noisy problem with $\sigma \neq 0$.

4.8 Example

In order to examine the solutions obtained by the methods of Sections 4.5 and 4.7, let us consider the one-dimensional system governed by the equation

$$dx = g(x)x\,dt + u\,dt + dz \tag{4.46}$$

where $g(x)$ is a scalar function analytic in x. The index of performance is

$$J = \mathcal{E}[\frac{1}{2}\int_0^T (x^2 + u^2)dt] \tag{4.47}$$

Y. Sawaragi and T. Ono [88] have treated a similar example formerly. The control duration T is assumed sufficiently large. Furthermore, we assume that the parameter $\tilde{\varepsilon}$ in Eq. (4.34) is σ itself:

$$\tilde{\varepsilon} = \sigma \tag{4.48}$$

and, as a consequence, $\tilde{\sigma} = 1$.

As shown in the example of Section 2.13(a), the optimal feedback function p for the corresponding deterministic problem is exactly given by

$$p = (g + \sqrt{g^2 + 1})x \tag{4.49}$$

Hence, for the present simple problem, Eqs. (4.38) and (4.41) are reduced to

$$\frac{\partial \phi^{(k)}}{\partial t} - \sqrt{g^2 + 1}\,x \frac{\partial \phi^{(k)}}{\partial x} = \theta^{(k)} \qquad (k = 1,\ 2,\ \ldots) \tag{4.50}$$

$$\begin{aligned} \theta^{(1)} &= -\frac{1}{2}(g + \sqrt{g^2 + 1})(1 + \frac{x}{\sqrt{g^2 + 1}}\frac{dg}{dx}), \\ \theta^{(k)} &= -\frac{1}{2}\frac{\partial^2 \phi^{(k-1)}}{\partial x^2} + \frac{1}{2}\sum_{i=1}^{k-1} \frac{\partial \phi^{(i)}}{\partial x}\frac{\partial \phi^{(k-i)}}{\partial x} \qquad (k = 2,\ 3,\ \ldots) \end{aligned} \tag{4.51}$$

Due to the one-dimensionality of the problem, Eqs. (4.50) may be analytically solvable. We apply the method of separation of variables to Eqs. (4.50). Let us put

$$\phi^{(k)}(x,\ t) = \xi^{(k)}(x) + \eta^{(k)}(t) \tag{4.52}$$

By making use of the quantities $\xi^{(k)}$ and $\eta^{(k)}$, we can rewrite Eqs. (4.50) into

$$\sqrt{g^2 + 1}\,x \frac{d\xi^{(k)}}{dx} + \theta^{(k)} = \dot{\eta}^{(k)} \triangleq \alpha^{(k)} \tag{4.53}$$

where all scalars $\alpha^{(k)}$ have to be constants, and therefore we obtain

$$\frac{\partial \phi^{(k)}}{\partial x} = \frac{d\xi^{(k)}}{dx} = \frac{\alpha^{(k)} - \theta^{(k)}}{\sqrt{g^2 + 1}\,x} \tag{4.54}$$

If $\alpha^{(k)}$ can be found, $\partial\phi^{(k)}/\partial x$ are calculated, and then the optimal control is obtained. The quantity $\alpha^{(k)}$ is given by the constant term in the Maclaurin expansion of $\theta^{(k)}$, because the numerator in the last side of Eq. (4.54) must have a factor x. For, if not, the numerator would not tend to zero as $x \to 0$ and consequently the function $\partial\phi^{(k)}/\partial x$ to infinity, which leads to meaninglessness. Equation (4.54) shows us that the optimal control function is independent of t.

For more explicit computations, we consider the case

$$g(x) = -1 + \varepsilon x^2 \tag{4.55}$$

In accordance with the above-mentioned procedure, we obtain the suboptimal control of the second order in σ:

$$u_\sigma = u_\sigma^{(0)}(x; \varepsilon) + \sigma u_\sigma^{(1)}(x; \varepsilon) + \sigma^2 u_\sigma^{(2)}(x; \varepsilon) \tag{4.56}$$

where

$$\begin{aligned}
u_\sigma^{(0)}(x; \varepsilon) &\triangleq -(g + a)x, \\
u_\sigma^{(1)}(x; \varepsilon) &\triangleq -\frac{\varepsilon}{a} bx, \\
u_\sigma^{(2)}(x; \varepsilon) &\triangleq -\frac{\varepsilon^2}{a^2}\left\{\frac{3}{2a} + \frac{3}{4a + 4\sqrt{2}} - \frac{b^2}{2a} - \frac{g^2}{a^3} + \frac{g}{a}\left[\frac{1 - g}{(\sqrt{2}a + 2)^2} - \frac{b}{a}\right]\right. \\
&\quad \left. + \frac{1 - g}{4}\left[\frac{3}{\sqrt{2}a + 2} - \frac{1}{a} - \frac{a + 2\sqrt{2}}{(\sqrt{2}a + 2)^2}\right]\right\}x,
\end{aligned} \tag{4.57}$$

$$a \triangleq \sqrt{g^2 + 1}, \qquad b \triangleq \frac{3}{2} + \frac{g}{a} + \frac{g - 1}{2a + 2\sqrt{2}}$$

From Eqs. (4.57), when $\varepsilon = 0$, $u_\sigma^{(k)} = 0$ for $k = 1, 2, \ldots$. As stated in Section 4.4, this means that the linear system with noise has the optimal control free from σ.

On the other hand, using the method of Section 4.5, we obtain the alternative form of the suboptimal control in a power series in ε:

$$u_\varepsilon = u_\varepsilon^{(0)}(x) + \varepsilon u_\varepsilon^{(1)}(x;\ \sigma) + \varepsilon^2 u_\varepsilon^{(2)}(x;\ \sigma) \tag{4.58}$$

where

$$\begin{aligned} u_\varepsilon^{(0)}(x) &\triangleq -(\sqrt{2}-1)x, \\ u_\varepsilon^{(1)}(x;\ \sigma) &\triangleq -\frac{\sqrt{2}-1}{4}(2\sqrt{2}x^2 + 3\sigma)x, \\ u_\varepsilon^{(2)}(x;\ \sigma) &\triangleq -\frac{\sqrt{2}}{32}[4x^4 + (12-\sqrt{2})\sigma x^2 + 3(6\sqrt{2}-5)\sigma^2]x \end{aligned} \tag{4.59}$$

We readily see that Eqs. (4.58) and (4.59) agree with Eqs. (4.56) and (4.57) rewritten into series form in ε.

We make comparison among the values of J resulting from the suboptimal controls of eight possible approximations:

$$\begin{aligned} u_{\sigma\ell} &= \sum_{k=0}^{\ell} \sigma^k u_\sigma^{(k)}(x;\ \varepsilon) && (\ell = 0,\ 1,\ 2), \\ u_{\varepsilon\ell} &= \sum_{k=1}^{\ell} \varepsilon^k u_\varepsilon^{(k)}(x;\ \sigma) && (\ell = 1,\ 2), \\ \tilde{u}_{\varepsilon 0} &= u_\varepsilon^{(0)}(x), \\ \tilde{u}_{\varepsilon\ell} &= u_\varepsilon^{(0)}(x) + \sum_{k=1}^{\ell} \varepsilon^k u_\varepsilon^{(k)}(x;\ 0) && (\ell = 1,\ 2) \end{aligned} \tag{4.60}$$

The control $u_{\sigma 0}$ is the optimal one for the corresponding deterministic problem. The controls $u_{\sigma\ell}$ and $u_{\varepsilon\ell}$ $(\ell = 1,\ 2)$ are the suboptimal ones of power-series forms in σ and ε, respectively; while $\tilde{u}_{\varepsilon\ell}$ $(\ell = 0,\ 1,\ 2)$ are the same as u_ℓ given by Eqs. (2.134) and (2.133), i.e., the suboptimal controls obtained in disregard of the noise.

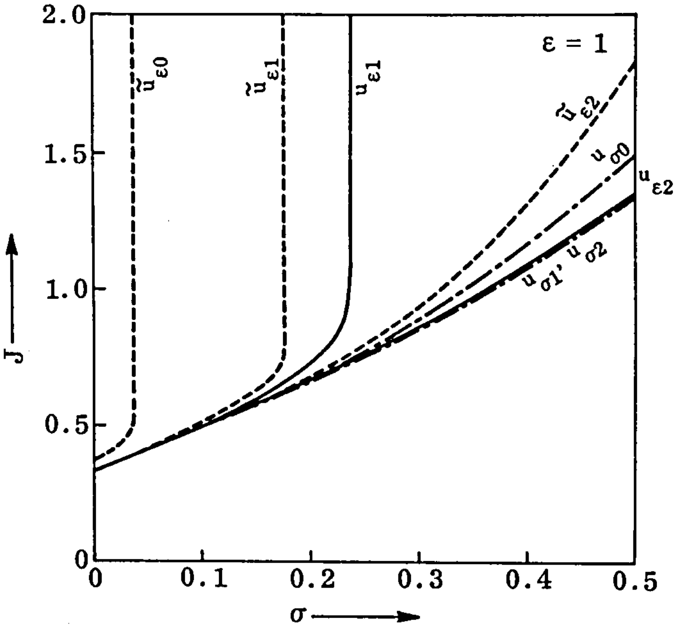
For $T = 5$, Figs. 4.1 and 4.2 illustrate examples of variation of J with σ and with ε, respectively, for the various suboptimal controls of Eqs. (4.60). A computer-generated pseudorandom-number sequence with Gaussian distribution is used in place of the Wiener process noise.* The averages of the results of 30 runs with various random-number sequences are plotted in each figure. We see that the control with consideration of noise is almost better than the same order control without it.

4.9 Concluding Remarks

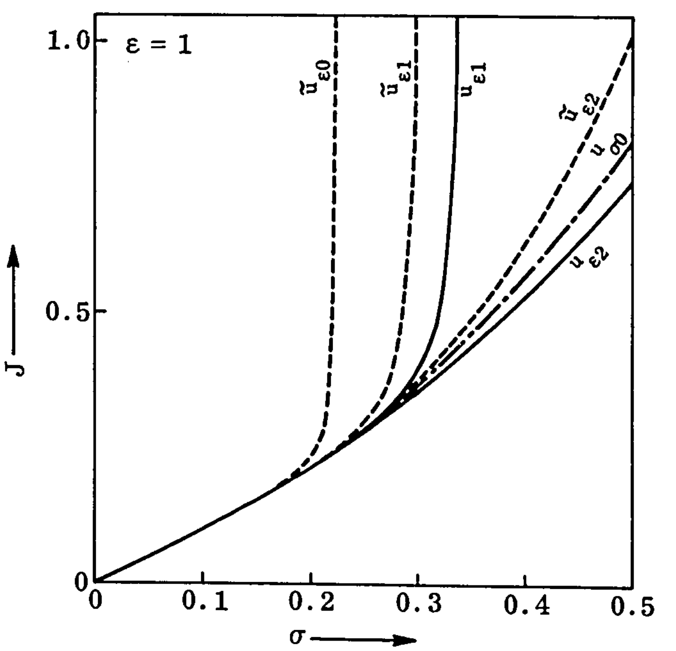
A successive procedure has been developed to construct a suboptimal control for nonlinear noisy systems. Theorems 4.1 and 4.2 are the stochastic versions of Theorems 2.3 and 2.4, respectively, in Chapter 2. Theorem 4.1 shows that, when the nonlinear function in the system equation is a polynomial, the feedback function for the stochastic case contains terms of more degrees in state than for the deterministic case. In a certain case even a term independent of the state is contained in the feedback function. Theorem 4.2 discusses the stability property of the feedback function for the problem with infinite duration time. The infinite control duration is meaningless in both mathematical and physical senses, because it leads to the infinite value of the performance index due to the noise. Hence, Theorem 4.2 should be used for the problem with very large control duration.

Theorem 4.3 reveals the fact that, if the linear feedback control

* A Gaussian pseudorandom-number sequence with specified mean and variance is generated by a digital computer with use of the method of Box-Muller [10].



(a) $x^0 = 1$



(b) $x^0 = 0$

Fig. 4.1. Variation of the performance index J with σ.

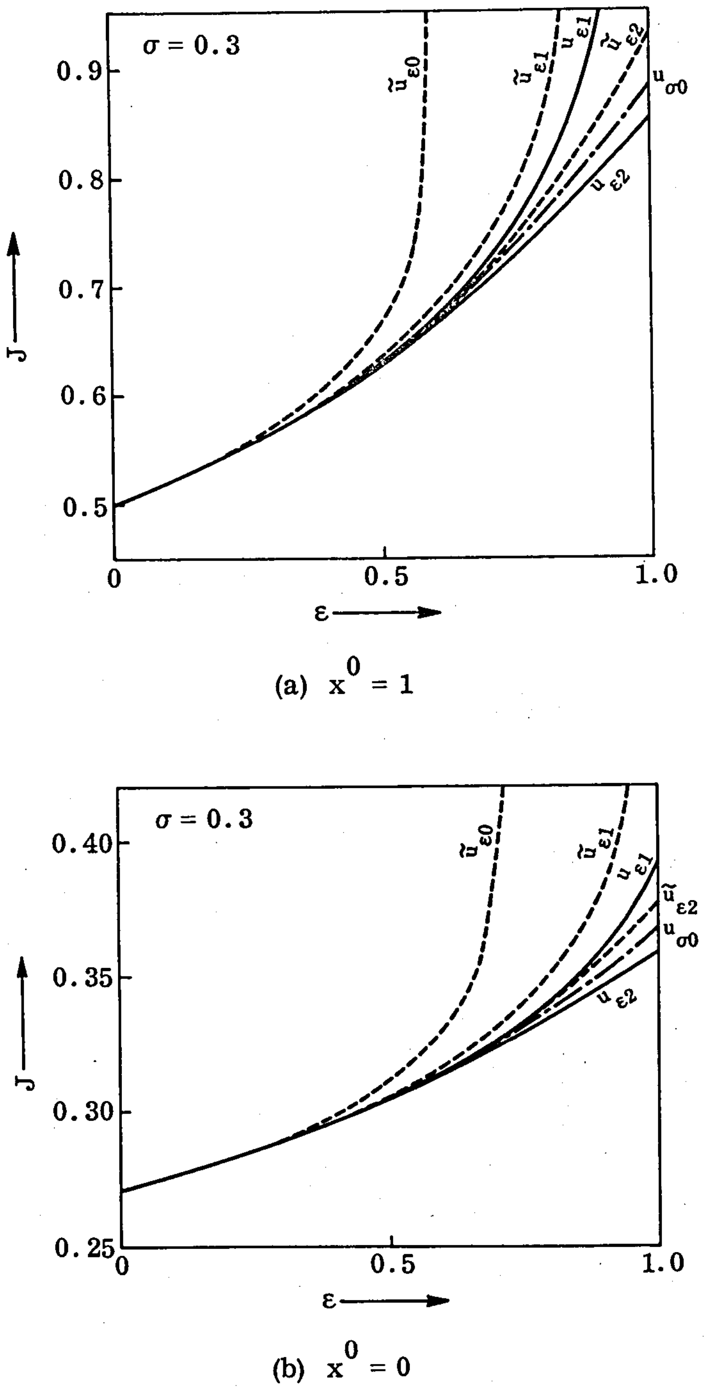


Fig. 4.2. Variation of the performance index J with ε.

is optimal for the nonlinear deterministic system, the same linear control is optimal also for the corresponding noisy system.

For a simple one-dimensional problem, the suboptimal solutions have been calculated in two different forms: the power-series form in ε and the power-series form in σ. Averaged values of the performance index resulting from these suboptimal controls have been compared with each other by using various random-number sequences generated by a computer. We have seen that the suboptimal control with consideration of noise is almost better than the same order control without it.

CHAPTER 5

SUBOPTIMAL SOLUTION TO NONLINEAR DIFFERENTIAL GAME PROBLEMS

5.1 Introduction

Since R. Isaacs initiated the study of differential games [35], there have been numerous works on the subject. Most of them have treated zero-sum two-person problems governed by linear dynamics. The games of linear pursuit-evasion type are extensively investigated in those works [33, 83]. However, in order to apply the differential-game theory to a wider class of problems, including those of economics and behavioral sciences, the restrictions such as zero-sum, two-person, and/or linearity must be removed. This chapter presents a first-step treatment of a class of zero-sum and nonzero-sum differential games with slightly nonlinear dynamics [65].

An elementary game, first discussed by J. von Neumann and O. Morgenstern [63], is a finite zero-sum two-person game which can be formulated as a rectangular game. In the zero-sum rectangular game, two players have completely conflicting interests. Under such a situation, at least one set of minimax strategies exists among mixed strategies [60, p. 21]. Since the minimax strategy set assures an equilibrium between the interests of two rational players, it can be regarded as an optimal solution of the zero-sum two-person game. However, if the restriction of zero-sum or two-person is removed, the interests of players do not necessarily conflict. Under this situation the notion of solution becomes more complicated. In fact, various notions for solution have been developed; among them the Nash equilibrium solution, the minimax solution, and the Pareto optimal solution are typical ones. In

Section 5.2 of the present chapter, we illustrate important differences between zero-sum and nonzero-sum games, referring to simple rectangular games.

Section 5.3 introduces a zero-sum two-person differential game whose state dynamics is governed by a slightly nonlinear differential equation. The nonlinearity is assumed analytic in states, and is taken to be a perturbation to the state dynamics. A parameter ε is introduced to stand for the perturbation. For simplicity, a cost function is considered to be quadratic in states and strategic controls. We develop a method which gives a minimax solution of the game in power-series form in ε. The zero-order term gives a well-known Riccati equation, and higher-order terms yield a series of linear differential equations. This procedure is an application of the method for suboptimal design of nonlinear regulators developed in Chapter 2.

Section 5.4 treats a nonzero-sum N-person differential game with slightly nonlinear dynamics. Each player has his own cost of the quadratic type. The method similar to that of Section 5.3 is applied to approximate determination of the Nash equilibrium solution, the minimax solution, and the Pareto optimal solution. This is an extension of the work done by A. W. Staar and Y. C. Ho [90] on the problem with linear dynamics.

Section 5.5 illustrates a simple example of two-player game with scalar dynamics and examines some features of a differential game.

5.2 Fundamental Notions of Solution of Games

Before discussing our differential games, we briefly explain some fundamental notions of the "solution" of game. Especially we try to illustrate the conceptual difference between the solutions of zero-sum and nonzero-sum games, using simple rectangular games.

5.2.1 Zero-Sum Rectangular Games

In Game I of Fig. 5.1, Player 1 chooses the strategy u_1 or v_1, while Player 2 chooses u_2 or v_2. The entries of the matrix show the cost J_1 of Player 1 and J_2 of Player 2 corresponding to a chosen pair of strategies. For example, the costs $J_1 = 1$ and $J_2 = -1$ result from the strategy pair (u_1, u_2). Since $J_1 + J_2 = 0$ for every strategy pair, Game I is zero-sum. In the zero-sum game the interests of two players completely conflict: What is best for one player is worst to the other. Hence each player must play taking the worst situation for him, the best for the opponent, into account.

Suppose that Player 1 chooses the strategy u_1. Then J_1 is maximum, i.e., $J_1 = 1$ when Player 2 uses his strategy u_2. Similarly if Player 1 chooses v_1 the maximum is $J_1 = 2$. Then, of course Player 1 will choose u_1 by which the maximum possible, i.e., the worst cost is minimized. This principle for decision of strategy is called the minimax principle. Player 2, if he is rational, will also choose his minimax strategy u_2. By use of the minimax strategy pair, the resulting costs are just ones expected by the players; then the game is solved, and the values $J_1 = 1$ and $J_2 = -1$ are called the security level of the players. A zero-sum game whose optimal solution exists among pure strategies is called a closed game.

Generalizing the above idea, we define a minimax strategy in an N-person game: A strategy u_i^* is a minimax strategy for Player i if, for any admissible strategies $u_1, u_2, \ldots, u_N$,

$$\max_{u_j, j\neq i} J_i(u_1, \ldots, u_{i-1}, u_i^*, u_{i+1}, \ldots, u_N) \leq \max_{u_j, j\neq i} J_i(u_1, \ldots, u_i, \ldots, u_N) \tag{5.1}$$

Player 2

Player 1	u_2	v_2
u_1	1, -1	0, 0
v_1	2, -2	-2, 2

(a) Game I: A closed game

Player 2

Player 1	u_2	v_2
u_1	-1, 1	0, 0
v_1	2, -2	-2, 2

(b) Game II: An open game

Fig. 5.1. Typical zero-sum rectangular games.

Player 2

Player 1	u_2	v_2
u_1	0, 1	2, 2
v_1	2, 2	1, 0

(a) Game III: Dating game

Player 2

Player 1	u_2	v_2
u_1	2, 2	10, 1
v_1	1, 10	5, 5

(b) Game IV: Prisoner's dilemma game

Fig. 5.2. Typical nonzero-sum rectangular games.

where $J_i(u_1, u_2, \ldots, u_N)$ is the cost to be minimized by Player i. The value of the left side in the inequality (5.1) is the security level of Player i.

Game II is also a zero-sum and the minimax pair is (u_1, u_2). However the resulting value $J_1 = -1$ is different from Player 1's expecting value $J_1 = 0$. Now no solution exists among pure strategies.

If Player 1 uses u_1, Player 2 will change the strategy from u_2 to v_2 to reduce his cost. Then Player 1 will use v_1 rather than u_1. Again Player 2 will change his strategy, and so forth. A zero-sum game which can not be solved by pure strategy set is called an open game. In every open zero-sum rectangular game the optimal solution exists among mixed strategies [60, p. 21]. In the mixed strategy, a probability is assigned to each choice of strategy.

5.2.2 Nonzero-Sum Rectangular Games

Game III and Game IV of Fig. 5.2 are nonzero-sum, i.e., $J_1 + J_2$ is not necessarily zero nor constant. In a nonzero-sum game a minimax strategy set does not yield an equilibrium between players. The notion of equilibrium or Nash equilibrium in an N-person game is defined as follows: A set of strategies $(\bar{u}_1, \bar{u}_2, \ldots, \bar{u}_N)$ is a Nash equilibrium strategy set if, for all $i = 1, 2, \ldots, N$,

$$J_i(\bar{u}_1, \ldots, \bar{u}_i, \ldots, \bar{u}_N) \leq J_i(\bar{u}_1, \ldots, \bar{u}_{i-1}, u_i, \bar{u}_{i+1}, \ldots, \bar{u}_N) \qquad (5.2)$$

where u_i is any admissible strategy for Player i.

The intuitive meaning of the Nash equilibrium set is this: The Nash equilibrium strategy is the optimal strategy for each player if all the other players adhere to their Nash strategies. In a zero-sum two-person closed game, a minimax set is a Nash equilibrium set and vice versa.

Game III has two Nash equilibria, (u_1, u_2) and (v_1, v_2), with different costs. Each of these equilibria may not necessarily be chosen without negotiation between the players. But, if Player 1 plays u_1 first and announces his choice to Player 2, Player 2 has no choice but to play u_2, giving an advantage to Player 1. It is advantageous in some nonzero-sum games to disclose one's strategy to the opponents in advance. In contrast, this is never true in zero-sum games.

The Nash equilibrium strategy may be regarded as a sort of solution of a nonzero-sum game. However it does not necessarily provide players with the most desirable results. In Game IV the only Nash equilibrium set is (v_1, v_2), yet (u_1, u_2) clearly gives a better result for both players. Suppose there exist two sets of strategies, S_1 and S_2. If, for every player, the cost due to S_1 is not larger than due to S_2, then S_1 is said to dominate S_2. In this terminology the strategy pair (u_1, u_2) dominates (v_1, v_2). In spite of dominance (u_1, u_2) is hard to be used, because this pair is vulnerable to cheating by one player.

If a negotiation among players is allowed before playing a nonzero-sum N-person game, and players' cooperation is guaranteed, the strategy set which is not dominated by any other strategy set is considered a reasonable object of the negotiation. Such a set is called the Pareto optimal set or the noninferior set of strategies: A strategy set $(\hat{u}_1, \hat{u}_2, \ldots, \hat{u}_N)$ belongs to the Pareto optimal set or the noninferior set if there exists no admissible strategy set $(u_1, u_2, \ldots, u_N)$ satisfying

$$J_i(u_1, u_2, \ldots, u_N) \leq J_i(\hat{u}_1, \hat{u}_2, \ldots, \hat{u}_N) \tag{5.3}$$

for all $i = 1, 2, \ldots, N$ with at least one strict inequality. In Game III the strategy pairs (u_1, u_2) and (v_1, v_2) belong to the Pareto optimal set and in Game IV (u_1, u_2), (u_1, v_2), and (v_1, u_2) do.

When no negotiation is allowed, there is not available so far any principle which can be considered completely reasonable. Although the minimax principle is a means for decision of strategy, the idea is too pessimistic for nonzero-sum games. Because the minimax strategy for a player, say, the i-th player is chosen on the assumption that all the other players wish to maximize the i-th player's cost without regard to their own costs. Also the notion of equilibrium may not be appreciated too highly. However we have to notice the following: When a game is played repeatedly for a fairly long period, by a group of people, it sometimes happens that they fall into the habit of playing an equilibrium [94, p. 65]. Thus the Nash equilibrium strategy could be accepted as a conventional standard of behavior for the group: One player violates the convention at one's peril, unless one succeeds in persuading others to violate it also.

Discussion about the notions of the solution of game is discontinued, because it is not a subject of the present chapter.

5.3 Zero-Sum Differential Games

5.3.1 General Zero-Sum Differential Games

We consider the slightly nonlinear system

$$\dot{x} = A(t)x + \varepsilon f(x, t) + B_1(t)u_1 + B_2(t)u_2 \tag{5.4}$$

with the initial condition

$$x(\tau) = x^0 \tag{5.5}$$

In Eq. (5.4),

x: n-state vector

u_i: m_i-strategic control vector of Player i ($i = 1, 2$)

$f(x, t)$: n-vector function analytic in x and continuous in t satisfying $f(0, t) = 0$ for any t

$A(t)$: $n \times n$-matrix continuous in t

$B_i(t)$: $n \times m_i$-matrix continuous in t $(i = 1, 2)$

ε: small scalar parameter

Player 1 wishes to minimize the cost

$$J(u_1, u_2) = \frac{1}{2}x'(T)Fx(T) + \frac{1}{2}\int_{\tau}^{T} [x'Q(t)x + u_1'R_1(t)u_1 + u_2'R_2(t)u_2]dt \qquad (5.6)$$

where

F: $n \times n$-positive semidefinite symmetric matrix

$Q(t)$: $n \times n$-positive semidefinite symmetric matrix continuous in t

$R_1(t)$: $m_1 \times m_1$-positive definite symmetric matrix continuous in t

$R_2(t)$: $m_2 \times m_2$-negative definite symmetric matrix continuous in t

On the other hand, Player 2 wishes to maximize the same cost of Eq. (5.6). Accordingly, the problem is of a zero-sum type and the game is solved by determining minimax strategies.

The minimax strategy pair u_1^* and u_2^* is such that satisfies

$$J(u_1^*, u_2) \leq J(u_1^*, u_2^*) \leq J(u_1, u_2^*) \qquad (5.7)$$

for arbitrary admissible strategies u_1 and u_2. The similarity of the differential game problem to the problem of optimal control is immediately apparent; it is only necessary to determine the strategies of feedback form. We can also apply the usual minimum principle to the problem of Eq. (5.4) to (5.6). The Hamiltonian is given by

$$H = \frac{1}{2}(x'Qx + u_1'R_1u_1 + u_2'R_2u_2) + p'(Ax + \varepsilon f + B_1u_1 + B_2u_2) \qquad (5.8)$$

where $p(x, t)$ is the n-costate vector satisfying

$$\dot{p} = -\frac{\partial H}{\partial x} = -Qx - A'p - \varepsilon\left(\frac{\partial f}{\partial x}\right)'p, \qquad p(x, T) = Fx(T) \tag{5.9}$$

A necessary condition for the minimax solution is that the Hamiltonian be minimum with respect to u_1 and maximum with respect to u_2. Hence the solution does not exist unless

$$\frac{\partial^2 H}{\partial u_1^2} = R_1 > 0, \qquad \frac{\partial^2 H}{\partial u_2^2} = R_2 < 0 \tag{5.10}$$

which are satisfied, by assumption, in our problem. Thereupon the solution is to be given by, for $i = 1, 2$,

$$\frac{\partial H}{\partial u_i} = R_i u_i^* + B_i' p = 0 \tag{5.11}$$

or

$$u_i^* = -R_i^{-1} B_i' p \tag{5.12}$$

It can readily be seen that, for the dynamics (5.4) with $\varepsilon = 0$, the solution is written as

$$u_i^* = -R_i^{-1} B_i' Px \tag{5.13}$$

where the matrix P, symmetric, is the solution of the Riccati equation

$$\dot{P} = -PA - A'P - Q + PEP, \qquad P(T) = F \tag{5.14}$$

with

$$E \triangleq B_1 R_1^{-1} B_1' + B_2 R_2^{-1} B_2' \tag{5.15}$$

The exact determination of the true minimax strategy is generally impossible for nonlinear problems. The present aim is to obtain an approximate strategy in a series form. For a sufficiently small value of ε, it may be a reasonable assumption that the minimax solution be analytic in ε for all $t \in [\tau, T]$. Therefore we develop the costate $p(x, t)$ into a power series with respect to ε:

$$p(x, t) = P(t)x + \sum_{k=1}^{\infty} \varepsilon^k p^{(k)}(x, t) \tag{5.16}$$

Substituting Eq. (5.16) into (5.9) with use of Eqs. (5.4) and (5.12) and equating the coefficients of the like powers of ε separately, we obtain a sequence of equations for the function $p^{(k)}$:

$$\frac{\partial p^{(k)}}{\partial t} + \frac{\partial p^{(k)}}{\partial x} Sx + S' p^{(k)} = h^{(k)} \qquad (k = 1, 2, \ldots) \tag{5.17}$$

and the boundary conditions

$$p^{(k)}(x, T) = 0 \tag{5.18}$$

The vector functions $h^{(k)}(x, t)$ are defined by

$$h^{(1)}(x, t) \triangleq -\left(\frac{\partial f}{\partial x}\right)' Px - Pf,$$

$$h^{(k)}(x, t) \triangleq -\left(\frac{\partial f}{\partial x}\right)' p^{(k-1)} - \frac{\partial p^{(k-1)}}{\partial x} f + \sum_{i=1}^{k-1} \frac{\partial p^{(i)}}{\partial x} E p^{(k-i)} \tag{5.19}$$

$$(k = 2, 3, \ldots)$$

and the matrix S is defined by

$$S \triangleq A - EP \tag{5.20}$$

The zero-order equation is identical with Eq. (5.14).

Equations (5.17) and (5.19) are the same as Eqs. (2.65) and (2.67), respectively, so that, if the solution of Eq. (5.14) exists for $t \in [\tau, T]$, all the statements in Section 2.8 are true for our game problem as well. Consequently Eqs. (5.17) are successively solved with increasing k to determine the functions $p^{(k)}(x, t)$. For more details of the procedure for obtaining the suboptimal strategies, we can refer to Section 2.8.

It should be noted, however, that the stability theorem for the suboptimal control function, Theorem 2.4, does not always apply, owing to no assurance of having a stable solution of Eq. (5.14) as $t \to -\infty$. In fact, even when Q is positive definite, the solution of Eq. (5.14) may have a finite escape time.

The stability theorem regarding the function $p^{(k)}(x, t)$ of the present problem is as follows:

<u>Theorem 5.1</u>

Let us suppose that all the parameters in the system equation (5.4) and the cost (5.6) are time-invariant. Furthermore, we suppose that $F = 0$. If the solution $P(t)$ of Eq. (5.14) is stable as $t \to -\infty$, i.e., there exists the constant matrix such that

$$P_0 \triangleq \lim_{t \to -\infty} P(t) \tag{5.21}$$

and if the constant matrix S_0 defined by

$$S_0 \triangleq A - EP_0 \tag{5.22}$$

is stable as $t \to \infty$, then the solution $p^{(k)}(x, t)$ of Eq. (5.17) is uniformly asymptotically stable as $t \to -\infty$, relative to the polynomial $p_0^{(k)}(x)$ which satisfies the equation

$$x'S_0'p_0^{(k)} = \theta_0^{(k)} \qquad (k = 1, 2, \ldots) \tag{5.23}$$

where

$$\begin{aligned} \theta_0^{(1)}(x) &\triangleq - x'P_0 f, \\ \theta_0^{(k)}(x) &\triangleq - f'p_0^{(k-1)} + \frac{1}{2}\sum_{i=1}^{k-1} p_0^{(i)\prime}Ep_0^{(k-i)} \qquad (k = 2, 3, \ldots) \end{aligned} \tag{5.24}$$

We can use again Lemma 2.5 in Chapter 2 to complete the proof of this theorem.

5.3.2 Pursuit-Evasion Games

A special class of zero-sum differential games is a familiar pursuit-evasion game. So far linear pursuit-evasion games have been studied most extensively [33, 83]. In this subsection, we treat a type of pursuit-evasion game whose state is governed by the following differential equations:

$$\begin{aligned} \dot{x}_p &= A_p(t)x_p + \varepsilon f_p(x_p, t) + B_p(t)u_p, \qquad & x_p(\tau) &= x_p^0, \\ \dot{x}_e &= A_e(t)x_e + \varepsilon f_e(x_e, t) + B_e(t)u_e, \qquad & x_e(\tau) &= x_e^0 \end{aligned} \tag{5.25}$$

where

x_p: n-state vector of the pursuer

x_e: n-state vector of the evader

u_p: m_p-strategic control vector of the pursuer

u_e: m_e-strategic control vector of the evader

$f_i(x_i, t)$: n-vector function analytic in x_i and continuous in t satisfying $f_i(0, t) = 0$ for any t ($i = p, e$)

$A_i(t)$: $n \times n$-matrix continuous in t ($i = p, e$)

$B_i(t)$: $n \times m_i$-matrix continuous in t ($i = p, e$)

The cost to be minimized by the pursuer is assumed to be quadratic:

$$J = \frac{1}{2}[x_p(T) - x_e(T)]'\tilde{F}[x_p(T) - x_e(T)] + \frac{1}{2}\int_{\tau}^{T}[u_p' R_p(t)u_p - u_e' R_e(t)u_e]dt \quad (5.26)$$

where

$\tilde{F}$: $n \times n$-positive semidefinite symmetric matrix

$R_i(t)$: $m_i \times m_i$-positive definite symmetric matrix continuous in t ($i = p, e$)

The pursuer wishes to minimize J of Eq. (5.26) and the evader to maximize it, meaning a zero-sum game. Thus the game is solved by determining a minimax strategy.

By defining

$$x \triangleq \begin{bmatrix} x_p \\ x_e \end{bmatrix}, \quad x^0 \triangleq \begin{bmatrix} x_p^0 \\ x_e^0 \end{bmatrix}, \quad u_1 \triangleq u_p, \quad u_2 \triangleq u_e,$$

$$f \triangleq \begin{bmatrix} f_p \\ f_e \end{bmatrix}, \quad A \triangleq \begin{bmatrix} A_p & 0 \\ 0 & A_e \end{bmatrix}, \quad B_1 \triangleq \begin{bmatrix} B_p \\ 0 \end{bmatrix}, \quad B_2 \triangleq \begin{bmatrix} 0 \\ B_e \end{bmatrix}, \quad (5.27)$$

$$F \triangleq \begin{bmatrix} \tilde{F} & -\tilde{F} \\ -\tilde{F} & \tilde{F} \end{bmatrix}, \qquad R_1 \triangleq R_p, \qquad R_2 \triangleq -R_e$$

the state equations and the cost are rewritten into

$$\dot{x} = Ax + \varepsilon f + B_1 u_1 + B_2 u_2, \qquad x(\tau) = x^0 \tag{5.28}$$

$$J = \frac{1}{2}x'(T)Fx(T) + \frac{1}{2}\int_{\tau}^{T}(u_1' R_1 u_1 + u_2' R_2 u_2)dt \tag{5.29}$$

which are of the same form as Eq. (5.4) to (5.6). Thus the minimax solution is given by

$$u_p^* = -R_p^{-1}B_p' p, \qquad u_e^* = R_e^{-1}B_e' p \tag{5.30}$$

The zero-order coefficient $P(t)$ in the power-series solution (5.16) satisfies

$$\dot{P} = -PA - A'P + PEP, \qquad P(T) = F \tag{5.31}$$

with

$$E \triangleq \begin{bmatrix} B_p R_p^{-1} B_p' & 0 \\ 0 & -B_e R_e^{-1} B_e' \end{bmatrix} \tag{5.32}$$

The higher-order equations are identical with Eqs. (5.17) to (5.19).

5.4 Nonzero-Sum Differential Games

5.4.1 Problem Statement

In this section we are concerned with a nonzero-sum N-person differential game of nonlinear-quadratic type, and examine its solutions of various kinds.

For $i = 1, 2, \ldots, N$, Player i wishes to choose his strategy u_i to minimize

$$J_i = \frac{1}{2}x'(T)F_i x(T) + \frac{1}{2}\int_{\tau}^{T} [x'Q_i(t)x + \sum_{j=1}^{N} u_j' R_{ij}(t)u_j]dt \tag{5.33}$$

subject to the state-dynamics constraint

$$\dot{x} = A(t)x + \varepsilon f(x, t) + \sum_{j=1}^{N} B_j(t)u_j, \qquad x(\tau) = x^0 \tag{5.34}$$

where

u_i: m_i-strategic control vector of Player i

F_i: $n \times n$-positive semidefinite symmetric matrix

$Q_i(t)$: $n \times n$-positive semidefinite symmetric matrix continuous in t

$R_{ij}(t)$: $m_j \times m_j$-symmetric matrix continuous in t

$B_i(t)$: $n \times m_i$-matrix continuous in t

N: number of players

The other quantities x, f, and A are the same as defined in Section 5.3.1. The sign definiteness of the matrices R_{ij} will be specified later, corresponding to the kinds of the game solution considered.

We must recall that the nonzero-sum rectangular games in Section 5.2 could not be approached until one specified what properties the solution should have. Similarly, in the nonzero-sum differential game, one must demand that the solution have some attribute such as the Nash equilibrium, the minimax, or the Pareto optimality. In this section solutions with these properties are investigated by using a power-series method analogous to that of Section 5.3.

5.4.2 Nash Equilibrium Solution

With consideration of the definition (5.2) of Nash equilibrium, we may use the idea of the usual minimum principle to obtain a solution. The Hamiltonian for Player i of the problem is given by

$$H_i = \frac{1}{2}x'Q_ix + \frac{1}{2}\sum_{j=1}^{N} u_j'R_{ij}u_j + p_i'(Ax + \varepsilon f + \sum_{j=1}^{N} B_ju_j)$$

$$(i = 1,\ 2,\ \ldots,\ N) \qquad (5.35)$$

The Nash strategy for Player i, $u_i = \bar{u}_i$, is the strategy which minimizes H_i on the assumption that all of the other players choose their Nash strategies. Therefore $p_i(x, t)$, the n-costate vector, in Eq. (5.35) satisfies the differential equation

$$\dot{p}_i = -\frac{\partial H_i}{\partial x} - \sum_{j=1,\, j\neq i}^{N} \left(\frac{\partial \bar{u}_j}{\partial x}\right)' \frac{\partial H_i(\bar{u}_1, \ldots, \bar{u}_{i-1}, u_i, \bar{u}_{i+1}, \ldots, \bar{u}_N)}{\partial u_j},$$

$$p_i(x,\ T) = F_ix(T) \qquad (i = 1,\ 2,\ \ldots,\ N) \qquad (5.36)$$

The Nash solution for Player i is

$$\bar{u}_i = -\,R_{ii}^{-1}B_i'p_i \qquad (5.37)$$

provided that

$$\frac{\partial^2 H_i}{\partial u_i^2} = R_{ii} > 0 \qquad (i = 1, 2, \ldots, N) \tag{5.38}$$

Here we develop the costate vector $p_i(x, t)$ into a power series with respect to ε:

$$p_i(x, t) = P_i(t)x + \sum_{k=1}^{\infty} \varepsilon^k p_i^{(k)}(x, t) \tag{5.39}$$

Proceeding in a way similar to that of Section 5.3, we obtain a sequence of equations for successive determination of P_i and $p_i^{(k)}$:

$$\dot{P}_i = -P_iA - A'P_i - Q_i - \sum_{j=1}^{N} (P_j\bar{E}_{ij}P_j - P_i\bar{E}_{jj}P_j - P_j\bar{E}_{jj}P_i)$$

$$(i = 1, 2, \ldots, N) \tag{5.40}$$

$$\frac{\partial p_i^{(k)}}{\partial t} + \frac{\partial p_i^{(k)}}{\partial x}\bar{S}x + \bar{S}'p_i^{(k)} + \sum_{j=1, j\neq i}^{N} \left\{ [\frac{\partial p_j^{(k)}}{\partial x}]'T_{ij}x + T'_{ij}p_j^{(k)} \right\} = h_i^{(k)}$$

$$(i = 1, 2, \ldots, N;\ k = 1, 2, \ldots) \tag{5.41}$$

with the boundary conditions

$$P_i(T) = F_i, \qquad p_i^{(k)}(x, T) = 0 \tag{5.42}$$

In Eq. (5.41)

$$h_i^{(1)}(x, t) \triangleq -(\frac{\partial f}{\partial x})'P_ix - P_if,$$

$$h_i^{(k)}(x, t) \triangleq -(\frac{\partial f}{\partial x})'p_i^{(k-1)} - \frac{\partial p_i^{(k-1)}}{\partial x} f + \sum_{\ell=1}^{k-1} \sum_{j=1}^{N} \{ [\frac{\partial p_i^{(\ell)}}{\partial x} \bar{E}_{jj}p_j^{(k-\ell)} \tag{5.43}$$

$$+ [\frac{\partial p_j^{(\ell)}}{\partial x}]'\bar{E}_{jj}p_i^{(k-\ell)} - [\frac{\partial p_j^{(\ell)}}{\partial x}]'\bar{E}_{ij}p_j^{(k-\ell)}\Big\} \qquad (k = 2, 3, \ldots)$$

$\bar{S}$, $\bar{E}_{ij}$, and T_{ij} are the matrices defined by

$$\begin{aligned} \bar{S} &\triangleq A - \sum_{j=1}^{N} \bar{E}_{jj}P_j, \\ \bar{E}_{ij} &\triangleq B_j R_{jj}^{-1} R_{ij} R_{jj}^{-1} B_j', \qquad T_{ij} \triangleq \bar{E}_{ij}P_j - \bar{E}_{jj}P_i \end{aligned} \tag{5.44}$$

If the system (5.34) is linear ($\varepsilon = 0$), the solutions of Eqs. (5.40) exactly give the Nash strategy. It is noted that, with $N = 1$, the problem reduces to the optimal regulator problem in Chapter 2 and Eq. (5.40) to the familiar Riccati equation (2.14). The equilibrium results only from a simultaneous use of the Nash strategies by all the players. The coupling of Eqs. (5.40) or (5.41) for all i reflects this fact.

If the matrices $\partial p_i^{(\ell)}/\partial x$ are symmetric for $i = 1, 2, \ldots, N$ and $\ell = 1, 2, \ldots, k - 1$, then Eqs. (5.43) are integrable with respect to x, and Eqs. (5.41) are rewritten into

$$\begin{aligned} \frac{\partial p_i^{(k)}}{\partial t} + \frac{\partial p_i^{(k)}}{\partial x}\bar{S}x + \bar{S}'p_i^{(k)} = \frac{\partial}{\partial x}\Big\{ &- \sum_{j=1, j\neq i}^{N} x'T_{ij}'p_j^{(k)} - f'p_i^{(k-1)} \\ &+ \sum_{\ell=1}^{k-1}\sum_{j=1}^{N} [p_i^{(\ell)\prime}\bar{E}_{jj}p_j^{(k-\ell)} - \frac{1}{2}p_j^{(\ell)\prime}\bar{E}_{ij}p_j^{(k-\ell)}]\Big\} \end{aligned} \tag{5.45}$$

Though the right side of the above equation contains the other players' unknown functions $p_j^{(k)}$ ($j = 1, 2, \ldots, N$; $j \neq i$), it is integrable with respect to x, which inductively shows the symmetry of the matrices $\partial p_i^{(k)}/\partial x$ for all k and i like as in Section 2.8.

Hence, in each step of the successive calculation of Eqs. (5.41), it is required to solve the equations

$$\frac{\partial g_i}{\partial t} + \frac{\partial}{\partial x}(x'\bar{S}'g_i + \sum_{j=1, j\neq i}^{N} x'T'_{ij}g_j) = \frac{\partial \theta_i}{\partial x} \qquad (i = 1, 2, \ldots, N) \tag{5.46}$$

where $g_i(x, t)$ are unknown n-vectors with the boundary conditions

$$g_i(x, T) = 0 \tag{5.47}$$

and $\theta_i(x, t)$ are given scalars. The functions θ_i are assumed to be homogeneous polynomials of the degree $r + 1$ in x, and are written as*

$$\theta_i(x, t) = \frac{1}{r+1}\Theta^i_{j_0 j_1 \ldots j_r}(t)x_{j_0}x_{j_1}\ldots x_{j_r} \tag{5.48}$$

Thus the j_0-th components of g_i may be written as

$$g_{ij_0}(x, t) = G^i_{j_0 j_1 \ldots j_r}(t)x_{j_1}x_{j_2}\ldots x_{j_r} \tag{5.49}$$

with the conditions

$$G^i_{j_0 j_1 \ldots j_r}(T) = 0 \tag{5.50}$$

The coefficients $\Theta^i_{j_0 j_1 \ldots j_r}$ and $G^i_{j_0 j_1 \ldots j_r}$ are symmetric with respect to the lower indices. By substituting Eqs. (5.48) and (5.49) into

---

* Here and in what follows the summation convention as in Section 2.8 is used again.

(5.46), we have

$$\dot{G}^i_{j_0 j_1 \ldots j_r} = - S_{kj_0} G^i_{kj_1 \ldots j_r} - S_{kj_1} G^i_{j_0 k j_2 \ldots j_r} - \ldots - S_{kj_r} G^i_{j_0 \ldots j_{r-1} k}$$

$$- \sum_{j=1, j \neq i}^{N} [T^{(ij)}_{kj_0} G^j_{kj_1 \ldots j_r} + T^{(ij)}_{kj_1} G^j_{j_0 k j_2 \ldots j_r} + \ldots$$

$$+ T^{(ij)}_{kj_r} G^j_{j_0 \ldots j_{r-1} k}] + \Theta^i_{j_0 j_1 \ldots j_r}$$

$$(i = 1, 2, \ldots, N; \; j_0, j_1, \ldots, j_r = 1, 2, \ldots, n) \quad (5.51)$$

where $S_{k\ell}$ and $T^{(ij)}_{k\ell}$ are the (k, ℓ)-elements of the matrices $\bar{S}$ and T_{ij}, respectively.

By virtue of Lemma 2.3 in Section 2.9, Eqs. (5.51) are rewritten into

$$\dot{v}_i = - S^{(r+1)}(t) v_i - \sum_{j=1, j \neq i}^{N} T^{(r+1)}_{ij}(t) v_j + b_i(t) \qquad (i = 1, 2, \ldots, N) \quad (5.52)$$

with the boundary conditions

$$v_i(T) = 0 \quad (5.53)$$

$S^{(r+1)}$ and $T^{(r+1)}_{ij}$ are the $(r + 1)$-times repeated Kronecker sums of $\bar{S}'$ and T'_{ij}, respectively; v_i and b_i are the n^{r+1}-vectors such that the ℓ-th components are $G^i_{j_0 j_1 \ldots j_r}$ and $\Theta^i_{j_0 j_1 \ldots j_r}$, respectively, ℓ being

$$\ell \triangleq \sum_{k=0}^{r} (j_k - 1) n^{r-k} + 1 \quad (5.54)$$

Assume that, as $t \to -\infty$, $\bar{S}$ and T_{ij} reduce to the constant matrices S_0 and T_{0ij}. Let $S_0^{(r)}$ and $T_{0ij}^{(r)}$ be the matrices of r-times repeated Kronecker sums of S_0' and T_{0ij}', respectively, and let $M^{(r)}$ be the $Nn^r \times Nn^r$-composite matrix defined by

$$M^{(r)} \triangleq \begin{bmatrix} S_0^{(r)} & T_{012}^{(r)} & \cdots & T_{01N}^{(r)} \\ T_{021}^{(r)} & S_0^{(r)} & \cdots & T_{02N}^{(r)} \\ & \cdots\cdots & & \\ T_{0N1}^{(r)} & T_{0N2}^{(r)} & \cdots & S_0^{(r)} \end{bmatrix} \tag{5.55}$$

The behavior of the solution of Eq. (5.52) is characterized by the stability of the matrix $M^{(r+1)}$. Contrary to the optimal control problem in Section 2.10,* the stability of $M^{(r)}$ is not the direct consequence of that of S_0 alone, as will be seen later in an example, and the stability of $M^{(r)}$ for $r \geq 2$ is not explicitly related to that of $M^{(1)}$. Therefore the stability theorem for the function $p_i^{(k)}$ can not be mentioned in a simple form as Theorem 2.4.

5.4.3 Minimax Solution

The minimax strategy for Player i, $u_i = u_i^*$, is the strategy satisfying the inequality (5.1). The strategy u_i^* is to minimize the Hamiltonian H_i of Eq. (5.35) on the assumption that all the other players try to maximize H_i, i.e., they choose

$$u_j = - R_{ij}^{-1} B_j' p_i \qquad (j \neq i) \tag{5.56}$$

---

* We have observed, in Section 2.10, that, if Eqs. (2.92) are time-invariant and if S is stable, the solution v(t) of Eqs. (2.92) is stable as $t \to -\infty$.

provided that

$$R_{ij} < 0 \qquad (j \neq i) \tag{5.57}$$

Then the minimax strategy for Player i is

$$u_i^* = - R_{ii}^{-1} B_i' p_i \tag{5.58}$$

provided that

$$R_{ii} > 0 \tag{5.59}$$

where p_i satisfies

$$\dot{p}_i = - \frac{\partial H_i}{\partial x} \tag{5.60}$$

When the condition $R_{ij} < 0$, the inequality (5.57), fails to satisfy, the expected cost by Player i, i.e., his security level, is infinite, reducing the minimax solution meaningless.

By developing the costate vector p_i into a power series of Eq. (5.39), we obtain

$$\dot{P}_i = - P_i A - A' P_i - Q_i + \sum_{j=1}^{N} P_i E_{ij} P_i \qquad (i = 1, 2, \ldots, N) \tag{5.61}$$

$$\frac{\partial p_i^{(k)}}{\partial t} + \frac{\partial p_i^{(k)}}{\partial x} S_i x + S_i' p_i^{(k)} = h_i^{(k)} \qquad (i = 1, 2, \ldots, N;\ k = 1, 2, \ldots) \tag{5.62}$$

where*

$$h_i^{(1)}(x, t) \triangleq -\left(\frac{\partial f}{\partial x}\right)' P_i x - P_i f,$$

$$h_i^{(k)}(x, t) \triangleq -\left(\frac{\partial f}{\partial x}\right)' p_i^{(k-1)} - \frac{\partial p_i^{(k-1)}}{\partial x} f + \sum_{\ell=1}^{k-1} \sum_{j=1}^{N} \frac{\partial p_i^{(\ell)}}{\partial x} E_{ij} p_i^{(k-\ell)} \tag{5.63}$$

$$(k = 2, 3, \ldots)$$

$$E_{ij} \triangleq B_j R_{ij}^{-1} B_j', \qquad S_i \triangleq A - \sum_{j=1}^{N} E_{ij} P_i \tag{5.64}$$

The boundary conditions are

$$P_i(T) = F_i, \qquad p_i^{(k)}(x, T) = 0 \tag{5.65}$$

Equations (5.61) and (5.62) are separated for each i. This is because Player i presumes, by himself, his opponents' strategies, all of which try to maximize his cost J_i. Corresponding to Theorem 5.1 for the zero-sum game, the following theorem is established:

<u>Theorem 5.2</u>

Assume that the system (5.34) is time-invariant, all the parameters in the performance indices (5.33) are constant, and $F_i = 0$. If the constant matrices

$$P_{0i} \triangleq \lim_{t \to -\infty} P_i(t) \tag{5.66}$$

---

* The matrices S, $\bar{S}$, and $\hat{S}$ defined in Eqs. (5.20), (5.44), and (5.79), respectively, are the resultant system matrices with the corresponding optimal strategies for $\varepsilon = 0$; while not so is S_i, because for the minimax solution $A - \sum_{j=1}^{N} E_{jj} P_j$ is the resultant system matrix.

exist and if all the constant matrices defined by

$$S_{0i} \triangleq A - \sum_{j=1}^{N} E_{ij} P_{0i} \tag{5.67}$$

are stable, $p_i^{(k)}$ is uniformly asymptotically stable as $t \to -\infty$, relative to $p_{0i}^{(k)}$ satisfying

$$x' S_{0i}' p_{0i}^{(k)} = \theta_{0i}^{(k)} \qquad (k = 1, 2, \ldots) \tag{5.68}$$

where

$$\begin{aligned} &\theta_{0i}^{(1)}(x) \triangleq - x' P_{0i} f, \\ &\theta_{0i}^{(k)}(x) \triangleq - f' p_{0i}^{(k-1)} + \frac{1}{2} \sum_{\ell=1}^{k-1} \sum_{j=1}^{N} p_{0i}^{(\ell)\prime} E_{ij} p_{0i}^{(k-\ell)} \qquad (k = 2, 3, \ldots) \end{aligned} \tag{5.69}$$

5.4.4 Pareto Optimal Solution

If cooperation among the players is allowable, clearly dominated strategies are disregarded and a strategy is chosen among the Pareto optimal set. Solving for the Pareto optimal set is equivalent to solving an optimal control problem with a vector cost function $(J_1, J_2, \ldots, J_N)$. When values of the costs form a convex polyhedron in the N-dimensional space, the problem is equivalent to an $N - 1$ parameter family of control problems with a scalar cost criterion [15]. That is to say, the Pareto optimal solution $(\hat{u}_1, \hat{u}_2, \ldots, \hat{u}_N)$ minimizes the scalar function

$$J = \sum_{i=1}^{N} \alpha_i J_i \tag{5.70}$$

for some $(\alpha_1, \alpha_2, \ldots, \alpha_N)$ satisfying

$$\sum_{i=1}^{N} \alpha_i = 1, \qquad \alpha_i > 0 \qquad (i = 1, 2, \ldots, N) \tag{5.71}$$

Conversely, for any α_i satisfying Eq. (5.71), the corresponding solution which minimizes J in Eq. (5.70) belongs to the set of Pareto optimal solutions.

The values of α_i satisfying Eq. (5.71) are to be determined by a negotiation among the players. This determination involves a conflict of the players' interests: Decreasing one player's cost enforces increasing at least one of the others'. Here we assume that the values of α_i have already been determined by some means.

The Hamiltonian of the problem (5.34) and (5.70) is given by

$$\begin{aligned} H &\triangleq \frac{1}{2}x'\hat{Q}x + \frac{1}{2}\sum_{i=1}^{N} u_i'\hat{R}_i u_i + p'(Ax + \varepsilon f + \sum_{i=1}^{N} B_i u_i), \\ \hat{Q} &\triangleq \sum_{j=1}^{N} \alpha_j Q_j, \qquad \hat{R}_i \triangleq \sum_{j=1}^{N} \alpha_j R_{ji} \end{aligned} \tag{5.72}$$

where $p(x, t)$ satisfies

$$\begin{aligned} \dot{p} &= -\frac{\partial H}{\partial x} = -\hat{Q}x - A'p - \varepsilon\left(\frac{\partial f}{\partial x}\right)'p, \qquad p(x, T) = \hat{F}x(T), \\ \hat{F} &\triangleq \sum_{i=1}^{N} \alpha_i F_i \end{aligned} \tag{5.73}$$

The Pareto optimal solution $\hat{u}_i$, to minimize H, is obtained:

$$\hat{u}_i = -\hat{R}_i^{-1}B_i'p \qquad (i = 1, 2, \ldots, N) \tag{5.74}$$

provided that

$$\hat{R}_i > 0 \qquad (i = 1, 2, \ldots, N) \tag{5.75}$$

The inequality (5.75) is necessary for existence of $\hat{u}_i$ to minimize H.

By use of the power-series expansion of Eq. (5.16) we obtain

$$\dot{P} = - PA - A'P - \hat{Q} + P\hat{E}P \tag{5.76}$$

$$\frac{\partial p^{(k)}}{\partial t} + \frac{\partial p^{(k)}}{\partial x}\hat{S}x + \hat{S}'p^{(k)} = h^{(k)} \qquad (k = 1, 2, \ldots) \tag{5.77}$$

where

$$h^{(1)}(x, t) \triangleq - (\frac{\partial f}{\partial x})'Px - Pf,$$

$$h^{(k)}(x, t) \triangleq - (\frac{\partial f}{\partial x})'p^{(k-1)} - \frac{\partial p^{(k-1)}}{\partial x}f + \sum_{i=1}^{k-1} \frac{\partial p^{(i)}}{\partial x}\hat{E}p^{(k-i)} \tag{5.78}$$

$$(k = 2, 3, \ldots)$$

$$\hat{E} \triangleq \sum_{i=1}^{N} B_i\hat{R}_i^{-1}B_i', \qquad \hat{S} \triangleq A - \hat{E}P \tag{5.79}$$

The boundary conditions are

$$P(T) = \hat{F}, \qquad p^{(k)}(x, T) = 0 \tag{5.80}$$

The problem of Eqs. (5.34) and (5.70) is now one of the simple optimal control problem rather than game problem. Thus, it is readily seen that the sufficient condition to have Theorem 2.4 is the following: $\hat{Q}$ is positive definite, $\hat{F} = 0$, and A and B are of complete controllability, where B is the $n \times (m_1 + m_2 + \ldots + m_N)$-composite matrix defined by

$$B \triangleq [B_1 \quad B_2 \quad \cdots \quad B_N] \tag{5.81}$$

Then Theorem 2.4 is also true for this case.

5.5 Example

In this section, we examine a simple example and illustrate some features of the differential game. For simplicity, here we treat a two-person game whose state dynamics is governed by the scalar differential equation:

$$\dot{x} = x + \varepsilon x^3 + u_1 + u_2, \qquad x(0) = x^0 \tag{5.82}$$

The cost functions of Players 1 and 2 are, respectively,

$$\begin{aligned} J_1 &= \frac{1}{2}F_1 x(T)^2 + \frac{1}{2}\int_0^T (x^2 + r_1 u_1^2 + r_2 u_2^2)dt, \\ J_2 &= \frac{1}{2}F_2 x(T)^2 + \frac{1}{2}\int_0^T (x^2 + r_2 u_1^2 + r_1 u_2^2)dt \end{aligned} \tag{5.83}$$

where F_1 and F_2 are nonnegative constants. These costs may be called symmetric with respect to u_1 and u_2.

To begin with, we examine values of the parameters r_1 and r_2 for which various kinds of solutions exist. Due to the condition (5.38) a Nash equilibrium solution exists with $r_1 > 0$, and a minimax solution with $r_1 > 0$ and $r_2 < 0$ due to the conditions (5.57) and (5.59). For Pareto optimality $r_1 + r_2 > 0$ is required by the condition (5.75).

(a) A Case in Which Three Kinds of Solutions Exist

Considering the above result we choose

$$r_1 = 2, \qquad r_2 = -1 \tag{5.84}$$

because all the three kinds of solutions exist for this parameter pair. The initial condition and the game period are taken

$$x^0 = 1, \qquad T = 10 \tag{5.85}$$

respectively, throughout the numerical calculation.

(a.1) Pareto optimal solution

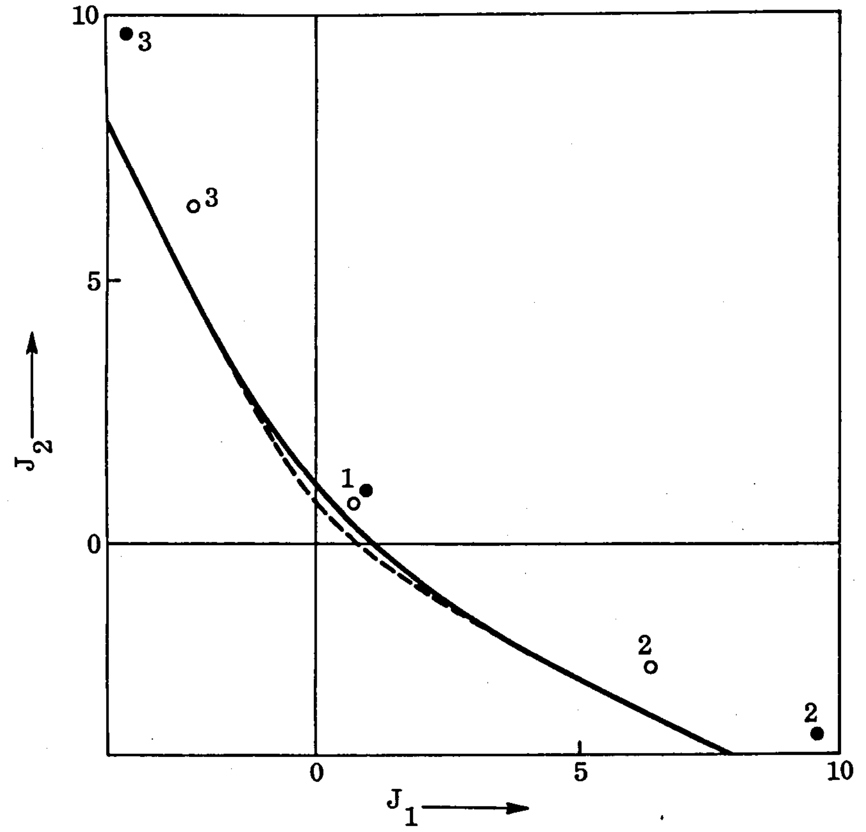
Since the Pareto optimality gives the best possible solutions in a sense, first we examine it. By using the approximation procedure of Section 5.4.4, the Pareto optimal strategies are obtained up to the second order in ε:

$$\hat{u}_{si} = -\frac{\nu_i}{\nu}\left\{(1+\sqrt{1+\nu})x + \frac{1+\nu+\sqrt{1+\nu}}{1+\nu}\varepsilon x^3 + \left[\frac{1+\sqrt{1+\nu}}{1+\nu} - \frac{2(1+\nu)+(2+\nu)\sqrt{1+\nu}}{2(1+\nu)^2}\right]\varepsilon^2 x^5\right\} \quad (i = 1,\ 2) \tag{5.86}$$

where

$$\nu_1 \triangleq \frac{1}{3\alpha - 1}, \qquad \nu_2 \triangleq \frac{1}{2 - 3\alpha}, \qquad \nu \triangleq \nu_1 + \nu_2, \qquad \frac{1}{3} < \alpha < \frac{2}{3} \tag{5.87}$$

To be exact, the coefficients in the power series of Eqs. (5.86) vary with time t, because they are solutions of differential equations. However, with a relatively long game period, e.g., $T = 10$ in our example, their steady-state values can be used through the game. It is noted that



——: Pareto ($\varepsilon = 1$) -----: Pareto ($\varepsilon = 0$)

●: Nash ($\varepsilon = 1$) o: Nash ($\varepsilon = 0$)

Fig. 5.3. Costs J_1 and J_2 due to the Pareto optimal strategies (5.86) and the Nash equilibrium strategies (5.88) or (5.89).

these steady-state values are unique and consequently not affected by the values of F_1 and F_2.

In Fig. 5.3 the sets of costs J_1 and J_2 attained by the Pareto optimal strategies are illustrated with $\varepsilon = 0$ and $\varepsilon = 1$. By definition, no cost below the curve of the Pareto optimality is possible to attain with respective value of ε; that is, the right-upper region to the curve is the attainable set of the costs.

(a.2) Nash equilibrium solution

When $F_1 = F_2$, the approximate strategies are

$$\bar{u}_{s1} = \bar{u}_{s2} = -0.7595x - 0.4580\varepsilon x^3 - 0.0548\varepsilon^2 x^5 \tag{5.88}$$

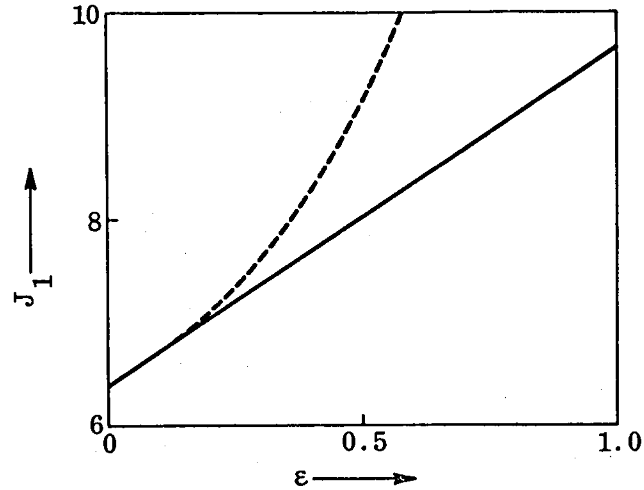
The resulting costs J_1 and J_2 are also equal. However, if $F_1 \neq F_2$, a drastic change occurs in the strategies and the costs. For $F_1 > F_2$

$$\begin{aligned} \bar{u}_{s1} &= -6.3589x - 6.5883\varepsilon x^3 + 0.1208\varepsilon^2 x^5, \\ \bar{u}_{s2} &= 2.3589x + 2.5883\varepsilon x^3 - 0.1208\varepsilon^2 x^5 \end{aligned} \tag{5.89}$$

and J_1 is far larger than J_2. If $F_1 < F_2$, the result is reversed for Player 1 and Player 2. That is to say, if the cost functions are in completely symmetric forms, the strategies and the cost values are same for both players, but a slight difference in the cost parameters results in a big change of the Nash equilibrium solution. This change might well be understood by examining transient variations of the solution, which is detailed for $\varepsilon = 0$ in Appendix V.

The costs attained by the Nash equilibrium strategies are plotted in Fig. 5.3. Points 1, 2, and 3 are for $F_1 = F_2$, $F_1 > F_2$, and $F_1 < F_2$, respectively. The numeral assigned to each point for the case of $\varepsilon = 0$ indicates the corresponding singular point on the phase-plane diagram of Fig. V.1(c) in Appendix V. Clearly every Nash strategy is dominated by some part of the Pareto optimal set.

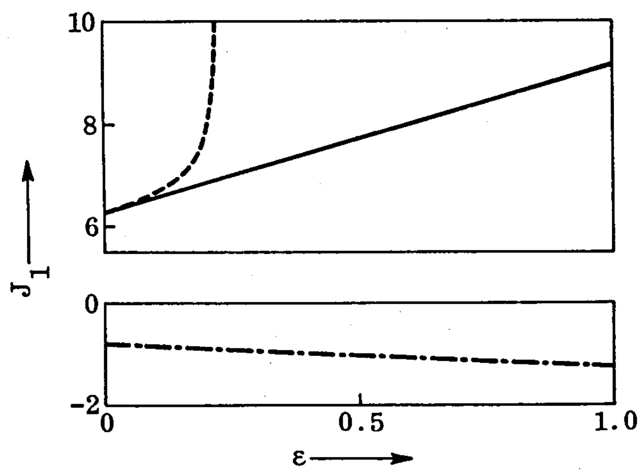
When the opponent uses his Nash strategy, the player can do best by also using the Nash strategy. This is illustrated in Fig. 5.4 for



——: J_1 due to $\bar{u}_{s1}$ of the first or the second order, or $\bar{u}_1$

-----: J_1 due to $\bar{u}_{s1}$ of the zero order

Fig. 5.4. Variation of the Nash equilibrium cost J_1 with ε. Player 2 is using his exact Nash strategy $\bar{u}_2$ of Eqs. (5.94).



——: J_1 due to u^*_{s1} of the first or the second order or u^*_1, and u_2 of Eq. (5.98) (security level of J_1)

-----: J_1 due to u^*_{s1} of the zero order and u_2 of Eq. (5.98)

—·—·—: J_1 due to u^*_1 and u^*_2 of Eqs. (5.97)

Fig. 5.5. Variation of the minimax cost J_1 with ε.

$F_1 > F_2$. Here Player 2 is choosing his Nash strategy. For Player 1, the cost is improved by using higher-order approximations to the Nash strategy. The effect of improvement becomes remarkable as ε increases.

(a.3) Minimax solution

The minimax solution diverges very quickly: In fact, the zero-order solutions of Eqs. (5.61) are given by

$$P_1(t) = P_2(t) = \frac{-[2 + (2 - \sqrt{2})F] + [2 + (2 + \sqrt{2})F]e^{\sqrt{2}(T-t)}}{2 + \sqrt{2} + F - (2 - \sqrt{2} + F)e^{\sqrt{2}(T-t)}}, \qquad (5.90)$$

$$F \triangleq F_1 = F_2$$

These have, for $F \geqq 0$, the final escape time

$$t = T - \frac{1}{\sqrt{2}} \log \frac{2 + \sqrt{2} + F}{2 - \sqrt{2} + F} \qquad (5.91)$$

The detail is omitted here.

(a.4) Application of the method of instantaneous linearization

The method of instantaneous linearization of J. D. Pearson [78] gives the exact solution to a nonlinear regulator problem having a scalar state, as mentioned in the example of Section 2.13(a). The method works also for the exact solution of the present differential game which is governed by the scalar equation.

The Pareto optimal and the Nash equilibrium solutions are calculated: Pareto optimal solutions are

$$\hat{u}_i = -\frac{\nu_i}{\nu}[1 + \varepsilon x^2 + \sqrt{(1 + \varepsilon x^2)^2 + \nu}]x \qquad (5.92)$$

Nash equilibrium solutions are, with $F_1 = F_2$,

$$\bar{u}_1 = \bar{u}_2 = -\frac{2}{7}[1 + \varepsilon x^2 + \sqrt{(1 + \varepsilon x^2)^2 + \frac{7}{4}}]x \tag{5.93}$$

and, with $F_1 > F_2$,

$$\begin{aligned} \bar{u}_1 &= -[2(1 + \varepsilon x^2) + \sqrt{20(1 + \varepsilon x^2)^2 - 1}]x, \\ \bar{u}_2 &= -[2(1 + \varepsilon x^2) - \sqrt{20(1 + \varepsilon x^2)^2 - 1}]x \end{aligned} \tag{5.94}$$

The power-series approximations of the second order, Eqs. (5.86), (5.88), and (5.89), turn out to agree with the formal expansions of the exact strategies, Eqs. (5.92), (5.93), and (5.94), respectively.

(b) A Case in Which the Minimax Solution Has No Finite Escape Time

We choose values of the parameters r_1 and r_2 as

$$r_1 = 2, \qquad r_2 = -3 \tag{5.95}$$

For these values of r_1 and r_2, there is no value of α_i satisfying the condition (5.75) for Pareto optimal solution, and, as seen in Appendix V, the Nash equilibrium solution has no steady-state value except for $F_1 = F_2$.

By the approximation method in Section 5.4.3, we obtain the minimax solution in the steady state up to the second order:

$$u^*_{s1} = u^*_{s2} = -6.2404x - 5.7775\varepsilon x^3 - 0.1984\varepsilon^2 x^5 \tag{5.96}$$

Application of the method of instantaneous linearization gives the exact minimax solution:

$$u_1^* = u_2^* = -3[1 + \varepsilon x^2 + \sqrt{(1 + \varepsilon x^2)^2 + \frac{1}{6}}]x \tag{5.97}$$

When one player, say Player 2, uses the strategy maximizing the other's cost J_1, that is,

$$u_2 = 2[1 + \varepsilon x^2 + \sqrt{(1 + \varepsilon x^2)^2 + \frac{1}{6}}]x \tag{5.98}$$

then the minimax strategy is optimal for Player 1. Figure 5.5 shows the values of J_1 when Player 1 uses the strategy (5.96) or (5.97) and Player 2 uses (5.97) or (5.98). Provided that Player 2 uses the strategy (5.98), then J_1 is almost at the security level, the solid line in the figure, when Player 1 uses his power-series form minimax strategies (5.96) of the more than first order. The value of J_1 resulting from the both players' minimax strategies (5.97) is fairly less than the security level. This shows the pessimistic feature of minimax strategies in a nonzero-sum game.

(c) A Case in Which the Steady-State Higher-Order Solution of Nash Equilibrium Can Not Be Determined Uniquely

By choosing values of the parameters r_1 and r_2 as

$$r_1 = 2, \qquad r_2 = 2\sqrt{17} - 8 \tag{5.99}$$

there arises an irregular situation in computing higher-order solutions of Nash equilibrium. In fact, the steady-state solution as $t \to -\infty$ of Eqs. (5.40) with $F_1 = F_2 \geqq 0$ is given by

$$P_{01} = P_{02} = \frac{\sqrt{17} + 3}{4} \tag{5.100}$$

and S_0, to which $\bar{S}$ defined in Eqs. (5.44) is reduced, is stable:

$$S_0 = - \frac{\sqrt{17} - 1}{4} \tag{5.101}$$

Nevertheless, in this case, the matrix $M^{(r)}$ of Eq. (5.55), by which higher-order solutions of Nash equilibrium is calculated,

$$M^{(r)} = - \frac{\sqrt{17} - 1}{4} r \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \tag{5.102}$$

is singular, and hence the coefficients of the higher order can not be uniquely determined.

5.6 Concluding Remarks

An approximation method has been presented to obtain various solutions of differential game problems with nonlinear dynamics. A zero-sum two-person differential game has been treated in Section 5.3. The minimax strategy of the game is obtained in a power-series form of the small parameter ε. If the nonlinearity of the system is of a polynomial in state, the coefficients of the series are also polynomials and can be determined successively with increasing order.

A nonzero-sum N-person differential game has been discussed in Section 5.4. The nonzero-sum game could not be approached until one specified what properties the solution should have. Typical solutions of the nonzero-sum game are the Nash equilibrium strategy, the minimax strategy, and the Pareto optimal strategy. The small parameter method has also been applied to approximate determination of these solutions.

The stability properties of the approximate minimax solutions have been discussed in Theorems 5.1 and 5.2 for a zero-sum and a nonzero-sum game, respectively. For the Nash equilibrium solution, such a stability property can not be mentioned in a simple form; and, for the Pareto optimal solution, the property is the same as that for the nonlinear regulator solution in Chapter 2.

A nonzero-sum two-person differential game with scalar dynamics has been examined in Section 5.5. Some features have been illustrated, including the above-mentioned stability properties of the approximate solutions.

APPENDIX I

KRONECKER PRODUCT AND SUM OF MATRICES

The definitions and important properties of the Kronecker product and sum of matrices are briefly mentioned [7, p. 227].* Let A and B be $n \times n$- and $m \times m$-matrices, respectively.

<u>Definition I. 1</u>

The $nm \times nm$-matrix whose (p, q)-element is $A_{ij}B_{k\ell}$, p being $m(i - 1) + k$ and q being $m(j - 1) + \ell$, is called the Kronecker product of A and B, and written $A \otimes B$.

<u>Definition I. 2</u>

The $nm \times nm$-matrix $A \otimes I^{(m)} + I^{(n)} \otimes B$, where $I^{(k)}$ is the $k \times k$-identity matrix, is called the Kronecker sum of A and B, and written $A \oplus B$.

Let λ_i $(i = 1, 2, \ldots, n)$ and μ_i $(i = 1, 2, \ldots, m)$ be the eigenvalues of A and B, respectively. Then the Kronecker product and sum have important properties of interest: The nm eigenvalues of $A \otimes B$ are $\lambda_i\lambda_j$ and nm eigenvalues of $A \oplus B$ are $\lambda_i + \lambda_j$ $(i = 1, \ldots, n;\ j = 1, \ldots, m)$.

The following fact has been used to establish Theorem 2.4 in Chapter 2. Let $\tilde{A}$ be the $n^r \times n^r$-matrix which is the r-times repeated Kronecker sum of A:

---

* In some textbooks on matrix theories the terms "direct product and sum" are used in place of the Kronecker product and sum [e.g., 57, p. 81].

$$\tilde{A} \triangleq \underbrace{A \oplus A \oplus \ldots \oplus A}_{r} \qquad \text{(I.1)}$$

It is inductively shown that the n^r eigenvalues of $\tilde{A}$ are given by $\lambda_{j_1} + \lambda_{j_2} + \ldots + \lambda_{j_r}$ $(j_1, j_2, \ldots, j_r = 1, 2, \ldots, n)$. In particular, if A is stable, so is $\tilde{A}$.

Consider the matrix equation for an unknown $n \times m$-matrix X:

$$AX + XB = C \qquad \text{(I.2)}$$

where A, B, and C are given $n \times n$-, $m \times m$-, and $n \times m$-matrices, respectively. It is known that Eq. (I.2) can be rewritten into the equation in an unknown vector x [7, p. 231]:

$$(A \oplus B')x = c \qquad \text{(I.3)}$$

where x and c are the nm-vectors whose ℓ-th components are X_{ij} and C_{ij}, respectively, ℓ being $m(i - 1) + j$.

APPENDIX II

DERIVATION OF DYNAMICAL EQUATIONS FOR SOME CONTROL PROBLEMS OF SPACE VEHICLES

Some examples of dynamical systems optimization problems are found in attitude control and trajectory optimization of a space vehicle. Dynamical equations describing the attitude rotation or the trajectory motion of a satellite are linearized under some reasonable assumptions. Many works are concerned with such linearized models [14, 19, 28, 32, 64, 91]. It may often be required, however, to treat a more precise model. Then we are inevitably faced with nonlinearities and find the ordinary linear theories helpless. In addition, taking nonlinearities into consideration may make the system large-scale due to nonlinear couplings between subsystems.

To solve such kinds of space problems with nonlinearities and/or with high dimensionalities is one of the motivations of the research developed in the present text. Some space problems are adopted as examples mainly in Chapter 3. The present appendix gives the derivation of dynamical equations for two typical space problems used there.

II.1 Equations for Attitude Control of an Orbiting Satellite [64]

Consider an active attitude control system of a satellite orbiting around the earth. The controlled element dynamics is characterized by a set of differential equations relating an external torque to the rotational response of a satellite. Equations for the attitude of a satellite are summarized in what follows.

The attitude of the body is defined by Euler angles ϕ, θ, and ψ, which determine the orientation of three body-fixed axes with respect to earth-centered reference axes. The angles ϕ, θ, and ψ represent the pitch, yaw, and roll motion, respectively, of the body about its principal axes of inertia. The angular velocity of the satellite is written, in terms of components along the body-fixed axes, as

$$\begin{aligned} \omega_1 &= -(\Omega + \phi') \sin\theta + \psi', \\ \omega_2 &= (\Omega + \phi') \sin\psi \cos\theta + \theta' \cos\psi, \\ \omega_3 &= (\Omega + \phi') \cos\psi \cos\theta - \theta' \sin\psi \end{aligned} \tag{II.1}$$

where Ω is the angular velocity of the radial line from the gravity center of the earth to the satellite; only here and throughout this appendix, a prime denotes differentiation with respect to real time t_r, although, in all the other parts of the text, it denotes transposition of a vector or a matrix.

If we assume the satellite to be a spherical body, the Euler equation of the rotational motion is, in component forms,

$$M_i = A\omega_i' \qquad (i = 1, 2, 3) \tag{II.2}$$

where M_i is the i-th component of an external torque applied to the satellite and A is a principal moment of inertia about a diameter of the spherical body.

For the satellite revolving along an elliptic orbit, the angular velocity Ω of the earth-centered frame is approximately given by

$$\Omega = \bar{\Omega}\,(1 + e \cos \bar{\Omega}\, t_r) \tag{II.3}$$

where $\bar{\Omega}$ is the mean value of Ω over one revolution along the orbit and e denotes the eccentricity of the orbit. The time at which the satellite passes through the perigee of the orbit is taken to be zero in Eq. (II.3).

Substitution of Eqs. (II.1) into (II.2) yields a set of second-order differential equations with respect to ϕ, θ, and ψ. Solving these equations for ψ'', θ'', and ϕ'', we obtain

$$
\begin{aligned}
\psi'' &= (\Omega + \phi' + \psi' \sin\theta)\theta' \\
&\quad + \frac{M_1}{A}\cos\theta + \frac{M_2}{A}\sin\psi\sin\theta + \frac{M_3}{A}\cos\psi\sin\theta, \\
\theta'' &= -(\Omega + \phi')\psi'\cos\theta + \frac{M_2}{A}\cos\psi - \frac{M_3}{A}\sin\psi, \\
\phi'' &= (\Omega + \phi')\theta'\tan\theta + \theta'\psi'\sec\theta - \Omega' \\
&\quad + \frac{M_2}{A}\sin\psi\sec\theta + \frac{M_3}{A}\cos\psi\sec\theta
\end{aligned}
\tag{II.4}
$$

Equations (II.4) describe the exact dynamical model of the attitude motion.

We introduce nondimensional quantities defined by

$$
\begin{aligned}
&x_1 \triangleq \frac{\psi}{\varepsilon}, \quad x_2 \triangleq \frac{\dot\psi}{\varepsilon}, \quad x_3 \triangleq \frac{\theta}{\varepsilon}, \quad x_4 \triangleq \frac{\dot\theta}{\varepsilon}, \quad x_5 \triangleq \frac{\phi}{\varepsilon}, \quad x_6 \triangleq \frac{\dot\phi}{\varepsilon}, \\
&u_i \triangleq \frac{M_i}{\varepsilon M_0} \quad (i = 1,\ 2,\ 3), \qquad t \triangleq \frac{t_r}{t_m}, \qquad t_m \triangleq \sqrt{\frac{A}{M_0}}, \\
&c \triangleq t_m\bar\Omega, \qquad \mu \triangleq 2ce
\end{aligned}
\tag{II.5}
$$

where M_0 is an appropriate base quantity of the torque and ε is a nondimensional parameter introduced for convenience. A dot over a quantity denotes differentiation with respect to the normalized time t. By use of the quantities in Eqs. (II.5), Eqs. (II.4) are rewritten into

$$
\begin{aligned}
\dot{x}_1 &= x_2, \\
\dot{x}_2 &= (c + \mu \cos ct + \varepsilon x_6 + \varepsilon x_2 \sin \varepsilon x_3)x_4 \\
&\qquad + u_1 \cos \varepsilon x_3 + u_2 \sin \varepsilon x_1 \sin \varepsilon x_3 + u_3 \cos \varepsilon x_1 \sin \varepsilon x_3, \\
\dot{x}_3 &= x_4, \\
\dot{x}_4 &= -(c + \mu \cos ct + \varepsilon x_6)x_2 \cos \varepsilon x_3 + u_2 \cos \varepsilon x_1 - u_3 \sin \varepsilon x_1, \\
\dot{x}_5 &= x_6, \\
\dot{x}_6 &= (c + \mu \cos ct + \varepsilon x_6)x_4 \tan \varepsilon x_3 + \varepsilon x_2 x_4 \sec \varepsilon x_3 + \frac{\mu c}{\varepsilon} \sin ct \\
&\qquad + u_2 \sin \varepsilon x_1 \sec \varepsilon x_3 + u_3 \cos \varepsilon x_1 \sec \varepsilon x_3
\end{aligned}
\tag{II. 6}
$$

Some special types of approximate equations, which are easy to handle, are derived under the corresponding adequate conditions. The nonlinear terms, including those nonlinear in x_i and u_i, may be disregarded under the assumption that the deviation angles and their time rates are sufficiently small. Retaining only linear terms yields

$$
\begin{aligned}
\dot{x}_1 &= x_2, \\
\dot{x}_2 &= cx_4 + \mu \cos ct \cdot x_4 + u_1, \\
\dot{x}_3 &= x_4, \\
\dot{x}_4 &= -cx_2 - \mu \cos ct \cdot x_2 + u_2
\end{aligned}
\tag{II. 7}
$$

and

$$
\begin{aligned}
\dot{x}_5 &= x_6, \\
\dot{x}_6 &= \frac{\mu c}{\varepsilon} \sin ct + u_3
\end{aligned}
\tag{II. 8}
$$

The method developed in Section 3.4 is applicable to the periodic system of Eqs. (II.7).

If the orbit is circular, the angular velocity Ω of the orbital motion is constant, that is, $\mu = 0$. Equations (II.7) then are reduced to

$$
\begin{aligned}
\dot{x}_1 &= x_2, \\
\dot{x}_2 &= cx_4 + u_1, \\
\dot{x}_3 &= x_4, \\
\dot{x}_4 &= -cx_2 + u_2
\end{aligned}
\tag{II.9}
$$

When $c = 0$, the system of Eqs. (II.9) is decoupled into two subsystems, each of them being a double integral system. Hence the method of Section 3.3.3 for weakly coupled linear systems is applicable to Eqs. (II.9).

Under special requirement, it is desirable to retain higher-order terms in x_i and u_i. If the terms up to the second order are retained and the orbit is circular, we obtain

$$
\begin{aligned}
\dot{x}_1 &= x_2, \\
\dot{x}_2 &= cx_4 + \varepsilon x_4 x_6 + u_1 + \varepsilon x_3 u_3, \\
\dot{x}_3 &= x_4, \\
\dot{x}_4 &= -cx_2 - \varepsilon x_2 x_6 + u_2 - \varepsilon x_1 u_3, \\
\dot{x}_5 &= x_6, \\
\dot{x}_6 &= \varepsilon x_2 x_4 + \varepsilon x_3 x_4 + u_3 + \varepsilon x_1 u_2
\end{aligned}
\tag{II.10}
$$

The method in Section 3.2 works effectively for the system of Eqs. (II.10) with the controls appearing linearly.

II.2 Equations for Orbit Transfer between Neighboring Elliptic Orbits [32]

This item is concerned with the motion of a space vehicle, in an inversely square gravitational force field, which transfers between two neighboring low-eccentricity elliptic orbits. In order to describe dynamical behabior of a vehicle, we have to define a reference coordinate system. We may choose a rotating rectangular system whose origin revolves along the final orbit in the specified direction.

The distance r of the origin from the center of the earth is related to the true anomaly θ by

$$r = \frac{\ell}{1 + e \cos \theta} \tag{II.11}$$

where ℓ and e denote the semilatus rectum and the eccentricity of the final orbit, respectively.

Let x, y, and z be three components of the vehicle displacement in the reference frame: y is the component directed outward along the local vertical, x the component perpendicular to y in the final orbit plane, and z the component normal to the orbit plane and directed so as to form the coordinates a right-handed triad. In component forms, the equation of motion of the vehicle is given by

$$\begin{aligned} x'' &= -2\theta' y' + \theta'^2 x - \frac{\mu}{R^3} x + A_1, \\ y'' &= 2\theta' x' + \theta'^2 (r + y) - \frac{\mu}{R^3} (r + y) + A_2, \\ z'' &= -\frac{\mu}{R^3} z + A_3 \end{aligned} \tag{II.12}$$

where μ is the product of the universal gravitational constant and the mass of the earth, and R is defined by

$$R \triangleq \sqrt{x^2 + (r + y)^2 + z^2} \tag{II.13}$$

A_1, A_2, and A_3 are circumferential, radial, and normal components of the thrust acceleration of the vehicle, respectively.

The time variation of θ is approximately given by

$$\theta = \omega t_r + 2e \sin \omega t_r \tag{II.14}$$

where ω is the mean value of the angular velocity of the origin over one revolution along the orbit. The time origin is chosen as $t_r = 0$ when the origin passes through the perigee of the orbit. Under the assumption that e is not so large, terms of order higher than the first in e are discarded in Eq. (II.14).

For convenience, we introduce nondimensional quantities defined by

$$\begin{gathered} x_1 \triangleq \frac{x}{d}, \quad x_2 \triangleq \frac{\dot{x}}{d}, \quad x_3 \triangleq \frac{y}{d}, \quad x_4 \triangleq \frac{\dot{y}}{d}, \quad x_5 \triangleq \frac{z}{d}, \quad x_6 \triangleq \frac{\dot{z}}{d}, \\ u_i \triangleq \frac{A_i}{d\omega^2} \quad (i = 1, 2, 3), \qquad t \triangleq \omega t_r, \qquad \varepsilon \triangleq \frac{3d}{2\ell} \end{gathered} \tag{II.15}$$

where d is an appropriate base quantity of the distance.

Since the initial and target orbits are located close to each other, we may assume that x, y, and z are sufficiently small relative to r. Hence, by neglecting terms of order higher than the first in x/r, y/r, and z/r in Eqs. (II.12), we obtain, in terms of the quantities defined by Eqs. (II.15),

$$
\begin{aligned}
\dot{x}_1 &= x_2, \\
\dot{x}_2 &= -2x_4 + e(\cos t \cdot x_1 + 2 \sin t \cdot x_3 - 4 \cos t \cdot x_4) + u_1, \\
\dot{x}_3 &= x_4, \\
\dot{x}_4 &= 2x_2 + 3x_3 - 2e(\sin t \cdot x_1 - 2 \cos t \cdot x_2 - 5 \cos t \cdot x_3) + u_2
\end{aligned}
\tag{II. 16}
$$

and

$$
\begin{aligned}
\dot{x}_5 &= x_6, \\
\dot{x}_6 &= -x_5 - 3e \cos t \cdot x_5 + u_3
\end{aligned}
\tag{II. 17}
$$

Equations (II. 16) are mutually related; while Eqs. (II. 17) are independent of Eqs. (II. 16). In other words, the out-of-plane motion is decoupled from the in-plane motion in the linearized system.

This is not the case if higher-order terms in Eqs. (II. 12) are introduced. In fact, taking the terms up to the second-order in x/r, y/r, and z/r yields

$$
\begin{aligned}
\dot{x}_1 &= x_2, \\
\dot{x}_2 &= -2x_4 + 2\varepsilon x_1 x_3 + u_1, \\
\dot{x}_3 &= x_4, \\
\dot{x}_4 &= 2x_2 + 3x_3 + \varepsilon(x_1^2 - 2x_3^2 + x_5^2) + u_2, \\
\dot{x}_5 &= x_6, \\
\dot{x}_6 &= -x_5 + 2\varepsilon x_3 x_5 + u_3
\end{aligned}
\tag{II. 18}
$$

In Eqs. (II. 18), the orbit of the vehicle is assumed circular, i.e., $e = 0$, for simplicity.

The method of Section 3.4 is suitable for treating the system of

Eqs. (II.16) and (II.17). The system of Eqs. (II.18) with $\varepsilon = 0$ is decoupled into two linear subsystems: the systems of Eqs. (II.16) with $e = 0$ and of Eqs. (II.17) with $e = 0$. Hence the method of Section 3.3 for large-scale systems is applicable to Eqs. (II.18).

APPENDIX III

SUBOPTIMAL SOLUTION TO THE PROBLEM WITH A NONLINEAR OUTPUT EQUATION

The recurrence equations are summarized for a suboptimal control of the system with a nonlinear output equation. Consider the system

$$\dot{x} = A(t)x + \varepsilon f(x, t) + B(t)u \tag{III. 1}$$

with the output and the performance index

$$y = C(t)x + \varepsilon g(x, t) \tag{III. 2}$$

$$J = \frac{1}{2}y'(T)Fy(T) + \frac{1}{2}\int_{\tau}^{T} [y'Q(t)y + u'R(t)u]dt \tag{III. 3}$$

where $g(x, t)$ is the s-vector function analytic in x and continuous in $t \in [\tau, T]$ such that

$$g(0, t) = 0 \qquad \text{for any } t \in [\tau, T] \tag{III. 4}$$

All the other quantities in Eq. (III. 1) to (III. 3) are the same as specified in Section 2. 2.

The optimal control for the above problem is

$$u^* = - R^{-1}(t)B'(t)[P(t)x + \sum_{k=1}^{\infty} \varepsilon^k p^{(k)}(x, t)] \tag{III. 5}$$

The unknown n-vectors $p^{(k)}(x, t)$ satisfy a sequence of equations:

$$\frac{\partial p^{(k)}}{\partial t} + \frac{\partial}{\partial x}[x'S'p^{(k)}] = \frac{\partial \theta^{(k)}}{\partial x} \qquad (k = 1, 2, \ldots) \tag{III.6}$$

with the boundary conditions

$$\begin{aligned} p^{(1)}(x, T) &= \frac{\partial}{\partial x}[x'C'(T)Fg(x, T)], \\ p^{(2)}(x, T) &= \frac{1}{2}\frac{\partial}{\partial x}[g'(x, T)Fg(x, T)], \\ p^{(k)}(x, T) &= 0 \qquad (k = 3, 4, \ldots) \end{aligned} \tag{III.7}$$

The scalar functions $\theta^{(k)}(x, t)$ are defined by

$$\begin{aligned} \theta^{(1)}(x, t) &\triangleq - x'(Pf + C'Qg), \\ \theta^{(2)}(x, t) &\triangleq - f'p^{(1)} + \frac{1}{2}p^{(1)\prime}Ep^{(1)} - \frac{1}{2}g'Qg, \\ \theta^{(k)}(x, t) &\triangleq - f'p^{(k-1)} + \frac{1}{2}\sum_{i=1}^{k-1} p^{(i)\prime}Ep^{(k-i)} \qquad (k = 3, 4, \ldots) \end{aligned} \tag{III.8}$$

$P(t)$, $S(t)$, and $E(t)$ are given by Eqs. (2.14), (2.18), and (2.12), respectively.

The recursive solution of Eqs. (III.6) proceeds in much the same way as in Chapter 2.

APPENDIX IV

SUBOPTIMAL SOLUTION TO THE PROBLEM WITH A NONQUADRATIC PERFORMANCE INDEX

The recurrence equations are summarized for a suboptimal control of the system with a nonquadratic performance index. Consider the system

$$\dot{x} = A(t)x + \varepsilon f(x, t) + B(t)u \tag{IV.1}$$

with the output and the performance index

$$y = C(t)x \tag{IV.2}$$

$$J = \frac{1}{2}y'(T)Fy(T) + \varepsilon\phi[y(T)] + \frac{1}{2}\int_{\tau}^{T}[y'Q(t)y + 2\varepsilon\psi(y, t) + u'R(t)u]dt \tag{IV.3}$$

where $\phi(y)$ and $\psi(y, t)$ are the scalar functions analytic in y, continuous in $t \in [\tau, T]$, and nonnegative for any y and $t \in [\tau, T]$, such that

$$\phi(0) = 0, \qquad \psi(0, t) = 0 \qquad \text{for any } t \in [\tau, T] \tag{IV.4}$$

respectively. All the other quantities in Eq. (IV.1) to (IV.3) are the same as specified in Section 2.2.

The optimal control for the above problem is

$$u^* = -R^{-1}(t)B'(t)[P(t)x + \sum_{k=1}^{\infty}\varepsilon^k p^{(k)}(x, t)] \tag{IV.5}$$

The unknown n-vectors $p^{(k)}(x, t)$ satisfy a sequence of equations:

$$\frac{\partial p^{(k)}}{\partial t} + \frac{\partial}{\partial x}[x'S'p^{(k)}] = \frac{\partial \theta^{(k)}}{\partial x} \qquad (k = 1, 2, \ldots) \tag{IV.6}$$

with the boundary conditions

$$\begin{aligned} p^{(1)}(x, T) &= \frac{\partial}{\partial x}\phi[C(T)x], \\ p^{(k)}(x, T) &= 0 \qquad (k = 2, 3, \ldots) \end{aligned} \tag{IV.7}$$

The scalar functions $\theta^{(k)}(x, t)$ are defined by

$$\begin{aligned} \theta^{(1)}(x, t) &\triangleq - x'Pf - \psi(Cx, t), \\ \theta^{(k)}(x, t) &\triangleq - f'p^{(k-1)} + \frac{1}{2}\sum_{i=1}^{k-1} p^{(i)'}Ep^{(k-i)} \qquad (k = 2, 3, \ldots) \end{aligned} \tag{IV.8}$$

$P(t)$, $S(t)$, and $E(t)$ are given by Eqs. (2.14), (2.18), and (2.12), respectively.

The recursive solution of Eqs. (IV.6) proceeds in much the same way as in Chapter 2.

APPENDIX V

EXAMINATION OF TRANSIENT VARIATIONS OF A NASH EQUILIBRIUM SOLUTION

As seen in Section 5.5(a), a slight difference in the cost parameters results in a drastic change of the Nash equilibrium solution. This change might well be understood by examining transient variations of the solution. In this appendix, we will do it for the simple example of Section 5.5 with $\varepsilon = 0$ and $r_1 = 2$.

From Eqs. (5.40), the Nash equilibrium strategies of each player for $\varepsilon = 0$ are given by

$$u_i = -\frac{1}{2}P_i x \qquad (i = 1,\ 2) \tag{V.1}$$

where P_i satisfy

$$\begin{aligned} \dot{P}_1 &= -2P_1 - 1 + \frac{1}{2}P_1^2 - \frac{r_2}{4}P_2^2 + P_1P_2, \\ \dot{P}_2 &= -2P_2 - 1 + \frac{1}{2}P_2^2 - \frac{r_2}{4}P_1^2 + P_1P_2 \end{aligned} \tag{V.2}$$

and the boundary conditions

$$P_i(T) = F_i \geqq 0 \qquad (i = 1,\ 2) \tag{V.3}$$

Figure V.1 shows the schematic patterns of phase-plane diagram for Eqs. (V.2) with variation of r_2. The diagrams are symmetrical with respect to the straight line $P_1 - P_2 = 0$. The arrows indicate the direction of decreasing time. There exist some singular points in each

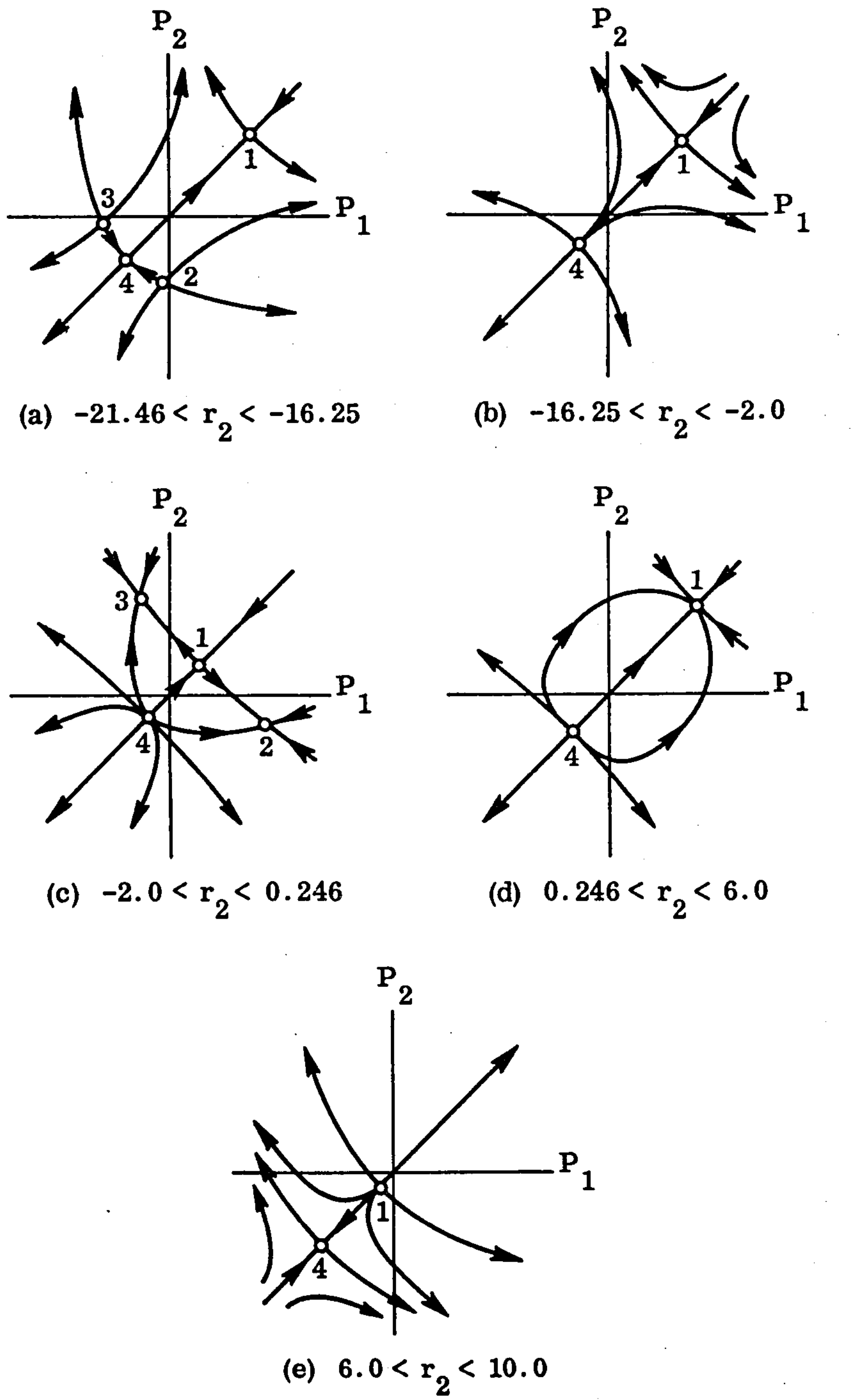


Fig. V. 1. Phase-plane diagrams for Eqs. (V.2) with various values of r_2.

diagram, and their classification is listed in Table V.1.

For $r_2 = -1$, chosen in Section 5.5(a), we readily see from Fig. V.1(c) that, if $F_1 = F_2$, a solution trajectory terminates at Point 1; while, if $F_1 \neq F_2$, it leads to Point 2 or 3 resulting a big difference between u_1 and u_2. Also note that, for some r_2, solutions diverge to infinity even with nonnegative F_i.

Table V.1 Classification of the Singular Points in Fig. V.1

Value of r_2	Singular point		
	1	2, 3	4
$r_2 < -21.46$	Saddle	Unstable focus	Saddle
$-21.46 < r_2 < -16.25$	Saddle	Unstable node	Saddle
$-16.25 < r_2 < -2.0$	Saddle	—	Unstable node
$-2.0 < r_2 < 0.246$	Saddle	Stable node	Unstable node
$0.246 < r_2 < 6.0$	Stable node	—	Unstable node
$6.0 < r_2 < 10.0$	Unstable node	—	Saddle
$10.0 < r_2$	—	—	—

REFERENCES

1. Al'brekht, E. G.: On the Optimal Stabilization of Nonlinear Systems, J. Applied Mathematics and Mechanics (PMM), vol. 25, pp. 1254-1266 (1961).
2. Aoki, M.: Control of Large-Scale Dynamic Systems by Aggregation, IEEE Trans. Automatic Control, vol. AC-13, pp. 246-253 (1968).
3. Athans, M., and P. L. Falb: "Optimal Control," McGraw-Hill Book Co., New York, 1966.
4. Athans, M., P. L. Falb, and R. T. Lacoss: Time-, Fuel-, and Energy-Optimal Control of Nonlinear Norm-Invariant Systems, IEEE Trans. Automatic Control, vol. AC-8, pp. 196-202 (1963).
5. Baldwin, J. F., and J. H. S. Williams: The Use of a Method of Perturbations in the Synthesis of Closed-Loop Optimal Control Laws for Non-Linear Systems, Automatica, vol. 5, pp. 357-367 (1969).
6. Bellman, R.: "Dynamic Programming," Princeton University Press, Princeton, 1957.
7. Bellman, R.: "Introduction to Matrix Analysis," McGraw-Hill Book Co., New York, 1960.
8. Bellman, R.: "Perturbation Techniques in Mathematics, Physics, and Engineering," Holt, Rinehart and Winston, Inc., New York, 1964.
9. Bogoliuboff, N. N., and Y. A. Mitropolsky: "Asymptotic Methods in the Theory of Nonlinear Oscillations," Gordon and Breach, Science Publishers, Inc., New York, 1961.
10. Box, G. E. P., and M. E. Muller: A Note on the Generation of Random Normal Deviates, Ann. Mathematical Statistics, vol. 29, pp. 610-611 (1958).
11. Burghart, J. H.: A Technique for Suboptimal Feedback Control of Nonlinear Systems, IEEE Trans. Automatic Control, vol. AC-14, pp. 530-533 (1969).
12. Coddington, E. A., and N. Levinson: "Theory of Ordinary Differential Equations," McGraw-Hill Book Co., New York, 1955.
13. Cole, J. D.: "Perturbation Methods in Applied Mathematics," Blaisdell Publishing Co., Waltham, 1968.

14. Craig, A. J., and I. Flügge-Lotz: Investigation of Optimal Control with a Minimum-Fuel Consumption Criterion for a Forth-Order Plant with Two Control Inputs; Synthesis of an Efficient Sub-Optimal Control, Preprints of 1964 Joint Automatic Control Conference, pp. 207-221, San Francisco, June 1964.

15. Da Cunha, N. O., and E. Polak: Constrained Minimization under Vector-Valued Criteria in Finite Dimensional Spaces, J. Mathematical Analysis and Applications, vol. 19, pp. 103-124 (1967).

16. D'Ans, G., P. V. Kokotivić, and D. Gottlieb: A Nonlinear Regulator Problems for a Model of Biological Waste Treatment, IEEE Trans. Automatic Control, vol. AC-16, pp. 341-388 (1971).

17. Doob, J. L.: "Stochastic Processes," John Wiley and Sons, New York, 1953.

18. Durbeck, R. C.: An Approximation Technique for Suboptimal Control, IEEE Trans. Automatic Control, vol. AC-10, pp. 144-149 (1965).

19. Edelbaum, T. N.: Optimum Low-Thrust Rendezvous and Station Keeping, AIAA J., vol. 2, pp. 1196-1201 (1964).

20. Eisenberg, B. R., and A. P. Sage: Closed Loop Optimization of Fixed Configuration Systems, International J. Control, vol. 3, pp. 183-194 (1966).

21. Eller, D. H., and J. K. Aggarwal: Sub-Optimal Control of Non-Linear Single Input Systems, International J. Control, vol. 8, pp. 113-121 (1968).

22. Faddeev, D. K., and V. N. Faddeeva: "Computational Methods of Linear Algebra," W. H. Freeman and Co., San Francisco, 1963.

23. Falb, P. L., and J. L. de Jong: "Some Successive Approximation Methods in Control and Oscillation Theory," Academic Press, New York, 1969.

24. Florentin, J. J.: Optimal Control of Continuous Time, Markov, Stochastic Systems, J. Electronics and Control, vol. 10, pp. 473-488 (1961).

25. Garrard, W. L.: Additional Results on Suboptimal Feedback Control of Non-Linear Systems, International J. Control, vol. 10, pp. 657-663 (1969).

26. Garrard, W. L.: Suboptimal Feedback Control for Nonlinear Systems, Automatica, vol. 8, pp. 219-221 (1972).

27. Garrard, W. L., N. H. McClamroch, and L. G. Clark: An Ap-

proach to Sub-Optimal Feedback Control of Non-Linear Systems, International J. Control, vol. 5, pp. 425-435 (1967).

28. Gobetz, F. W.: Optimal Variable-Thrust Transfer of a Power-Limited Rocket between Neighboring Circular Orbits, AIAA J., vol. 2, pp. 339-343 (1964).

29. Greenberg, S. G.: On Quadratic Optimization in Distributed Parameter Systems, IEEE Trans. Automatic Control, vol. AC-16, pp. 153-159 (1971).

30. Halme, A., and R. P. Hämäläinen: On the Nonlinear Regulator Problem, Control Engineering Laboratory Research Report, Tampere Univ. of Technology, Tampere, Finland, no. 1 (1973).

31. Hayashi, C.: "Nonlinear Oscillations in Physical Systems," McGraw-Hill Book Co., New York, 1964.

32. Hayashi, C., Y. Nishikawa, and N. Sannomiya: Optimal Variable-Thrust Transfer between Neighboring Elliptic Orbits, Proc. 3rd Congress of IFAC, vol. 1, pp. 39E.1-39E.8, London, June 1966.

33. Ho, Y. C., A. E. Bryson, Jr., and S. Baron: Differential Games and Optimal Pursuit-Evasion Strategies, IEEE Trans. Automatic Control, vol. AC-10, pp. 385-389 (1965).

34. Inoue, K., H. Akashi, K. Ogino, and Y. Sawaragi: Sensitivity Approaches to Optimization of Systems with Time Delay, Preprints of IFAC Kyoto Symp. on Systems Engineering Approach to Computer Control, pp. 286-291, Kyoto, Japan, Aug. 1970.

35. Issacs, R.: "Differential Games," John Wiley and Sons, New York, 1956.

36. Jazwinski, A. H.: "Stochastic Processes and Filtering Theory," Academic Press, New York, 1970.

37. Kalman, R. E.: Contributions to the Theory of Optimal Control, Boletin de la Sociedad Matematica Mexicana, vol. 5, pp. 102-119 (1960).

38. Kalman, R. E., and R. S. Bucy: New Results in Linear Filtering and Prediction Theory, Trans. ASME, J. Basic Engineering, vol. 83, pp. 95-108 (1961).

39. Kamel, A. A., and S. D. Hassan: A Perturbation Treatment for Optimal Slightly Nonlinear Systems with Linear Control and Quadratic Criteria, J. Optimization Theory and Applications, vol. 11, pp. 386-404 (1973).

40. Kirillova, F. M.: Contribution to the Problem on the Analytical Construction of Regulators, J. Applied Mathematics and Mechanics (PMM), vol. 25, pp. 644-653 (1961).

41. Kleinman, D. L., and M. Athans: The Design of Suboptimal Linear Time-Varying Systems, IEEE Trans. Automatic Control, vol. AC-13, pp. 150-159 (1968).

42. Kokotović, P. V., and J. B. Cruz, Jr.: An Approximation Theorem for Linear Optimal Regulators, J. Mathematical Analysis and Applications, vol. 27, pp. 249-252 (1969).

43. Kokotović, P. V., W. R. Perkins, J. B. Cruz, Jr., and G. D'Ans: ε-Coupling Method for Near-Optimum Design of Large-Scale Linear Systems, Proc. IEE, vol. 116, pp. 889-892 (1969).

44. Kokotović, P. V., and P. Sannuti: Singular Perturbation Method for Reducing the Model Order in Optimal Control Design, IEEE Trans. Automatic Control, vol. AC-13, pp. 377-384 (1968).

45. Kokotivić, P. V., and R. A. Yackel: Singular Perturbation of Linear Regulators: Basic Theorems, IEEE Trans. Automatic Control, vol. AC-17, pp. 29-37 (1972).

46. Kriechbaum, G. K. L., and E. Noges: Sub-Optimal Control of Non-Linear Dynamical Systems via Linear Approximation, International J. Control, vol. 13, pp. 1183-1195 (1971).

47. Kumar, K. S. P.: On the Optimum Stabilization of a Satellite, IEEE Trans. Aerospace and Electronic Systems, vol. AES-1, pp. 82-83 (1965).

48. Kushner, H. J.: Optimal Stochastic Control, IRE Trans. Automatic Control, vol. AC-7, pp. 120-122 (1962).

49. Kushner, H. J.: Some Problems and Some Recent Results in Stochastic Control, IEEE International Convention Record, Part 6, pp. 108-116, New York, Mar. 1965.

50. La Salle, J., and S. Lefschetz: "Stability by Liapunov's Direct Method with Applications," Academic Press, New York, 1961.

51. Lass, H.: "Vector and Tensor Analysis," McGraw-Hill Book Co., New York, 1950.

52. Levine, M. D.: Parameterized Feedback Control of Non-Linear Processes, International J. Control, vol. 3, pp. 39-49 (1966).

53. Longmuir, A. G., and E. V. Bohn: The Synthesis of Suboptimal Feedback Control Laws, IEEE Trans. Automatic Control, vol. AC-

12, pp. 755-758 (1967).

54. Luenberger, D. G.: Observing the State of a Linear System, IEEE Trans. Military Electronics, vol. MIL-8, pp. 74-80 (1964).

55. Luenberger, D. G.: Observers for Multivariable Systems, IEEE Trans. Automatic Control, vol. AC-11, pp. 190-197 (1966).

56. Lukes, D. L.: Optimal Regulation of Nonlinear Dynamical Systems, SIAM J. Control, vol. 7, pp. 75-100 (1969).

57. MacDuffee, C. C.: "The Theory of Matrices," Chelsea Publishing Co., New York, 1961.

58. Markus, L.: Controllability of Nonlinear Processes, SIAM J. Control, vol. 3, pp. 78-90 (1965).

59. Markus, L., and E. B. Lee: On the Existence of Optimal Controls, Trans. ASME, J. Basic Engineering, vol. 84, pp. 13-22 (1962).

60. McKinsey, J. C. C.: "Introduction to the Theory of Games," McGraw-Hill Book Co., New York, 1952.

61. Morse, P. M., and H. Feshbach: "Methods of Theoretical Physics," McGraw-Hill Book Co., New York, 1953.

62. Nayfeh, A. H.: "Perturbation Methods," John Wiley and Sons, New York, 1973.

63. Neuman, J. von, and O. Morgenstern: "Theory of Games and Economic Behavior," Princeton University Press, Princeton, 1944.

64. Nishikawa, Y., C. Hayashi, and N. Sannomiya: Fuel and Energy Minimization in Three Dimensional Attitude Control of an Orbiting Satellite, Proc. 1st IFAC Symp. on Automatic Control in Peaceful Uses of Space, pp. 287-298, Stavanger, Norway, June 1965.

65. Nishikawa, Y., H. Itakura, and H. Kitagawa: Nonzero-Sum Differential Games with Nonlinear Dynamics, Reports of Professional Group on Nonlinear Problems (Inst. Electronics and Communication Engineers of Japan), NLP71-18, Tokyo, Japan, Jan. 1972.

66. Nishikawa, Y., N. Sannomiya, and H. Itakura: Suboptimal Design of a Nonlinear Feedback System, Memoirs of Faculty of Engineering, Kyoto Univ., Kyoto, Japan, vol. XXXII, pp. 334-347 (1970).

67. Nishikawa, Y., N. Sannomiya, and H. Itakura: Suboptimal Nonlinear Feedback-Control Systems and Their Stability, Reports of Professional Group on Nonlinear Problems (Inst. Electronics and Communication Engineers of Japan), NLP70-15, Kyoto, Japan, Dec. 1970

(in Japanese).

68. Nishikawa, Y., N. Sannomiya, and H. Itakura: Suboptimal Feedback Control of a Nonlinear System, Trans. Soc. Instrument and Control Engineers, vol. 7, pp. 79-85 (1971) (in Japanese).

69. Nishikawa, Y., N. Sannomiya, and H. Itakura: Suboptimal Feedback Control of Linear Systems with Periodic Coefficients, Preprints of 10th Convention of Soc. Instrument and Control Engineers, pp. 69-75, Tokushima, Japan, Aug. 1971 (in Japanese).

70. Nishikawa, Y., N. Sannomiya, and H. Itakura: Suboptimal Feedback Control of Stochastic Nonlinear Systems, Preprints of 10th Convention of Soc. Instrument and Control Engineers, pp. 139-140, Tokushima, Japan, Aug. 1971 (in Japanese).

71. Nishikawa, Y., N. Sannomiya, and H. Itakura: A Method for Suboptimal Design of Nonlinear Feedback Systems, Automatica, vol. 7, pp. 703-712 (1971); also, in J. B. Cruz, Jr. (ed.): "System Sensitivity Analysis," pp. 317-326, Dowden, Hutchinson and Ross, Inc., Stroudsburg, 1973.

72. Nishikawa, Y., N. Sannomiya, and H. Itakura: Suboptimal Feedback Control of Linear Periodic Systems, Trans. Soc. Instrument and Control Engineers, vol. 8, pp. 73-77 (1972) (in Japanese).

73. Nishikawa, Y., N. Sannomiya, and H. Itakura: Suboptimal Feedback Control of Linear Periodic Systems, Reports of Professional Group on Nonlinear Problems: "Time-Varying Systems" (Inst. Electronics and Communication Engineers of Japan), NLP72-1, pp. 31-41, Tokyo, Japan, May 1972.

74. Nishikawa, Y., N. Sannomiya, H. Itakura, and T. Fukuda: Suboptimal Feedback Control of Weakly Coupled Systems, Preprints of Symp. on Systems Engineering (Soc. Instrument and Control Engineers), pp. 9-12, Tokyo, Japan, Sept. 1969 (in Japanese).

75. Nishikawa, Y., N. Sannomiya, H. Itakura, and T. Fukuda: Suboptimal Solution of Regulator Problems for Coupling Systems, Preprints of 12th Joint Convention of Automatic Control, pp. 155-156, Osaka, Japan, Nov. 1969 (in Japanese).

76. Nishikawa, Y., N. Sannomiya, H. Itakura, and T. Fukuda: Methods for Suboptimal Control of Large-Scale Systems, Preprints of IFAC Kyoto Symp. on Systems Engineering Approach to Computer Control, pp. 219-224, Kyoto, Japan, Aug. 1970.

77. Nishikawa, Y., N. Sannomiya, and T. Ojika: An Algorithm for In-

teraction Coordination in Multilevel Control of Nonlinear Systems, Trans. Soc. Instrument and Control Engineers, vol. 10, pp. 15-21 (1974) (in Japanese).

78. Pearson, J. D.: Approximation Methods in Optimal Control, I. Sub-Optimal Control, J. Electronics and Control, vol. 13, pp. 453-469 (1962).

79. Polak, E.: "Computational Methods in Optimization," Academic Press, New York, 1971.

80. Pontryagin, L. S., V. C. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko: "The Mathematical Theory of Optimal Processes," Interscience Publishers, New York, 1962.

81. Rekasius, Z. V.: Suboptimal Design of Intentionally Nonlinear Controllers, IEEE Trans. Automatic Control, vol. AC-9, pp. 380-386 (1964).

82. Rink, R. E., and R. R. Mohler: Completely Controllable Bilinear Systems, SIAM J. Control, vol. 6, pp. 477-486 (1968).

83. Sakawa, Y.: Solution of Linear Pursuit-Evasion Games, SIAM J. Control, vol. 8, pp. 100-112 (1970).

84. Salmon, D. M., and P. V. Kokotović: Design of Linear Feedback Controllers for Nonlinear Plants, IEEE Trans. Automatic Control, vol. AC-14, pp. 289-292 (1969).

85. Sannomiya, N.: A Method for Construction of Non-Linear Regulators by Minimax Algorithm, International J. Control, vol. 18, pp. 417-427 (1973).

86. Sannomiya, N., Y. Nishikawa, and H. Itakura: Suboptimal Design of Nonlinear Regulators by Perturbation Method and Parameter Optimization Technique, Trans. Soc. Instrument and Control Engineers, vol. 8, pp. 95-101 (1972) (in Japanese).

87. Sato, T., and A. Ichikawa: On a Decomposition Technique for Multilevel Control Systems, IEEE Trans. Automatic Control, vol. AC-12, pp. 457-458 (1967).

88. Sawaragi, Y., and T. Ono: Approximation Methods for Optimal Control of Nonlinear Systems with White Noise, Preprints of 10th Joint Convention on Automatic Control, pp. 83-84, Osaka, Japan, Nov. 1967 (in Japanese).

89. Sawaragi, Y., T. Ono, and Y. Matsui: A Study on the Sub-Optimal Control of Nonlinear Systems by the Method of Instantaneous Linear-

ization, J. Japan Association of Automatic Control Engineers, vol. 11, pp. 131-138 (1967) (in Japanese).

90. Starr, A. W., and Y. C. Ho: Nonzero-Sum Differential Games, J. Optimization Theory and Applications, vol. 3, pp. 184-206 (1969).

91. Stewart, B.: Satellite Attitude Control, Control, vol. 6, pp. 84-89 and pp. 145-148 (1962).

92. Sundararajan, N., and J. B. Cruz; A. H. Haddad and J. B. Cruz: ε-Coupling Method for Near-Optimum Design of Large-Scale Linear Systems, Proc. IEE, vol. 117, pp. 223-224 (1970).

93. Thau, F. E.: On the Inverse Optimum Control Problem for a Class of Nonlinear Autonomous Systems, IEEE Trans. Automatic Control, vol. AC-12, pp. 674-681 (1967).

94. Toda, M., and J. Nakahara: "Game Theory and Behavior Theory," Kyoritsu, Tokyo, 1968 (in Japanese).

95. Urabe, M.: An Existence Theorem for Multi-Point Boundary Value Problems, Funkcialaj Ekvacioj, vol. 9, pp. 43-60 (1966).

96. Werner, R. A., and J. B. Cruz, Jr.: Feedback Control Which Preserves Optimality for Systems with Unknown Parameters, IEEE Trans. Automatic Control, vol. AC-13, pp. 621-629 (1968).

97. Wismer, D. A. (ed.): "Optimization Methods for Large-Scale Systems, with Applications," McGraw-Hill Book Co., New York, 1971.