OPTIMAL CONTROL OF STOCHASTIC DIFFERENTIAL SYSTEMS WITH APPLICATIONS

BY

Hon-Wing Wong

Submitted to the Department of Electrical Engineering, University of Ottawa, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Department of Electrical Engineering
Faculty of Science and Engineering
University of Ottawa
Ottawa-Canada

TO

MY PARENTS

AND

MY WIFE

ABSTRACT

The subject of this thesis is the optimal control of stochastic differential equations. In chapter 1, optimality conditions related to this subject are briefly reveiwed. In chapter 2, two gradient algorithms are developed and a numerical comparison is made. An application in a satellite attitude control problem is also presented. In chapter 3, a Markov chain is introduced into the stochastic differential equation. Both the dynamic programming approach and the variational approach are examined, and certain optimality conditions are obtained. Examples with closed form solutions are presented which consist of a machine maintenance problem, an investor's problem, and a linear Gaussian quadratic problem.

ACKNOWLEDGEMENTS

The author expresses a deep gratitude to his supervisor Dr. N.U. Ahmed for his guidance and for his stimulating suggestions.

The author would also like to thank Dr. K.L. Teo, for his helpful and interesting discussions on the methods of computations; Dr. N.D. Georganas, Dr. K.F. Schenk, Dr. C.M. Deo, Dr. C. Lemyre, for their constructive comments.

Thanks are also due to Dr. R.G. Lyons for arranging a convenient vacation schedule for the purpose of this thesis; and also to my wife for her patient typing.

The postgraduate scholarship, awarded by the National Research Council of Canada, is gratefully acknowledged.

NOTATIONS FOR THE PROBLEM FORMULATIONS IN SECTIONS 2.1 AND 3.2

Section 2.1

I	Time interval, $I = (0,T)$
w	Standard Wiener process with values in R^d
U	Any closed and convex set in R^s
u	Measurable control with values in U
$\mathcal{U}$	Set of admissible controls
B	open bounded and connected set in R^s
∂B, $\bar{B}$	Boundary and closure of B respectively
Q	Cylinder defined by $B \times I$
S	Lateral boundary of Q defined by $\partial B \times I$
ξ	Dynamic state vector with values in B
π, q	Distribution and density of ξ at $t = 0$
τ	First exit time of ξ from Q
a_{ij}, b_i	Matrices as defined following (2.1.4)
$L_{\alpha,\beta}(Q)$, $L_\alpha(Q)$	Function spaces as defined after (2.1.6)
$V_2(Q)$, $W_2^{1,0}(Q)$, $W_2^{1,1}(Q)$	Function spaces as defined after (2.1.9)

Section 3.2

y	Markov chain
$\mathcal{N}$	State space for y consisting of a finite set of points $\{z_1, \ldots, z_N\}$ in R^m

P	Transition matrix of y, $P = \{p_{ij}\}$
Q	Differential transition matrix of y, $Q = \{q_{ij}\}$
I	Time interval, $I = (0,T)$
w	Standard Wiener process with values in R^d
U	Any closed convex set in R^s
u	Measurable control vector with values in U^N
$\mathcal{U}$	Set of admissible controls
B	Open bounded and connected set in R^n
$\partial B, \bar{B}$	Boundary and closure of B respectively
Q_T	Cylinder defined by $B \times I$
S	Lateral boundary of Q_T defined by $\partial B \times I$
ξ	Dynamic state vector with values in B
π, q	Distribution and density of ξ at $t = 0$
τ	First exit time of ξ from Q
A, A_{ij}, $B_i, \bar{B}_i$	Linear operators as defined in (3.2.4)—(3.2.6)
P_0	Stochastic optimization problem as defined in (3.2.1) and (3.2.2)
P_1	Deterministic optimization problem (equivalent to P_0) as defined in (3.2.7)
F	Vector of cost integrands as defined following (3.2.3)
$W_2^{1,0}(Q_T)$, $V_2(Q_T)$, $L_{q,r}(Q_T)$, $W_2^1(B)$, $\overset{0}{W}{}_2^1(B)$	Function spaces as defined following (3.2.9)

TABLE OF CONTENTS

TABLE OF CONTENTS

TABLE OF CONTENTS

INTRODUCTION

The subject of this thesis concerns the problem of finding optimal controls for stochastic differential systems. In practice, these optimal control problems arise when one attempts to control a system of ordinary differential equations which contains additive white noise.

In chapter 1, existing optimality conditions for obtaining optimal controls are briefly reviewed. In chapter 2, one of the optimality conditions is used to develop two gradient algorithms for computing the optimal control. As an application, a satellite attitude control problem is considered. In chapter 3, we consider a more general class of optimal control problems in which a Markov chain appears in the stochastic differential equation. Optimality conditions are derived and some applications, in the form of numerical and analytical examples, are presented.

Original work in this thesis include the analyses and computation related to the following: (i) gradient algorithms of section 2.2 [62]; (ii) satellite attitude control problem of section 2.3.1 [63]; (iii) three dimensional satellite attitude control problem of section 2.3.4 [61]; (iv) machine maintenance problem and the investor's problem, both of section 3.1; and (v) derivation of optimality condition for the optimal control problem which involves Markov chain, section 3.2 [7].

CHAPTER 1

A BRIEF REVIEW ON OPTIMAL CONTROL OF STOCHASTIC DIFFERENTIAL SYSTEMS

CHAPTER 1 A BRIEF REVIEW OF OPTIMAL CONTROL OF STOCHASTIC DIFFERENTIAL SYSTEMS

The most commonly used model in the study of stochastic optimal control is the stochastic differential equation (or Ito differential equation). To introduce the notations used in this chapter, we state a typical optimal control problem:

$$\begin{cases} dx(t) = f(t,x(t),u(t,x(t)))dt + g(t,x(t),u(t,x(t)))dw(t) \\ x(0) = x_0, \text{ with distribution } \pi \\ J(u) = E\{\int_0^T L(t,x(t),u(t,x(t)))dt + \psi(T)\} \equiv \min\limits_{u \in \mathcal{U}} \end{cases}$$

where x is an n-dimensional state vector, $t \in [0,T]$ denotes time; u denotes the control function with values in R^r; w is a d-dimensional standard Wiener process; f and g are matrix - valued functions. The problem is to find a control u in the admissible control class $\mathcal{U}$ so that a cost functional J is minimized. The cost functional J as shown above consists of the expected value of the sum of an integral cost, with cost integrand L, and a terminal cost ψ.

Although there are many variations of this typical optimal control problem, the above notations will be retained throughout this chapter. For brevity, most technical assumptions will not be mentioned. For more detail, the reader is referred to either the original works or the surveys conducted by Ahmed [1], Teo [58] and Fleming [19].

1.1 THEORETICAL DEVELOPMENTS

In this section we present a brief review of the developments in optimality conditions for control problems related to stochastic differential equations.

Stochastic versions of Pontryagin's maximum principle have been developed by many authors, among others, Kushner [31] - [36], Sworder [53] - [56], Davis and Varaiya [13], and Bismut [11], [10]. These 'random necessary conditions' involve random adjoint differential equations and random Hamiltonians, from which random controls are determined. In [31], Kushner considers the control of the system

$$dx(t) = f(t,x(t),u(t))dt + dz(t) \tag{1.1.1}$$

where z is not necessary a Wiener process. More recently, he has also developed random necessary conditions for the system

$$dx(t) = f(t,x(t),u(t))dt + g(t,x(t))dw(t) \tag{1.1.2}$$

where state constraints are imposed on x. Random necessary conditions usually lead to a stochastic two point boundary value problem, for which there is no known method of numerical solution. One exception is the random necessary conditions derived by Sworder [54] [56] for a class of linear systems of the form

$$\frac{dx}{dt} = A(t,y(t))x(t) + B(t,y(t))u(t,x(t),y(t)) \tag{1.1.3}$$

where A and B depend on a Markov chain $y(t)$. The cost function to be minimized is of the form

$$(1.1.4) \quad J(u) = E\{\int_0^T L(t,x(t),u(t,x(t),y(t)))dt\}.$$

In the special case where L is quadratic, i.e.,

$$(1.1.5) \quad L(t,x,u) = x'Rx + u'Su,$$

his random necessary conditions can be used to compute the optimal control which is given by,

$$u^*(t,x,i) = S^{-1}B(t,i)K(t,i)x$$

where $K(t,i)$ satisfies an ordinary differential equation. The above result is a special case of a more general problem considered in section 3.2 of this thesis.

More recently, Bismut [11], [10] has considered a more general system than (1.1.2),

$$(1.1.6) \quad dx(t) = f(t,x(t),u(t))dt + g(t,x(t),u(t))dw$$

In some special cases, his random necessary conditions can be reduced to those given by Kushner [32] and Davis and Varaiya [13]. He has also offered some interesting interpretations of his result in certain applied problems [12].

The major disadvantage of random necessary conditions is their unsuitability for computation purposes. As opposed to this 'probabilistic' approach, the 'analytic' approach has the advantage of computability. The analytic technique involves Ito's lemma and Bellman's dynamic programming method, the resulting optimality conditions usually consist of quasilinear parabolic partial differential equations. Some interesting results in this area are due to Fleming [15] - [20], Wonham [64] - [66], Rishel [49] - [50] and Ahmed and Teo [2] - [6]. The following is a brief account of their work.

Using dynamic programming technique, Fleming considers the following problem

$$(1.1.7)\quad \begin{cases} dx(t) = f(t,x(t),u(t,x(t)))dt + g(t,x(t))dw(t), \\ \qquad\qquad\qquad t \in [0,T], \\ J(u) = E \int_0^\tau L(t,x(t),u(t,x(t)))dt \end{cases}$$

where $u \in \mathcal{U}$ is smooth in t and x taking values in a convex compact set U, x takes values in a compact subset B of R^n, and τ is an exit time defined as the minimum of two quantities: the terminal time T, and the first time x reaches the boundary of B. Define the expected cost ϕ^u corresponding to a control u and an initial state ξ as

$$(1.1.8)\quad \phi^u(s,\xi) = E\{\int_s^\tau L(t,x(t),u(t,x(t)))dt | x(s) = \xi\}.$$

Under some smoothness and boundedness assumption of f, g and L, and a uniform parabolicity condition on g, the function ϕ^u satisfies the following linear parabolic partial differential equation with first boundary condition:

$$(1.1.9)\quad \begin{cases} \Lambda(\phi^u) + f^u\phi^u_\xi + L^u = 0, \ (s,\xi) \in (0,T) \times B \\ \phi^u = 0, \ (s,\xi) \in (I \times \partial B) \cup (\{T\} \times B) \equiv \Omega \end{cases}$$

where

$$(1.1.10)\begin{cases} \Lambda(h) = h_t + \sum_{i,j=1}^{n} \frac{1}{2}(gg')_{ij} h_{\xi_i} h_{\xi_j} \\ f^u \phi_\xi^u = \sum_{i=1}^{n} f_i(s,\xi,u(s,\xi)) h_{\xi_i} \\ L^u = L(s,\xi,u(s,\xi)). \end{cases}$$

Define the functions

$$(1.1.11)\begin{cases} z(s,\xi) = \operatorname*{Inf}_{u\varepsilon\mathcal{U}} \phi^u(s,\xi), \ (s,\xi) \ \varepsilon \ \Omega \\ H(s,\xi,\psi) = \min_{u\varepsilon U} \{L(s,\xi,u) + f(s,\xi,u)\psi\} \end{cases}$$

where $fz = \sum_{i=1}^{n} f_i z_i$. Fleming then shows that the function z satisfies the quasilinear parabolic equation

$$(1.1.12)\begin{cases} \Lambda(z) + H(s,x,z_x) = 0, \ (s,x) \ \varepsilon \ (0,T) \times R \\ z(s,x) = 0, \ (s,x) \ \varepsilon \ \Omega \end{cases}$$

It is noted that if jump parameters are present in the drift and diffusion coefficients, the equation (1.1.12) would become a <u>system</u> of non-linear parabolic equation (details in chapter 3 of this thesis).

In many applications, the dynamic states has to be measured in a noisy environment. Estimates of dynamic states can be obtained by filtering. However the use of state estimates, instead of their exact values, in the optimality conditions discussed so far would result in suboptimal performance in general. Therefore there is a need to jointly consider the filtering and control problem. For linear systems and linear measurement channels, however, Wonham has shown that filtering and control can be performed separately. The linear system dynamics and the measurement model are

$$(1.1.13) \begin{cases} dx(t) = [A(t)x(t) + b(t,u(t))]dt + B(t)dw_1(t) \\ x(0) = x_0 \end{cases}$$

$$(1.1.14) \begin{cases} dy(t) = F(t)x(t)dt + G(t)dw_2(t) \\ y(0) = 0 \end{cases}$$

The cost function is given by

$$(1.1.15) \quad J(u) = E\int_0^T L(t,x(t),u(t))dt.$$

Wonham [64] has shown that the optimal control of the problem (1.1.13) - (1.1.15) is identical to the optimal control of the problem (1.1.13) and (1.1.15) using the filtered estimates from the measurements in (1.1.14). Further the optimal control which uses only current measurements performs identically to the optimal control which uses all past information, thereby reducing greatly the storage requirements of the controller. This separation principle have not been demonstrated to hold for non-linear systems.

In some practical situations, only some of the current states can be observed. This problem has been formulated by Fleming [20], and the control problem is

$$(1.1.16)\quad \begin{cases} dx(t) = f(t,x,u(t,\tilde{x}))dt + g(t,x,u(t,\tilde{x}))dw \\ \\ J(u) = E\int_0^T L(t,x(t),u(t,\tilde{x}(t)))dt \end{cases}$$

where $\tilde{x}$ is the vector consisting of the observable components of x, and f, g, L are assumed smooth. The resulting necessary conditions lead to two coupled non-linear integro-partial differential equations whose numerical solution are by no means trivial. Fleming also has presented similar necessary conditions for the case with measurable f and L and without control in g [20].

In some optimal control problems, we may choose both controls and parameters. Ahmed and Teo [5] have considered the following problem

$$(1.1.17)\quad \begin{cases} dx(t) = f(t,x(t),\sigma,u(t,\tilde{x}(t)))dt + g(t,x(t,\sigma)dw(t) \\ \\ J(\sigma,u) = E\{\int_0^T L(t,x(t),\sigma,u(t,\tilde{x}(t)))dt + \eta(x(t))\} \end{cases}$$

where σ is a parameter vector. The resulting necessary conditions covers Fleming's result for (1.1.16).

The results on stochastic differential systems disscussed so far are based on the assumption that gg' is positive definite. In some applications, the system is degenerate, i.e., gg' is not positive definite. In this case two immediate problems arise. One is the question of existence of solutions to the dynamic programming equation. The other problem is to establish that the solution of the dynamic programming equation is equal to the value function. For the first problem Fleming [16] has demonstrated the existence of weak solution under certain conditions. As for the second problem, Rishel [50] has considered the following degenerate problem

$$(1.1.18)\quad \begin{cases} dx_1(t) = f^1(t,x_1(t),x_2(t))dt \\ dx_2(t) = f^2(t,x_1(t),x_2(t),u(t,x_1(t),x_2(t)))dt + g(t,x_2(t))dw \\ J(u) = E\{\int_0^T L(s,x_1(s),x_2(s),u(s,x_1(s),x_2(s)))ds \mid \\ \qquad\qquad x_1(0) = \xi, x_2(0) = \eta\}. \end{cases}$$

He has shown that under similar conditions as Fleming's, the weak solution equals the value function, thereby establishing the optimality conditions. In the sequel we return to the discussion of non-degenerate problems.

As mentioned earlier, Sworder [54] has considered the optimal control problem of linear ordinary differential equations with jump Markov disturbance. The corresponding non linear problem has later been studied by Rishel [50], who has considered the following problem

$$(1.1.19)\quad \begin{cases} \dfrac{dx}{dt} = f(t,x(t),y(t),u(t,x(t),y(t))) \\ J(u) = E \int_0^T f_0(s,x(s),y(s),u(s,x(s),y(s)))ds \end{cases}$$

where y(t) is the Markov jump process as considered by Sworder. For the above problem, he has established the weak solution of the dynamic programming equations, and he has shown that the weak solution is equal to the value function, thereby establishing the optimality conditions.

Recently Rishel [49] has presented necessary and sufficient conditions under weaker assumptions on the drift coefficient f, which is allowed to be discontinuous in both time and space. This permits the inclusion of discontinuous controls. His control problem is

$$(1.1.20)\quad \begin{cases} \dfrac{dx}{dt} = f(t,x(t),y(t),u(t,x(t),y(t))) \\ J(u) = E\{\psi(\tau_u,x(\tau_u)) \mid \xi(0) = s,\ y(0) = y^i\} \end{cases}$$

where τ_u is the first time the process x reaches a closed manifold $M \subset [0,\infty) \times R^n$. The necessary and sufficient condition for optimality of a control law is that the corresponding cost function satisfies the associated dynamic programming equation except on the set of discontinuities of f.

In chapter 3 of this thesis, the following control problem will be considered (Ahmed and Wong [7])

$$(1.1.21)\begin{cases} dx(t) = f(t,x(t),y(t),u(t,x(t),y(t)))dt + g(t,x(t),y(t))dw(t) \\ J(u) = E\int_0^{\tau} L(t,x(t),y(t),u(t,x(t),y(t)))dt \end{cases}$$

The existence of weak solutions of the corresponding dynamic programming equation is difficult to establish. However, a set of necessary conditions can be derived by a variational approach. These conditions consist of systems of parabolic equations whose weak solution is shown to exist under sufficiently general conditions.

1.2 NUMERICAL METHODS

While the theoretical developments in the optimal control problem of stochastic differential systems have been exciting and fruitful, the actual numerical methods for computing optimal controls have received less attention. It appears that problem (1.1.7), with completely observable controller, and problem (1.1.16), with partially observable controller, are two key problems. Of these two problems, only the completely observable problem is considered in this thesis; the partially observable control problem (1.1.16) is more difficult.

For the completely observable control problem (1.1.7), Fleming's review [19] has pointed out two approaches, one involving quasilinearization [21] of the dynamic programming equation resulting in exact solutions, and the other involves approximated solutions for small constant diffusion coefficients [22]. These and other developments are briefly mentioned ~~below~~. Our emphasis will be on the general problem (1.1.7) instead of particular cases, for which some results, mostly on linear quadratic problems, can be found in [29], [43], [44], [45].

If the diffusion coefficients are small constants, a method due to Fleming [22] can be used. In this case the deterministic problem is solved for all possible terminal states, and the resulting cost functions are used to approximate the cost function of the stochastic problem. The numerical procedures have been considered by Holland [27]. This method requires small noise intensities and sufficiently well behaved deterministic solution.

To obtain an optimal control without involving approximations, a direct solution of the dynamic programming equation may be attempted. However, this would require the numerical solution of a non-linear parabolic equation. Bellman [9] [21] suggested a quasilinearization method, whereby the non-linear parabolic equation is not solved directly. Instead the control is modified successively by the solutions of linear parabolic equations. The author is not aware of any numerical study of this approach.

Following a variation approach [4] [20] , Wong and Ahmed [62] [63] have used a variational inequality to obtain the gradient of the cost function with respect to the control. The resulting procedure involves repeated numerical solution of linear parabolic equations and successive modifications of the control by the gradient*.

Still another approach of solving (1.1.7) has recently been investigated by Teo, Reid and Boyd [59]. This approach also involves the repeated solution of the linear parabolic equations. At each iteration, the controls are modified in a neighborhood within the relevant state-time space. The modified control is computed by minimizing a Hamiltonian, which is obtained from the term to be minimized in the dynamic programming equation. For other interesting results, the reader may consult [46] and [47].

---

* The theoretical development and computation results constitute part of chapter 2 of this thesis.

For the problem of optimal control of stochastic differential systems with Markov jump parameters, computations have been attempted only for Linear Quadratic Gaussian Problem [28] [57], which can be reduced to a problem of solving ordinary differential equations*. In chapter 3 of this thesis, a gradient algorithm will be proposed for solving the general problem.

* A general result on the Linear Quadratic Gaussian problem is developed in section 3.2.3.

CHAPTER 2

OPTIMAL CONTROL OF STOCHASTIC DIFFERENTIAL SYSTEMS WITHOUT JUMP PARAMETERS

CHAPTER 2 OPTIMAL CONTROL OF STOCHASTIC DIFFERENTIAL SYSTEM WITHOUT JUMP PARAMETERS

Numerical solution of the optimal control problem (1.1.7) is usually computed by Bellman's dynamic programming method [36]. In this chapter we make use of the optimality conditions given by Fleming [20] and Ahmed and Teo [4] to formulate a gradient method. This gradient method has a significant advantage over the dynamic programming method: the gradient method generates directions of steepest descent for computing the optimal control.

The optimality conditions given in [4], [20] involve a parabolic equation, an adjoint equation and an inequality. Based on these results, we investigate, in section 2.2, two computational procedures to obtain the optimal control. The first method involves a two point boundary value problem which consists of two coupled non-linear parabolic equations. This method is computationally complex and will not be considered in detail. The main emphasis of this chapter is on the second method, which is a gradient-type technique. This gradient technique involves solving two uncoupled parabolic equations, and modifying the control successively by the gradient of the cost function.

A simplified version of this gradient technique is also considered. A comparison of these two gradient algorithms is made in a numerical example.

In section 2.3, we apply the gradient method to a two dimensional satellite attitude control problem. We then consider the satellite attitude control problem in three dimensions and obtain a closed form expression for the optimal control.

Markov chains, which do not appear in the optimal control problems considered in this chapter, will be included in the system dynamics in the next chapter.

2.1 PROBLEM FORMULATION AND OPTIMALITY CONDITIONS

In this section we formulate the problem of optimal control of stochastic differential equations. Known optimality conditions are presented.

Denote the time interval $(0,T)$ by I. Let $\{w(t): t \in I\}$ be the standard Wiener process with values in R^d and let U be any compact and convex set in R^s with piecewise differentiable boundary ∂U. Consider the stochastic differential equation in R^n given by

$$(2.1.1) \quad \begin{cases} d\xi = f(\xi,t,u)dt + g(\xi,t)dw \\ \xi(0) = \xi_0, \text{ with initial distribution } \pi, \end{cases}$$

where u is any measurable function on I with values in U. Let B be an open bounded and connected set in R^n with boundary ∂B. Denote the cylinder $B \times (0,T)$ by Q and the lateral boundary $\partial B \times (0,t)$ by S and let τ be the first time the process ξ hits the set $S \cup (B \times \{T\})$. Define $\mathcal{U}$ to be the class of admissible control laws consisting of functions u defined and measurable on $\bar{B} \times I$ with values in U. Let $f_0 : R^n \times I \times U \to R^+ = [0,\infty)$ so that for any fixed and arbitrary $u \in U$, f_0 is measurable in Q and so that for any fixed and arbitrary $(x,t) \in Q$, f_0 is continuous in u on U.

The problem is to find an optimal control law u^0 so that the cost functional

$$(2.1.2) \quad \bar{J}(u) = E\{\int_0^\tau f_0(\xi(t),t,u(t))dt\}$$

attains its minimum at $u(t) = u^0(\xi(t),t)$ subject to the dynamic constraints (2.1.1).

The optimization problem (2.1.1), (2.1.2) can be transformed into an equivalent deterministic problem involving a parabolic partial differential equation with first boundary conditions, which is now presented. As preparation, we define the conditional expectation

$$(2.1.3) \qquad \phi(x,s) = E\{\int_s^T f_0(\xi(t),t,u(\xi(t),t))dt \,|\, \xi(s) = x\}.$$

The convention of repeated indices will be used throughout this chapter. Using this convention, we define the operator $L(u)$ for $u \in \mathcal{U}$ by

$$(2.1.4) \qquad L(u)\phi = \phi_t + a_{ij}\phi_{x_i x_j} + b_i(u)\phi_{x_i},$$

where $a_{ij}(x,t) = \frac{1}{2}(gg')_{ij}\,(x,t)$ and $b_i(u)(x,t) = f_i(x,t,u(x,t))$ and " ' " denotes matrix transpose.

<u>Lemma 2.1.1</u>

Suppose for each [illegible] $\mathcal{U}$, the system (2.1.1) has a solution ξ. Then the optimization problem (2.1.1), (2.1.2) is equivalent to a deterministic problem given by

$$(2.1.5) \qquad \begin{cases} L(u)\phi + f_0(u) = 0, & (x,t) \in B \times [0,T) \\ \phi(x,t)|_S = 0 & \\ \phi(x,T) = 0, & x \in B \end{cases}$$

(2.1.6) $$\min_{u \in \mathcal{U}} \{J(u) = \int_B \phi(x,0)\ \pi(dx)\}$$

where $f_0(u) = f_0(x,t,u(x,t))$.

The proof follows from standard arguments involving Ito's lemma ([52], theorem 2.5).

When necessary, $\phi(u)$ will be used to denote the solution of (2.1.5) corresponding to the control $u \in \mathcal{U}$. Let $L_{\alpha,\beta}(Q)$, $1 \leq \alpha,\beta < \infty$, be the Banach space of scalar functions on Q with the finite norm

$$||z||_{\alpha,\beta,Q} = \{\int_I [\int_B |z|^\alpha dx\ ]^{\beta/\alpha}\ dt\}^{1/\beta}.$$

If $\alpha = \beta$, $L_\alpha(Q)$ will denote $L_{\alpha,\beta}(Q)$. Throughout this chapter, we assume the coefficients of (2.1.4) and the initial distribution π satisfy the following assumptions (i) - (iii):

(i) For each $i,j = 1,2,\ldots,n$, $a_{ij}: B \times I \to R$ satisfies the parabolicity conditions

$$\nu|\eta|^2 \leq a_{ij}(x,t)\eta_i\eta_j \leq \mu|\eta|^2$$

for all $\eta \in R^n$ uniformly with respect to $(x,t) \in Q$ for some positive constants ν and μ.

(ii) for each $i = 1,2,\ldots,m$, $\bar{b}_i(u) \equiv b_i(u) - \dfrac{\partial a_{ij}}{\partial x_j} : B \times I \to R$ and $f_0(u): B \times I \to R$ are dominated by functions $G \in L_{2q,2r}(Q)$

and $H \in L_{q_1,r_1}(Q)$ respectively and independently of the choice of $u \in \mathcal{U}$, where q, r, q_1, r_1 satisfy

$$
(2.1.7)\quad \begin{cases}
1/r + n/2q = 1, \\
q \in (n/2, \infty]\ ,\ r \in [1,\infty] \text{ for } n \geq 2, \\
q \in [1,\infty]\ ,\ r \in [1,2] \text{ for } n = 1, \\
\\
1/r_1 + n/2q_1 = 1 + n/4, \\
q_1 \in [2n/(n-2), 2],\ r_1 \in [1,2] \text{ for } n \geq 3, \\
q_1 \in (1,2]\ ,\ r_1 \in [1,2) \text{ for } n = 2, \\
q_1 \in [1,2]\ ,\ r_1 \in [1,4/3] \text{ for } n = 1.
\end{cases}
$$

(iii) q belongs to the Sobolev space $W_2^1(B)$.

The adjoint equation to (2.1.5) is defined as

$$
(2.1.8)\quad \begin{cases}
L^*(u)\psi = 0\ ,\ (x,t) \in B \times (0,T] \\
\psi|_S = 0\ , \\
\psi(x,0) = q(x),\ x \in B,\ \pi(E) = \int_E q(x)\, dx
\end{cases}
$$

where $u = u^0$ is the optimal control and L^* (the formal adjoint to L) is given by

$$
(2.1.9)\quad L^*(u)\psi \equiv -\psi_t + (a_{ij}\, \psi_{x_j} - \bar{b}_i(u)\psi)_{x_i}.
$$

The weak solutions of (2.1.5) and (2.1.8) are defined in the space $V_2(Q)$, which is the Banach space that consists of all elements of the Sobolev space $W_2^{1,0}(Q)$ on Q having a finite norm

$$|z|_Q = \operatorname*{ess\,sup}_{0 \leq t \leq T} ||z(x,t)||_{2,B} + ||z_x||_{2,Q}.$$

A weak solution (or solution, for simplicity) ϕ of the first boundary value problem (2.1.5) is defined as an element in $V_2(Q)$ such that $\phi|_S = 0$ and such that for all $\eta \in W_2^{1,1}(Q)$ with $\eta(x,0) = 0$, the following equality is satisfied:

$$(2.1.10) \quad \int_Q \{-\phi\eta_t - a_{ij}\phi_{x_i}\eta_{x_j} + \bar{b}_i(u)\phi_{x_i}\eta + f_0(u)\eta\} dx\, dt = 0.$$

Similarly a weak solution (or solution) ψ of the problem (2.1.8) is defined as an element in $V_2(Q)$ such that $\psi|_S = 0$ and such that for all $\gamma \in W_2^{1,1}(Q)$ with $\gamma(x,T) = 0$,

$$(2.1.11) \quad \int_Q \{\psi\gamma_t - a_{ij}\psi_{x_j}\gamma_{x_i} + \bar{b}_i(u)\psi\gamma_{x_i}\} dx\, dt = -\int_B \psi\gamma\, dx|_{t=0}.$$

Under the assumptions (i) – (iii), we quote a result in [37] (thm. 4.1, p.153) for the existence of solution to parabolic equations (2.1.5) and (2.1.8).

<u>Theorem 2.1.2</u>

Under the assumptions (i) – (iii) the systems (2.1.5) and (2.1.8) have (weak) solutions in $V_2(Q)$.

We proceed to present optimality conditions for the deterministic control problem (2.1.5) (2.1.6).

Let Γ be a map from a Banach space E_1 into a Banach space E_2 and let Γ_u denote the strong (linear) Gateaux derivative of Γ on E_1 evaluated at $u \in E_1$, i.e.,

$$(2.1.12) \qquad \lim_{\epsilon \to 0} \left|\left| \frac{\Gamma(u + \epsilon v) - \Gamma(u)}{\epsilon} - \Gamma_u(v) \right|\right|_{E_2} = 0$$

for any $v \in E_1$. $\Gamma_u(v)$ is the Gateaux differential of Γ at u in the direction $v \in E_1$. Let $\phi(u)$ be the solution of (2.1.5) corresponding to u. The coefficients of (2.1.5) are assumed to satisfy the following additional assumption:

(iv) The strong Gateaux differentials $b_{iu}(v)$, $f_{0u}(v)$ exist at every $u,v \in \mathcal{U}$ and are continuous in $v \in U$, further they have the same properties as b_i and f_0 respectively.

<u>Theorem 2.1.3</u>

Suppose the assumptions (i) - (iv) are satisfied. If $u^0 \in \mathcal{U}$ is optimal for the problem (2.1.5), (2.1.6), then there exists a solution ψ of the adjoint system (2.1.8) corresponding to u^0 such that for every $u \in \mathcal{U}$.

$$(2.1.13) \qquad \sum_{\ell=1}^{s} \int_Q [b_{iu_\ell}(x,t,u^0)\phi_{x_i}(u^0)(x,t) + f_{0u_\ell}(x,t,u^0)] \psi(x,t)(u_\ell - u_\ell^0)\,dx\,dt \geq 0.$$

Further for almost every $(x,t) \in Q$ and $u \in U$,

$$(2.1.14) \quad \sum_{\ell=1}^{s} \{[b_{iu_\ell}(x,t,u^0)\phi_{x_i}(u^0)(x,t) + f_{ou_\ell}(x,t,u^0)] \psi(x,t)(u_\ell - u_\ell^0)\} \geqslant 0.$$

The proof is given in [4],[7]. The inequalities (2.1.13) and (2.1.14) will be referred to as the optimality condition in the integral form and in the pointwise form respectively.

Although the above optimality conditions are known, studies on computational aspects and actual numerical applications have not appeared in literature. The following section addresses the implementation of the optimality conditions (2.1.13) and (2.1.14).

2.2 GRADIENT METHOD

The optimality conditions presented in the last section consist of a parabolic equation, an adjoint equation, and an inequality. Two versions of the inequality have been presented in theorem 2.1.3, one in a pointwise form and the other in an integral form. In this section we consider the procedures for implementing these optimality conditions. It is shown that the inequality in the pointwise form leads to a two point boundary value problem, which is difficult to solve numerically (subsection 2.2.1). The main objective of this section is to develop gradient methods from the inequality in the integral form. Two versions of the gradient method are considered. The first one involves solving two uncoupled parabolic equations and modifying the control successively by the gradient of the cost function (subsection 2.2.2). The second one is a modification of the first one, and requires solving only one linear parabolic equation (subsection 2.2.3). A numerical comparison of these two gradient algorithms is made in subsection 2.2.4.

2.2.1 COMPUTING OPTIMAL CONTROL THROUGH A TWO POINT BOUNDARY PROBLEM

Before presenting the gradient methods in the next two subsections, we consider the possibility of applying the pointwise optimality condition (2.1.14), which leads to a first boundary value problem involving coupled non-linear parabolic equations.

For convenience of notation we define the map $(x,t,u,v,w) \to G(x,t,u,v,w)$ from $B \times I \times U \times R^n \times R$ into U with components of G given by

$$G_k(x,t,u,v,w) = [b_{iu_k}(x,t,u)v_i + f_{0u_k}(x,t,u)]w \tag{2.2.1}$$

for $k = 1,2,\ldots,r$. The pointwise optimality condition (2.1.14) can then be written as

$$G(x,t,u^0,\phi_x(u^0),\psi(u^0)) \cdot (u - u^0) \geq 0 \text{ for all } u \in U,$$

where "$\cdot$" denotes vector multiplication $(W \cdot U = \sum_i W_i U_i)$.

Define $H(x,t,u,v,w) \equiv G(x,t,u,v,w) \cdot u$ and $M(x,t,v,w) = \min_{u \in U} H(x,t,u,v,w)$. Let $\mu : B \times I \times R^n \times R \to U$ such that

$$H(x,t,\mu(x,t,v,w),v,w) = M(x,t,v,w) \tag{2.2.2}$$

for all $(x,t,v,w) \in B \times I \times R^n \times R$. Suppose there is an optimal control to the (control) problem (2.1.5), (2.1.6), with $\phi^0(x,t)$ and $\psi^0(x,t)$ the corresponding responses to (2.1.5) and (2.1.8). Then it is clear that the optimal control is given by

$$u^0(x,t) = \mu(x,t,\phi_x^0(x,t), \psi^0(x,t)). \tag{2.2.3}$$

Using μ in the system equation (2.1.5) and (2.1.8) in place of u, we have the following coupled system of non-linear parabolic equations:

$$\begin{cases} L(\mu)\phi + f_0(\mu) = 0, & (x,t) \in B \times [0,T), \\ \phi(x,t)|_S = 0, & \\ \phi(x,T) = 0, & x \in B, \\ L^*(\mu)\psi = 0, & (x,t) \in B \times (0,T], \\ \psi(x,t)|_S = 0, & \\ \psi(x,0) = q(x), & x \in B, \ \pi(E) = \int_E q(x)dx, \end{cases} \tag{2.2.4}$$

where $\mu = \mu(x,t,\phi_x,\psi)$.

We now consider the existence of (weak) solution of the above system.

Theorem 2.2.1

Let the assumptions (i) - (iv) hold. Suppose U is closed and convex, the function H is continuous on U uniformly in $(x,t,v,w) \in B \times I \times R^n \times R$ and strictly convex in the variable $u \in U$, and H satisfies

(2.2.5) $\quad H_{u_i u_j} \beta_i . \beta_j \geq c|\beta|^2$ for every $(x,t,u,v,w) \in B \times I \times R^n \times R$; c = positive constant.

Then the non-linear two point boundary value problem (2.2.4) has a solution whenever the optimal control problem (2.1.5), (2.1.6) has a solution.

Proof:

Suppose the problem (2.1.5), (2.1.6) has a solution $u^0 \in \mathcal{U}$ so that $J(u^0) \leq J(u)$ for all $u \in \mathcal{U}$. Then by theorem 2.1.2, there exists ϕ^0 and $\psi^0 \in V_2(Q)$ that solve (2.1.5) and (2.1.8) with control u^0. Using μ as given by (2.2.2) we define

(2.2.6) $\quad \tilde{u}^0(x,t) = \mu(x,t,\phi^0_x(x,t),\psi^0(x,t)).$

By the definition of $\tilde{u}^0$, we have

(2.2.7) $\quad H(x,t,\tilde{u}^0(x,t),\phi^0_x(x,t),\psi^0(x,t)) \leq H(x,t,u,\phi^0_x(x,t),\psi^0(x,t))$

for all $u \in U$ and $(x,t) \in Q$. Further since u^0 is optimal, it satisfies the optimality condition

(2.2.8) $\quad H(x,t,u^0(x,t),\phi^0_x(x,t),\psi^0(x,t)) \leq H(x,t,u,\phi^0_x(x,t),\psi^0(x,t)),$

for all $u \in U$ and $(x,t) \in Q$. Under the given assumptions, H has a unique minimum on U for each fixed but arbitrary $(x,t,v,w) \in B \times I \times R^n \times R$. Therefore it follows from (2.2.7) and (2.2.8) that $\tilde{u}^0(x,t) = u^0(x,t)$ a.e. on Q. Consequently the system (2.2.4) with $\mu = \tilde{u}^0$ has ϕ^0 and ψ^0 as solutions. This completes the proof.

As a result of the above theorem, we have the following procedure for obtaining the optimal control for the problem (2.1.1), (2.1.2):

Obtain μ as in (2.2.2). Solve the two point boundary value problem (2.2.4) to obtain ϕ^0 and ψ^0. Set the optimal control $u^0(x,t) = \mu(x,t,\phi_x^0(x,t),\psi^0(x,t))$.

The numerically procedures of the above approach are too complex and will not be further persued. In the following we consider two gradient-type algorithms which are easier to implement.

2.2.2 COMPUTING OPTIMAL CONTROL THROUGH GRADIENT METHOD

In this subsection we show that the gradient of the cost function J with respect to the controls can be obtained from the optimality condition (2.1.13). Based on this gradient we present an iterative scheme to compute the optimal control. For preparation, we have the following theorem.

Theorem 2.2.2

Let $b_i^{(m)}$, $f_0^{(m)}$, $m = 1,2,\ldots,$ satisfy the assumptions that

$$\begin{cases} b_i^{(m)} \to b_i \text{ strongly in } L_{2q,2r}(Q), \\ f_0^{(m)} \to f_0 \text{ strongly in } L_{q_1,r_1}(Q), \end{cases}$$

and let a_{ij} , $b_i^{(m)}$, $f_0^{(m)}$, b_i, f_0 satisfy assumptions (i) and (ii) of section 2.1. Then the solution $\phi^{(m)}$ of

$$(2.2.9)\quad \begin{cases} \phi_t^{(m)} + a_{ij}\,\phi_{x_ix_j}^{(m)} + b_i^{(m)}\phi_{x_i}^{(m)} + f_0^{(m)} = 0, \ (x,t) \in B \times [0,T), \\ \phi^{(m)}(t,x)|_S = 0, \\ \phi^{(m)}(T,x) = 0, \ x \in B \end{cases}$$

converges strongly in $V_2(Q)$ to ϕ, which is the solution of

$$(2.2.10)\quad \begin{cases} \phi_t + a_{ij}\,\phi_{x_i x_j} + b_i\phi_{x_i} + f_0 = 0, \ (x,t) \in B \times [0,T), \\ \phi(t,x)|_S = 0, \\ \phi(T,x) = 0, \ x \in B. \end{cases}$$

Proof:

By the definitions of the weak solutions of (2.2.9) and (2.2.10), we have

$$(2.2.11)\quad \int_Q \{-\phi^{(m)}\eta_t - a_{ij}\,\phi^{(m)}_{x_i}\eta_{x_j} + \bar{b}_i^{(m)}\phi^{(m)}_{x_i}\eta + f_0^{(m)}\eta\}\,dx\,dt = 0$$

where $\bar{b}_i^{(m)} = b_i^{(m)} - \frac{\partial a_{ij}}{\partial x_j}$, and

$$(2.2.12)\quad \int_Q \{-\phi\eta_t - a_{ij}\phi_{x_i}\eta_{x_j} + \bar{b}_i\phi_{x_i}\eta + f_0\eta\}\,dx\,dt = 0, \ \bar{b}_i = b_i - \frac{\partial a_{ij}}{\partial x_j},$$

for every $\eta \in W_2^{1,1}(Q)$, $\eta(x,0) = 0$. Let $v^m = \phi^{(m)} - \phi$. Subtracting equality (2.2.12) from (2.2.11), we obtain

$$(2.2.13)\quad \int_Q \{-v^m\eta_t - a_{ij}\,v^m_{x_i}\eta_{x_j} + \bar{b}_i^{(m)}v^m_{x_i}\eta + F^m\eta\}\,dx\,dt = 0$$

where

$$F^m = (f_0^{(m)} - f_0) + (\bar{b}_i^{(m)} - \bar{b}_i)\phi_{x_i}$$

Since $\phi_{x_i} \varepsilon L_2(Q)$, therefore $F^m \varepsilon L_{q_2,r_2}(Q)$, where

$$q_2 \equiv \min(q_1, 2q/(1+q))$$

$$r_2 \equiv \min(r_1, 2r/(1+r)).$$

The integral equality (2.2.13) is a weak form [37, p.137] for the following system:

$$\begin{cases} v^m + a_{ij} v^m_{x_i x_j} + \bar{b}_i^{(m)} v^m_{x_i} + G_j^m v_{x_j} + F^m = 0, \ (x,t) \ \varepsilon \ B \times [0,T), \\ v^m(x,t)|_S = 0, \\ v^m(x,T) = 0, \ x \ \varepsilon \ B. \end{cases}$$

where v^m has an energy estimate ([37], lemma 2.1, p.139) given by

$$|v^m|_Q \leq C \ ||F^m||_{q_2,r_2,Q} \ .$$

But by assumption $F^m \to 0$ strongly in $L_{q_2,r_2}(Q)$.

Therefore

$$|v^m|_Q \to 0$$

or equivalently

$$\phi^{(m)} \to \phi \text{ strongly in } V_2(Q).$$

This completes the proof.

We also need the following preparation concerning the behaviour of solutions of (2.2.10) with integral average coefficients. Let $\{B^m\}$ be a sequence of open domains with smooth boundaries such that $B^m \subset B^{m-1} \subset \overline{B^{m-1}} \subset B$ and $\lim B^m \to B$. Let b_i^m, f_0^m be the integral average of b_i, f_0 respectively such that $b_i^m \to b_i$ and $f_0^m \to f$ strongly in $L_{2q,2r}(Q)$ and $L_{q_1,r_1}(Q)$ respectively. Then it is known that the problem

$$(2.2.14)\begin{cases} \phi_t^m + a_{ij}\phi_{x_ix_j}^m + \bar{b}_i^m\phi_{x_i} + f_0^m = 0, \forall(x,t) \in B^m \times [0,T), \\ \qquad \bar{b}_i^m = b_i^m - \dfrac{\partial a_{ij}}{\partial x_j}, \\ \phi^m|_{S^m} = 0, \ S^m = \partial B^m \times [0,T), \\ \phi^m(x,T) = 0, \ x \in B^m, \end{cases}$$

has a classical solution ϕ^m ([37], p.320). By theorem 2.2.2 $\phi^m \to \phi$, $\phi_{x_i}^m \to \phi_{x_i}$, $i = 1,\ldots,n$, strongly in $L_2(Q)$. Similarly, it can be shown that the corresponding adjoint problem

$$(2.2.15)\begin{cases} -\psi_t^m + (a_{ij}\psi_{x_j}^m - \bar{b}_i^m\psi^m)_{x_i} = 0, \ (x,t) \in B^m \times (0,T], \\ \psi^m|_{S^m} = 0, \ S^m = B^m \times (0,T], \\ \psi^m(x,0) = \psi_0^m(x), \ x \in B^m, \ \psi_0^m \text{ integral average of} \\ \quad \text{initial density } q(x), \end{cases}$$

has a classical solution ψ^m such that $\psi^m \to \psi$, $\psi^m_{x_i} \to \psi_{x_i}$, $i = 1,2,\ldots,n$, strongly in $L_2(Q)$. Let $\langle\cdot,\cdot\rangle$ denote inner product in $L_2(Q)$. For further preparation, we state a lemma from Aronson [8]:

Lemma 2.2.3

Suppose the assumptions (i) - (iii) from section 2.1 are satisfied. Then

$$\lim_{m\to\infty} \langle L(u)(\phi(u) - \phi(v)), \psi^m \rangle = -\int_B (\phi(u) - \phi(v))\psi dx|_{t=0}$$

for any $u, v \in \mathcal{U}$.

The gradient of the cost function is given in the following Lemma.

Lemma 2.2.4

Suppose that assumptions (i) - (iv) from section 2.1 are satisfied and U a convex compact set in R^s. Then the Gateaux differential $J_u(v)$ of J evaluated at $u \in \mathcal{U}$ along the direction $v \in \mathcal{U}$ exists and is given by

(2.2.16) $\quad J_u(v) = \int_Q G(x,t,u,\phi_x(u),\psi(u)) \cdot v \, dx \, dt.$

<u>Proof</u>:

Multiplying the first equation of (2.1.5) by $\eta \in W_2^{1,1}(Q)$ and integrating we obtain

$$\int_Q \{\phi_t(u)\eta + (a_{ij}\, \phi_{x_i}(u))_{x_j}\eta - \bar{b}_i(u)\, \phi_{x_i}(u)\eta + f_0(u)\eta\} dx\, dt = 0$$

After integration by parts,

$$(2.2.17) \quad \int_{Q_T} \{-\phi(u)\eta_t - a_{ij}\phi_{x_i}(u)\eta_{x_j} - \bar{b}_i(u)\phi_{x_i}(u)\eta + f_0(u)\eta\} = \int_B \phi(u)\eta dx\big|_{t=0}$$

Similarly the weak solution $\phi(u + \epsilon v)$ of (2.1.5) corresponding to the control $u + \epsilon v$ satisfies

$$(2.2.18) \quad \int_Q \{-\phi(u + \epsilon v)\eta_t - a_{ij}\phi_{x_i}(u + \epsilon v)\eta_{x_j} - \bar{b}_i(u + \epsilon v)\phi_{x_i}(u+ \epsilon v)\eta + f_0(u + \epsilon v)\eta\} dx\, dt = \int_B \phi(u + \epsilon v)\eta\, dx\big|_{t=0}.$$

Subtracting (2.2.18) from (2.2.17) and rearranging we obtain

$$< L(u)\, \phi(u) - L(u + \epsilon v)\phi(u + \epsilon v) + f_0(u) - f_0(u + \epsilon v),\, \eta> = 0.$$

Let $\eta = \psi^m$ where ψ^m is the solution of (2.2.15). Then we have

$$< L(u)(\phi(u) - \phi(u + \epsilon v)), \psi^m > + < (L(u) - L(u + \epsilon v))\phi(u + \epsilon v), \psi^m >$$
$$+ < f_0(u) - f_0(u + \epsilon v), \psi^m > = 0$$

By lemma 2.2.3,

$$\lim_{m\to\infty} < L(u)(\phi(u) - \phi(u + \epsilon v)), \psi^m > = -\int_B (\phi(u) - \phi(u + \epsilon v))\psi \, dx|_{t=0}$$

and we arrive at

$$(2.2.19) \quad -\int_B [\phi(u) - \phi(u + \epsilon v)]\psi \, dx|_{t = 0}$$
$$= < (L(u + \epsilon v) - L(u))\phi(u + \epsilon v) + f_0(u + \epsilon v) - f_0(u), \psi >.$$

Dividing both sides of the above equation by ϵ and letting $\epsilon \to 0$ we obtain the expression for the gradient

$$(2.2.20) \quad J_u(v) = \sum_{\ell=1}^{r} \int_Q [b_{iu}(u)\phi_{x_i}(u) + f_{ou_\ell}(u)]\psi \cdot v_\ell \, dx \, dt.$$

This completes the proof.

From the gradient of the cost function as given above, we can obtain an iterative scheme to compute the optimal control. This procedure is now considered.

Let $\mathcal{W}$ be the set of measurable functions mapping Q into R^s. Define the 'truncation function' $\chi : \mathcal{W} \to \mathcal{U}$ as follows. For any $(x,t) \in Q$ and $w \in \mathcal{W}$, if $w(x,t) \in U$ then define

$\chi(w)(x,t) = w(x,t)$. Otherwise define $\chi(w)(x,t)$ to be the point y in U s.t.

$$d(y, w(x,t)) = \min_{y' \varepsilon U} d(y', w(x,t))$$

where d is the Euclidean distance in R^s.

<u>Theorem 2.2.5</u>

Suppose assumptions (i) - (iv) in section 2.1 hold and U a convex compact set in R^s. Further assume the derivatives $f_{ou_\ell}(x,t,v)$, $b_{iu_\ell}(x,t,v)$, $(i = 1,\dots,n;\ \ell = 1,\dots,r)$, are continuous in $v \in U$ for $(x,t) \in Q$. Define $G(u^i) \equiv G(x,t,u^i, \phi_x(u^i),\psi(u^i))$. Then for any sequence of sufficiently small positive numbers $\{\rho^i\}$, the cost sequence $\{J(u^i)\}$ corresponding to the controls $\{u^i\}$ generated by

$$(2.2.21)\qquad u^{i+1} = \chi\left[u^i - \rho^i \frac{G(u^i)}{|G(u^i)|}\right],\ (|\cdot| \equiv \text{Euclidean norm}),$$

is monotonically decreasing (i.e. $J(u^{i+1}) \leq J(u^i)$) and $\{J(u^i)\}$ has a limit.

<u>Proof:</u>

By lemma 2.2.3, the Gateaux differential $J_u(v)$ exists for u, $v \varepsilon \mathcal{U}$. Therefore $J(u^{i+1})$ can be written as

$$J(u^{i+1}) = J(u^i) + J_{u^i}(u^{i+1} - u^i) + \eta(\rho^i),\ i = 1,2,\ldots, \tag{2.2.22}$$

where $\eta(\rho^i)$ denotes the remainder. We first show that $\eta(\rho^i)$ is of the order ρ^i, i.e.,

$$\lim_{\rho^i \to 0} \eta(\rho^i)/\rho^i = 0 \tag{2.2.23}$$

Using Lagrange formula we have

$$J(u^{i+1}) = J(u^i) + \int_0^1 J_{u^i + \theta(u^{i+1} - u^i)}(u^{i+1} - u^i)d\theta. \tag{2.2.24}$$

Comparing (2.2.22) and (2.2.24), the remainder $\eta(\rho^i)$ can be expressed as follows,

$$\eta(\rho^i) = \int_0^1 J_{u^i + \theta(u^{i+1} - u^i)}(u^{i+1} - u^i)d\theta - J_{u^i}(u^{i+1} - u^i).$$

Dividing by ρ^i and using the expression (2.2.16) for the gradient in the above equality we obtain

$$\frac{\eta(\rho^i)}{\rho^i} = \frac{1}{\rho^i}\{\int_0^1\int_Q G(u^i + \theta(u^{i+1} - u^i))\cdot(u^{i+1} - u^i)dx\,dt\,d\theta \tag{2.2.25}$$

$$- \int_Q G(u^i)\cdot(u^{i+1} - u^i)dx\,dt\}.$$

Define $w = [w_1,\ldots,w_s]^T$ such that $u^i - \rho^i w = \chi[u^i - \rho^i G(u^i)/|G(u^i)|]$. Using the expression (2.2.21) in the above equation we arrive at

$$(2.2.26)\quad \frac{\eta(\rho^i)}{\rho^i} = \int_0^1 \int_Q [G(u^i) - G(u^i - \theta\rho^i w)]\cdot w \, dx\, dt\, d\theta .$$

Let $\phi(u)$ and $\psi(u)$ be the weak solutions of (2.1.5) and (2.1.8) respectively corresponding to $u \in \mathcal{U}$ and denote $b_{ju_\ell}(x,t,v(x,t))$ by $b_{ju_\ell}(v)$. From the definition of G in (2.2.1), the right side of (2.2.26) can be expressed as the sum $I_1^i + I_2^i + I_3^i + I_4^i$ where

$$(2.2.27)\quad I_1^i = \sum_{\ell=1}^{r} \int_0^1 \int_Q b_{ju_\ell}(u^i-\theta\rho^i w)\psi(u^i-\theta\rho^i w)[\phi_{x_j}(u^i-\theta\rho^i w) - \phi_{x_j}(u^i)]\, w_\ell \, dx\, dt\, d\theta$$

$$(2.2.28)\quad I_2^i = \sum_{\ell=1}^{r} \int_0^1 \int_Q [b_{ju_\ell}(u^i-\theta\rho^i w)\psi(u^i-\theta\rho^i w) - b_{ju_\ell}(u^i)\psi(u^i)]\, \phi_{x_j}(u^i) v_\ell \, dx\, dt\, d\theta .$$

$$(2.2.29)\quad I_3^i = \sum_{\ell=1}^{r} \int_0^1 \int_Q f_{0u_\ell}(u^i-\theta\rho^i w)\, [\psi(u^i-\theta\rho^i w) - \psi(u^i)]\, v_\ell \, dx\, dt\, d\theta$$

$$(2.2.30)\quad I_4^i = \sum_{\ell=1}^{r} \int_0^1 \int_Q [f_{0u_\ell}(u^i-\theta\rho^i w) - f_{0u_\ell}(u^i)]\, \psi(u^i) v_\ell \, dx\, dt\, d\theta .$$

We now show that $\lim\limits_{\rho^i\to 0} I_k^i = 0,\ k = 1,\ldots,4.$ By assumption (iv) b_i is continuous in u on U, therefore by theorem 2.2.2, we have

$$(2.2.31)\quad \phi_{x_j}(u^i - \theta\rho^i w) \to \phi_{x_j}(u^i),\ j = 1,\dots,n,$$

$$(2.2.32)\quad \psi(u^i - \theta\rho^i w) \to \psi(u^i),$$

strongly in $L_2(Q)$ as $\rho^i \to 0$. We first demonstrate that

$$\lim_{\rho^i \to 0} I_1^i = 0.$$

Since

$$(2.2.33)\quad u^i - \theta\rho^i w \xrightarrow{\text{a.e.}} u_i,\ \text{on } Q \text{ as } \rho^i \to 0,$$

and $b_{ju_\ell}(v)$ is continuous in $v \in U$, we have

$$(2.2.34)\quad b_{ju_\ell}(u^i - \theta\rho^i w) \xrightarrow{\text{a.e.}} b_{ju_\ell}(u^i),\ \text{on } Q \text{ as } \rho^i \to 0.$$

Also since $\psi \in V_2(Q)$, it follows from an inequality in [37, p.75] that

$$(2.2.35)\quad ||\psi||_{\tilde q,\tilde r,Q} \leq C_1|\psi|_Q,$$

where $\tilde q, \tilde r$ are subject to the following conditions:

$$(2.2.36)\quad \begin{cases} n/2\tilde q + 1/\tilde r = n/4, & \\ \tilde q \in [2, 2n/(n-2)], & \tilde r \in [2,\infty] \text{ for } n > 2, \\ \tilde q \in [2,\infty), & \tilde r \in (2,\infty] \text{ for } n = 2, \\ \tilde q \in [2,\infty], & \tilde r \in [4,\infty] \text{ for } n = 1. \end{cases}$$

We choose $\tilde{q} = 2q/(q-1)$ and $\tilde{r} = 2r/(r-1)$ where q and r are as defined in (2.1.7). It is easily verified that $\tilde{q}$, $\tilde{r}$ as chosen above satisfy the conditions (2.2.36). With these choices and due to (2.2.35) we have $\psi \in L_{\tilde{q},\tilde{r}}(Q)$ and by assumption, $b_{ju_\ell} \in L_{2q,2r}(Q)$. Thus $b_{ju_\ell}\,\psi \in L_2(Q)$. Since by assumption b_{ju_ℓ} is dominated uniformly on Q by a function in $L_{2q,2r}(Q)$, it follows from (2.2.34) and dominated convergence theorem that

(2.2.37) $b_{ju_\ell}(u^i-\theta\rho^i w) \to b_{ju_\ell}(u^i)$ strongly in $L_{2q,2r}(Q)$ as $\rho^i \to 0$.

Similarly due to theorem 2.2.2 and (2.2.35), we have

(2.2.38) $\psi(u^i-\theta\rho^i w) \to \psi(u^i)$ strongly in $L_{\tilde{q},\tilde{r}}(Q)$ as $\rho^i \to 0$.

If sequences of functions $\{f_n\} \in L_{2q,2r}(Q)$ and $\{g_n\} \in L_{\tilde{q},\tilde{r}}(Q)$ converge strongly to f_0 and g_0 respectively, then it follows from the inequality

$$||f_ng_n-f_0g_0||_{2,Q} \leq ||f_n||_{2q,2r,Q}\,||g_n-g_0||_{\tilde{q},\tilde{r},Q} + ||g_0||_{\tilde{q},\tilde{r},Q}\cdot||f_n-f_0||_{2q,2r,Q}$$

that $f_ng_n \to f_0g_0$ strongly in $L_2(Q)$. This implies that

(2.2.39) $b_{ju_\ell}(u^i-\theta\rho^i w)\psi(u^i-\theta\rho^i w) \to b_{ju_\ell}(u^i)\psi(u^i)$, strongly in

$L_2(Q)$ as $\rho^j \to 0$.

Using (2.2.36), (2.2.39) and the fact that $w_\ell \in L_\infty(Q)$, we have $I_1^i \to 0$. Similarly, by using (2.2.39), $I_2^i \to 0$. We now show that $\lim_{\rho^i \to 0} I_3^i = 0$. By the continuity of $f_{0u_\ell}(v)$ in $v \in U$, we have

(2.2.40) $f_{0u_\ell}(u^i-\theta\rho^i w) \xrightarrow{a.e.} f_{0u_\ell}(u^i)$ on Q as $\rho^i \to 0$.

If $|\alpha_k|$ is a bounded sequence of numbers converging to zero, then it is clear that

(2.2.41) $||f_{0u_\ell}(u^i-\alpha_k w)||_{q_1,r_1,Q} \leq C$, C a constant independent of k.

By (2.2.40) and (2.2.41), it follows that [26 , thm. 13.44, p.207]

(2.2.42) $f_{0u_\ell}(u^i-\theta\rho^i w) \to f_{0u_\ell}(u^i)$ weakly in $L_{q_1,r_1}(Q)$ as $\rho^i \to 0$.

Define $\bar{q} = q_1/(q_1-1)$, $\bar{r} = r_1/(r_1-1)$. It is clear that the pair $(\bar{q},\bar{r})$ satisfies condition (2.2.36) and therefore, by (2.2.35), we have $\psi \in L_{\bar{q},\bar{r}}(Q)$ and

(2.2.43) $\psi(u^i-\theta\rho^i w) \to \psi(u^i)$ strongly in $L_{\bar{q},\bar{r}}(Q)$ as $\rho^i \to 0$.

From (2.2.42), (2.2.43) and the fact that $w_\ell \in L_\infty(Q)$, we have $\lim_{\rho^i\to 0} I_3^i = 0$. Also from (2.2.42), it follows that $\lim_{\rho^i\to 0} I_4^i = 0$.

Consequently,

$$\lim_{\rho^i\to 0} \eta(\rho^i)/\rho^i = I_1^i + I_2^i + I_3^i + I_4^i = 0,$$

or equivalently

(2.2.44) $\eta(\rho^i) = o(\rho^i)$.

Thus using (2.2.16), (2.2.21) and (2.2.44) into (2.2.22), we obtain

(2.2.45) $J(u^{i+1}) - J(u^i) = -\int_Q G(u^i)\cdot w(u^i)\,dx\,dt + o(\rho^i)$

where $w(u^i) \equiv u^i - \chi[u^i - \rho^i G(u^i)/|G(u^i)|]$. By the convexity of U, and by the piecewise differentiability of ∂U, it can be shown that if ρ^i is small, then the integral in (2.2.45) is positive and is bounded from below by $K\rho^i$ for some positive constant K. Hence for sufficiently small positive ρ^i, $J(u^{i+1}) \le J(u^i)$. Further $0 \le J(u^i) \le J(u^1)$ for $i = 1,2,\ldots$, therefore the cost sequence $\{J(u^i)\}$ has a limit. This completes the proof.

The following theorem shows that the limiting control as obtained in theorem 2.2.5 satisfy the necessary conditions in the integral form and in the pointwise form.

Let $u^{i+1}(x,t)$ be generated from u^i according to (2.2.21). $u^{i+1}(x,t)$ will be called 'a limiting control generated by (2.2.21)' if $u^{i+1}(x,t) = u^i(x,t)$ a.e. in Q for any $\rho^i > 0$.

<u>Theorem 2.2.6</u>

The limiting control generated by (2.2.21) satisfies necessary conditions (2.1.13) and (2.1.14).

Proof (by contradiction):

Let u' be a limiting control generated by (2.2.21), and let $m_{n+1}(\cdot)$ and $m_s(\cdot)$ denote Lebesgue measure in R^{n+1} and R^s respectively. Suppose there exists a set $Q_1 \subset Q$, with $m_{n+1}(Q_1) > 0$, such that for $(x,t) \in Q_1$,

$$(*) \quad G(x,t,u',\phi_x(u'),\psi(u')) \cdot (u - u') < 0$$

for every u belonging to some $U_1 \subset U$ where $m_s(U_1) > 0$. The set Q_1 can be partitioned into the following two subsets:

$$Q_A = \{(x,t) \in Q_1 | u'(x,t) \text{ belongs to the interior of } U\}$$

$$Q_B = \{(x,t) \in Q_1 | u'(x,t) \text{ belongs to the boundary of } U\}$$

Since Q_A is a subset of Q_1, $G(x,t,u',\phi_x(u'),\psi(u')) = 0$ for a.e. $(x,t) \in Q_A$; otherwise the generation of the new control $u' - \rho\left[G(u')/|G(u')|\right]$ by (2.2.21) would violate the assumption that u' is a limiting control. We also note that if $m_{n+1}(Q_A) > 0$, the observation that $G = 0$ for a.e. $(x,t) \in Q_A$ will contradict (*). Thus $m_{n+1}(Q_A) = 0$. Since $Q_1 = Q_A \cup Q_B$ and $m_{n+1}(Q_1) > 0$, therefore $m_{n+1}(Q_B) > 0$. We will show that '$m_{n+1}(Q_B) > 0$' leads to a contradiction.

Consider an arbitrary $(x_B,t_B) \in Q_B$ Let $u_B' = u'(x_B,t_B)$, and let u_B be an element of U such that (*) holds for $u = u_B$. By (*), the angle between the vectors $u_B - u_B'$ and $-G(x_B,t_B,u',\phi_x(u'),\psi(u'))$ is less than 90^o. Consequently, for a suitable ρ, there exists $u_B'' \in U$ such that

$$u_B'' = \chi\left[u'(x_B,t_B) - \frac{G(x_B,t_B,u',\phi_x(u'),\psi(u'))}{|G(x_B,t_B,u',\phi_x(u'),\psi(u''))|}\right] \neq u_B'.$$

Since this is true for every $(x_B,t_B) \in Q_B$ and since $m_{n+1}(Q_B) > 0$, therefore u' cannot be a limiting control. This completes the proof.

<u>Remark</u>:

To save computer time, the expression

$$(2.2.46)\quad u^{i+1} = \chi\left[u^i - \rho^i\, G(u^i)\right]$$

can be used instead of (2.2.21), avoiding the computation of $|G(u^i)|$. Also, the last two theorems do not show that any sequence generated by (2.2.21) will converge to an extremal control (one that satisfies (2.1.13) or (2.1.14)). Rather, they show that (2.2.21) can generate a sequence of controls with monotonically decreasing costs, and, if at any iteration, (2.2.21) does not generate an improved control for any step size, then we have obtained an extremal control. Without assuming any step size rule, these two theorems appear to be the best results obtainable.

Suboptimal Step Size

Let $\bar{\rho}^i = \text{diag}(\rho_1^i, \ldots, \rho_r^i)$. For computation purposes, consider the following scheme:

$$(2.2.47) \qquad u^{i+1} = \chi[u^i - \bar{\rho}^i G(u^i)].$$

Then we have

$$(2.2.48) \qquad J(u^{i+1}) - J(u^i) = -\int_Q G(u^i)[u^{i+1} - u^i]dx\, dt + \eta^1(\bar{\rho}^i),$$

where η^1 denotes the remaining terms. By the same argument as in theorem 2.2.5 the first term on the left side of (2.2.48) is negative. Therefore, ignoring η^1, we have $J(u^{i+1}) - J(u^i) \leq 0$ for small positive ρ_j^i, $j = 1,\ldots,s$, showing successive cost improvement.

Let $J_u^{(2)}(\tilde{u}) \equiv (J_u^{(2)}\tilde{u}, \tilde{u})$ be the second Gateaux differential of J with respect to u along the direction $\tilde{u} \in \mathcal{U}$. Assume $J_{u^i}^{(2)}(v)$ $(v \in \mathcal{U})$ exists and $u^{i+1} = u^i - \bar{\rho}^i G(u^i)$, we have

$$(2.2.49) \qquad J(u^{i+1}) - J(u^i) = -J_{u^i}(\bar{\rho}^i G(u^i)) + \tfrac{1}{2}J_{u^i}^{(2)}(\bar{\rho}^i G(u^i)).$$

To simplify the evaluation of the second differential in (2.2.49), assume

$$(2.2.50) \qquad \phi_x(u^i) = \phi_x(u^{i+1});\ \psi(u^i) = \psi(u^{i+1}) \text{ on } Q,$$

which is valid when $\bar{\rho}_k^i$ are sufficiently small or when u^i is close to the optimal control. Using (2.2.50), the second differential is given by

$$(2.2.51)\quad J_{u^i}^{(2)}(\bar{\rho}^i\, G(u^i)) = \int_Q \sum_{\ell=1}^{r} (\rho_\ell^i)^2\, [b_{ju_\ell u_\ell}(u^i)\phi_{x_j}(u^i) f_{0u_\ell u_\ell}(u^i)] \cdot [b_{ju_\ell}(u^i)\phi_{x_j}(u^i) f_{0u_\ell}(u^i)]^2 \psi^3(u^i)\,dx\,dt.$$

By (2.2.51), the matrix $\tilde{\rho}^i$ that minimizes the right side of (2.2.49) is given by $\tilde{\rho}^i = \mathrm{diag}(\tilde{\rho}_1^i,\ldots,\tilde{\rho}_r^i)$ where

$$(2.2.52)\quad \tilde{\rho}_\ell^i = \frac{\int_Q [b_{ju_\ell}(u^i)\phi_{x_j}(u^i) + f_{0u_\ell}(u^i)]^2 \psi^2(u^i)\,dx\,dt}{\int_Q [b_{ju_\ell}(u^i)\phi_{x_j}(u^i) + f_{0u_\ell}(u^i)]^2 [b_{j,u_\ell u_\ell}(u^i)\phi_{x_j}(u^i) + f_{0u_\ell u_\ell}(u^i)]\, \psi^3(u^i)\,dx\,dt}$$

For convenience, call $\tilde{\rho}^i$ a suboptimal step size. Since the assumption (2.2.50) may not be valid at each iteration, the matrix $\tilde{\rho}^i$ does not in general give the fastest convergence. However, in an example in the next subsection, it will be shown that (2.2.52) is more effective than arbitrary choices of step size.

The gradient method for solving (2.1.5), (2.1.6) is summarized below.

<u>Algorithm A</u>

Provide an initial guess u^1 for the optimal control and $\bar{\rho}^0$ for the step size. In the ith iteration, perform the following steps:

Step 1. Solve equations (2.1.5), (2.1.8) for $\phi(u^i)$ and $\psi(u^i)$.

Step 2. Compute $J(u^i)$. If $|J(u^i) - J(u^{i-1})| < \epsilon$ for some preassigned small positive number ϵ, no more iterations will be required.

Step 3. Either (i) determine the suboptimal step size $\tilde{\rho}^i$ by (2.2.52) and obtain u^{i+1} by (2.2.47) or (ii) set $\rho^i = \rho^{i-1}$ if $J(u^{i-1}) > J(u^i)$, otherwise set $\rho^i = \rho^{i-1}/c$, $(c > 1)$.

Step 4. Compute u^{i+1} from (2.2.46) and repeat Step 1 with the new control $u^{i+1} \varepsilon \mathcal{U}$.

Algorithm A requires, at each iteration, the numerical solution of the adjoint equation (2.1.8). In the next subsection, we present a simplified algorithm that does not require the numerical solution of the adjoint equation.

2.2.3 A SIMPLIFIED GRADIENT METHOD

Since the solution ψ of the adjoint equation is non-negative, it is possible to use this property of ψ to reduce the complexity of algorithm A.

Define the mapping $H : Q \times U \times R^n \to R^r$ with the ℓ-th component given by

$$(2.2.53)\quad H_\ell(x,t,u,w) = \sum_{j=1}^{n} b_{ju_\ell}(x,t,u)w_j + f_{0u_\ell}(x,t,u),$$

$$\ell = 1,\ldots,r.$$

For $v \in \mathcal{U}$, let $\phi(v)$ and $\psi(v)$ denote the solutions of (2.1.5) and (2.1.8) respectively corresponding to v, and for convenience of notation, we write

$$H(v) \equiv H(x,t,v,\phi_x(v)).$$

Theorem 2.2.7

Suppose the assumptions of theorem 2.2.5 hold. Then for any sequence of sufficiently small positive numbers $\{\rho^i\}$, the cost sequence $\{J(u^i)\}$ corresponding to the controls $\{u^i\}$ generated by

$$(2.2.54) \quad u^{i+1} = \chi\left[u^i - \rho^i \frac{H(u^i)}{|H(u^i)|}\right],$$

is monotonically decreasing (i.e. $J(u^{i+1}) \leq J(u^i)$) and has a limit.

Proof:

Write

$$(2.2.55) \quad J(u^{i+1}) = J(u^i) + J_{u^i}(u^{i+1} - u^i) + \bar{\eta}(\rho^i),$$

where it can be shown as in theorem 2.2.5 that the remainder $\bar{\eta}(\rho^i)$ is of the order $O(\rho^i)$. Substituting the expression (2.2.54) for the control in (2.2.55) we have, for sufficiently small ρ^i,

$$J(u^{i+1}) - J(u^i) = -J_{u^i}(\rho^i H(u^i)) + \bar{\eta}(\rho^i)$$

$$= \sum_{\ell=1}^{r} \sum_{j=1}^{n} \int_Q \rho^i [b_{ju_\ell}(u^i)\phi_{x_j}(u^i) + f_{0u_\ell}(u^i)]^2 \psi(u^i)\, dx\, dt$$

$$\leqslant 0,$$

for a small positive number ρ^i. The last step is due to the fact that $\psi(u^i)(x,t)$ is non-negative [8] for $(x,t) \in \bar{B} \times [0,T]$. Hence $J(u^{i+1}) - J(u^i) \leqslant 0$. Further $0 \leqslant J(u^{i+1}) \leqslant J(u^1)$, $(i = 1,2,\ldots)$ and therefore the cost sequence $\{J(u^{i+1})\}$ has a limit. This completes the proof.

Remark:

The expression (2.2.54) for generating successive controls does not require solving the adjoint equation (2.1.8). Compared to algorithm A, the above procedure is thus easier to implement. To save computer time we can use the expression

$$(2.2.56) \quad u^{i+1} = \chi[u^i - \gamma^i H(u^i)]$$

instead of (2.2.54), avoiding the computation of $|H(u^i)|$.

The above procedure for solving (2.1.5) and (2.1.6) is summarized below.

Algorithm B

Provide an initial guess $u^1 \varepsilon \mathcal{U}$ and a small positive number ρ^0. In the i-th iteration, perform the following steps:

Step 1. Solve (2.1.5) for $\phi(u^i)$.

Step 2. Compute $J(u^i) = \int_B \phi(u^i)(x,t)\ \pi(dx)$. If $|J(u^i) - J(u^{i-1})| < \epsilon$ for some preassigned small positive number ϵ, no more iterations will be requires.

Step 3. If $J(u^{i-1}) > J(u^i)$, set $\rho^{i+1} = \rho^i$. Otherwise set $\rho^{i+1} = \rho^i/C$, $(C > 1)$.

Step 4. Obtain u^{i+1} by (2.2.56). Repeat Step 1 with the new control $u^{i+1} \varepsilon \mathcal{U}$.

So far we have examined three possibilities of using the optimality conditions in theorem 2.1.3 to obtain optimal controls. The method of obtaining the optimal control by (2.1.14) in terms of ϕ_x and ψ leads to a two point boundary value problem, which is numerically complicated. Using (2.1.13), a gradient method (algorithm A) has been developed, along with a simplified version (algorithm B). Algorithm A with suboptimal step size has in general the advantage of faster convergence whereas algorithm B has the advantage of solving only one parabolic equation. In the next subsection, a numerical comparision between algorithms A and B is made.

2.2.4 NUMERICAL COMPARISON OF THE TWO GRADIENT METHODS

In this subsection, an optimal control problem* is solved by algorithms A and B separately. The convergence of the two algorithms will be compared. Consider the plant dynamics

$$(2.2.57)\quad \begin{bmatrix} d\xi_1 \\ d\xi_2 \end{bmatrix} = \begin{bmatrix} a\xi_2 + u_1 \\ -a\xi_1 + u_2 \end{bmatrix} dt + \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} dw(t), \quad t \in I = (0,.4)$$

where $\xi = (\xi_1,\xi_2)'$, $a \equiv .0154$, and $\sigma_1 = \sigma_2 = 10^{-2}$. The initial state $(\xi_1(0),\xi_2(0))'$ is a random vector with uniform distribution in the set $B = \{(x_1,x_2) : |x_1|,|x_2| < 1\}$. The optimal control problem is to find a control $u = (u_1(x,t),\ u_2(x,t))'$, $u \in U \triangleq [-1,1] \times [-1,1]$, that minimizes the cost functional

$$(2.2.58)\quad J(u) = E \int_0^T (\xi_1^2 + \xi_2^2 + u_1^2 + u_2^2)\ dt$$

---

* This sample problem arises from the attitude control of a spin stabilized satellite. The details are elaborated in the next section (2.3).

where τ is the first exit time of the process ξ from the set B.

Remark

Due to the control constraints, this problem cannot be solved by standard methods using Ricatti equations.

The equivalent deterministic minimization problem, from lemma 2.1.1, is

$$(2.2.59)\begin{cases} \phi_t + \tfrac{1}{2}\sigma_1^2\phi_{x_1x_1} + \tfrac{1}{2}\sigma_2^2\phi_{x_2x_2} + (ax_2 + u_1)\phi_{x_1} + (-ax_1 + u_2)\phi_{x_2} \\ \qquad + x_1^2 + x_2^2 + u_1^2 + u_2^2 = 0, \ (x,t) \in B \times [0,.4), \\ \phi(x,t) = 0, \ (x,t) \in \partial B \times (0,.4), \\ \phi(x,.4) = 0, \ x \in B. \end{cases}$$

$$(2.2.60) \quad \min_{u \in \mathcal{U}} \ \tfrac{1}{2} \int_B \phi(u)(x,0)\,dx$$

The corresponding adjoint system, from (2.1.8), is

$$(2.2.61)\begin{cases} -\psi_t + \tfrac{1}{2}\sigma_1^2\psi_{x_1x_1} + \tfrac{1}{2}\sigma_2^2\psi_{x_2x_2} - (ax_2 + u_1)\psi_{x_1} + (ax_1 - u_2)\psi_{x_2} \\ \qquad - (u_{1x_1} + u_{2x_2})\psi = 0, \ (x,t) \in B \times (0,.4], \\ \psi(x,t) = 0, \ (x,t) \in \partial B \times (0,.4), \\ \psi(x,0) = \tfrac{1}{2}, \ x \in B. \end{cases}$$

From the pointwise optimality condition (2.1.14), the optimal control attains its minimum at u^0, which can be shown to be

$$u_1^0 = \mathrm{Sat}(-\phi_{x_1}/4), \tag{2.2.62}$$

$$u_2^0 = \mathrm{Sat}(-\phi_{x_2}/4). \tag{2.2.63}$$

where

$$\mathrm{Sat}(y) \triangleq \begin{cases} 1 \text{ if } y > 1, \\ y \text{ if } |y| \leq 1, \\ -1 \text{ if } y < -1. \end{cases} \tag{2.2.64}$$

Using algorithm A, we have the following iterative formula from (2.2.47) to obtain the optimal control:

$$u^{i+1} = u^i - [\tilde{\rho}_1^i(\phi_{x_1}(u^i)+2u_1)\psi(u^i), \tilde{\rho}_2^i(\phi_{x_2}(u^i)+2u_2)\psi(u^i)]', \tag{2.2.65}$$

where the suboptimal step size $\tilde{\rho}^i$ obtained by (2.2.52) is given by

$$\begin{cases} \tilde{\rho}_1^i = \int_{B\times I}(\phi_{x_1}(u^i) + 2u_1)^2\psi^2(u^i)\,dx\,dt/2\int_{B\times I}(\phi_{x_i}(u^i)+2u_1)^2\psi^3(u^i)\,dx\,dt, \\ \tilde{\rho}_2^i = \int_{B\times I}(\phi_{x_2}(u^i)+2u_2)^2\psi^2(u^i)\,dx\,dt/2\int_{B\times I}(\phi_{x_2}(u^i)+2u_2)^2\psi^3(u^i)\,dx\,dt. \end{cases} \tag{2.2.66}$$

For comparision, we will also use 'fixed' step sizes as in algorithm A with step 3(ii).

Further, we will also use algorithm B, which avoids the computation of ψ. In this case the iterative formula (2.2.56) for the optimal control is

$$(2.2.67) \quad u^{i+1} = u^{i} - \rho^{i}[\phi_{x_1}(u^{i}) + 2u_1^{i}, \phi_{x_2}(u^{i}) + 2u_2^{i}]' .$$

The convergence of algorithm A with suboptimal step sizes (2.2.66) is shown in figure 1. The optimal control is obtained in 7 iterations, demonstrating fast convergence. In the same figure, convergence using 'fixed' step sizes is also shown; the algorithm converges relatively slowly. For example, from figure 1, when the 'fixed' step size is too large ($\rho^{i} = 50$) or too small ($\rho^{i} = .1$), the convergence is very slow.

The suboptimal step sizes as calculated from (2.2.66) is given in figure 2. From this graph it appears that if the fixed step sizes are close to 2, the convergence should be fast. Indeed, the 'fixed' step sizes with values 1 and 5 give rise to reasonably fast convergence (figure 1).

The convergence of algorithm B is shown in figure 3. For this algorithm, (2.2.61) is not solved and hence the suboptimal step sizes cannot be obtained as in algorithm A. The usual difficulty of choices of step sizes in gradient methods are evident. In this particular case convergence is in-

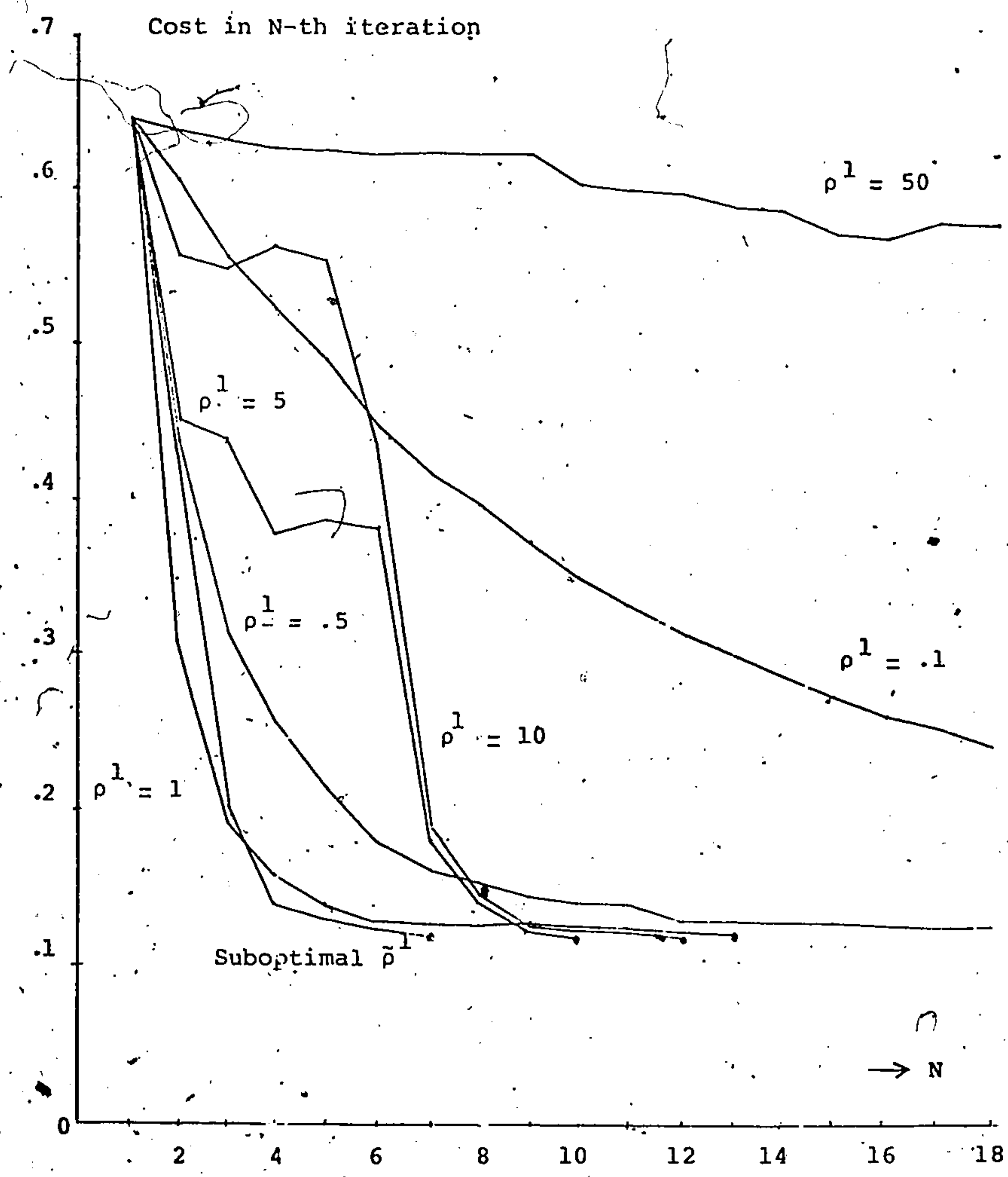


Figure 1. Rate of Convergence (algorithm A)
Scaling factor C = 2

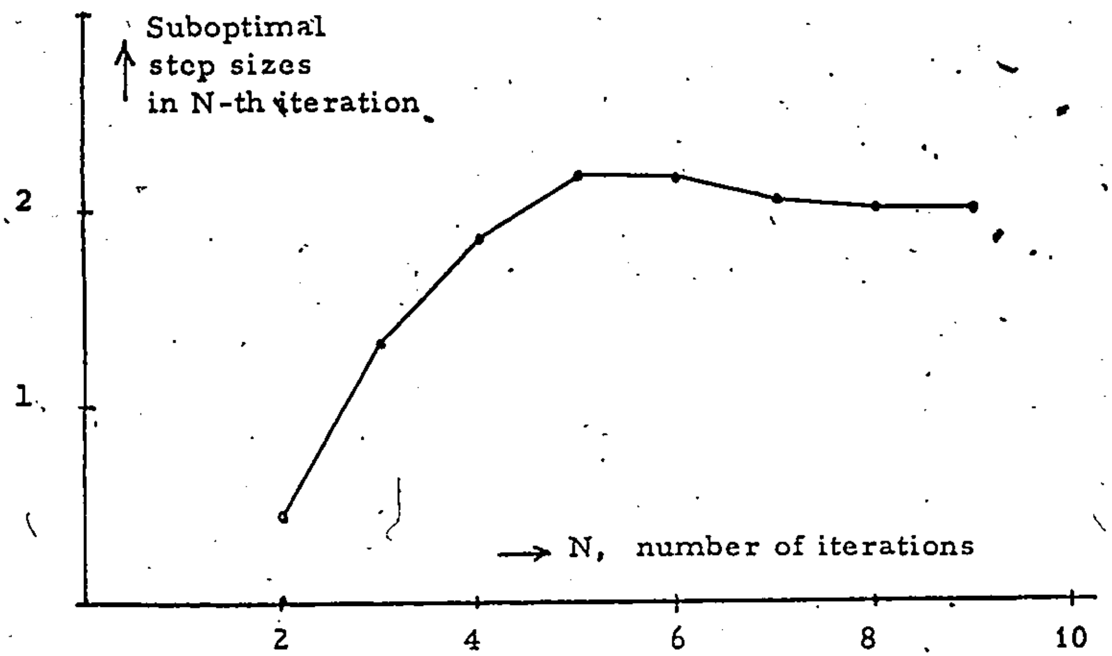


Figure 2 Suboptimal step sizes (algorithm A)

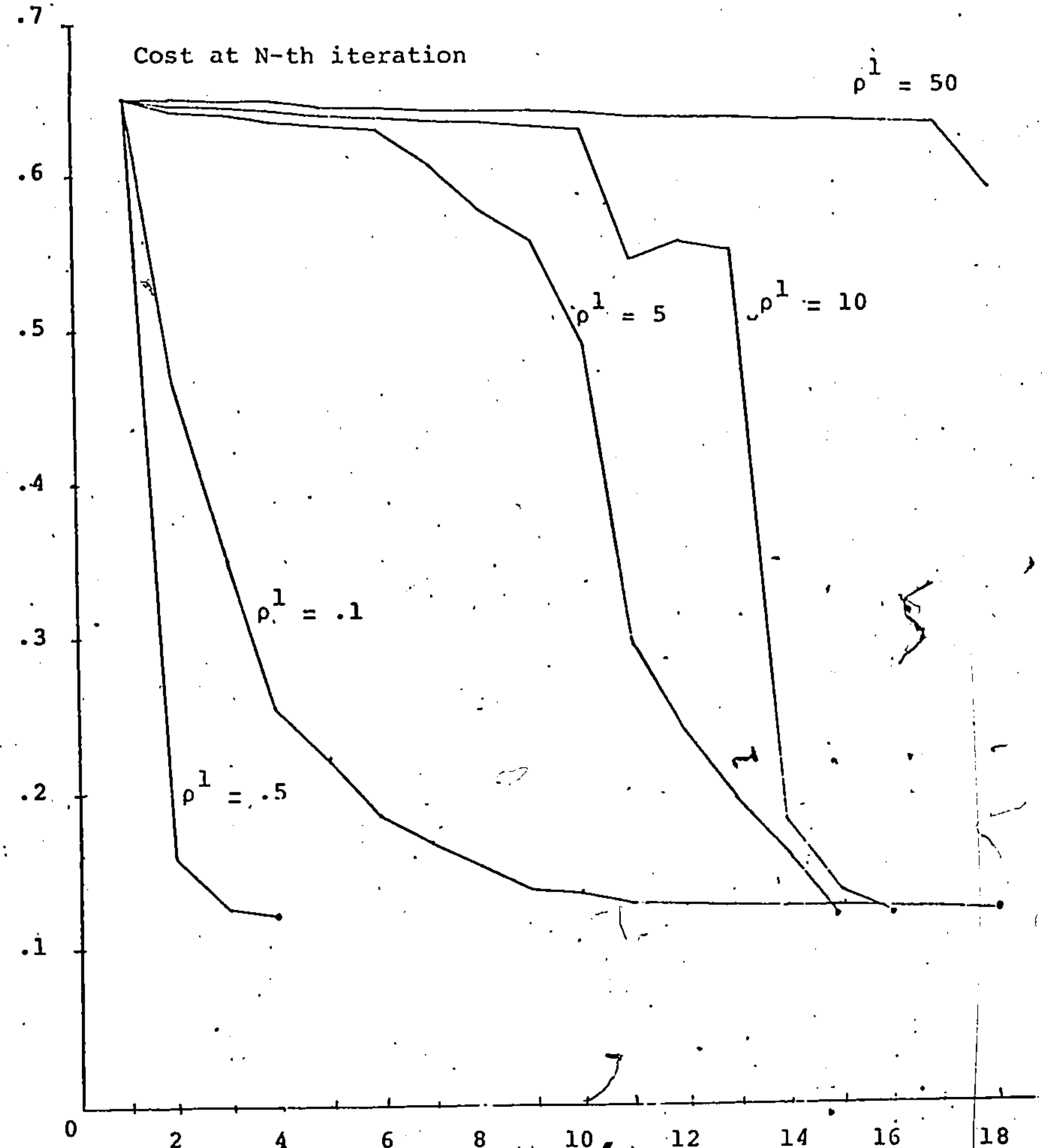


Figure 3. Rate of Convergence (algorithm B)
(Different values for the step size ρ^1 are used. Scaling factor C = 2.

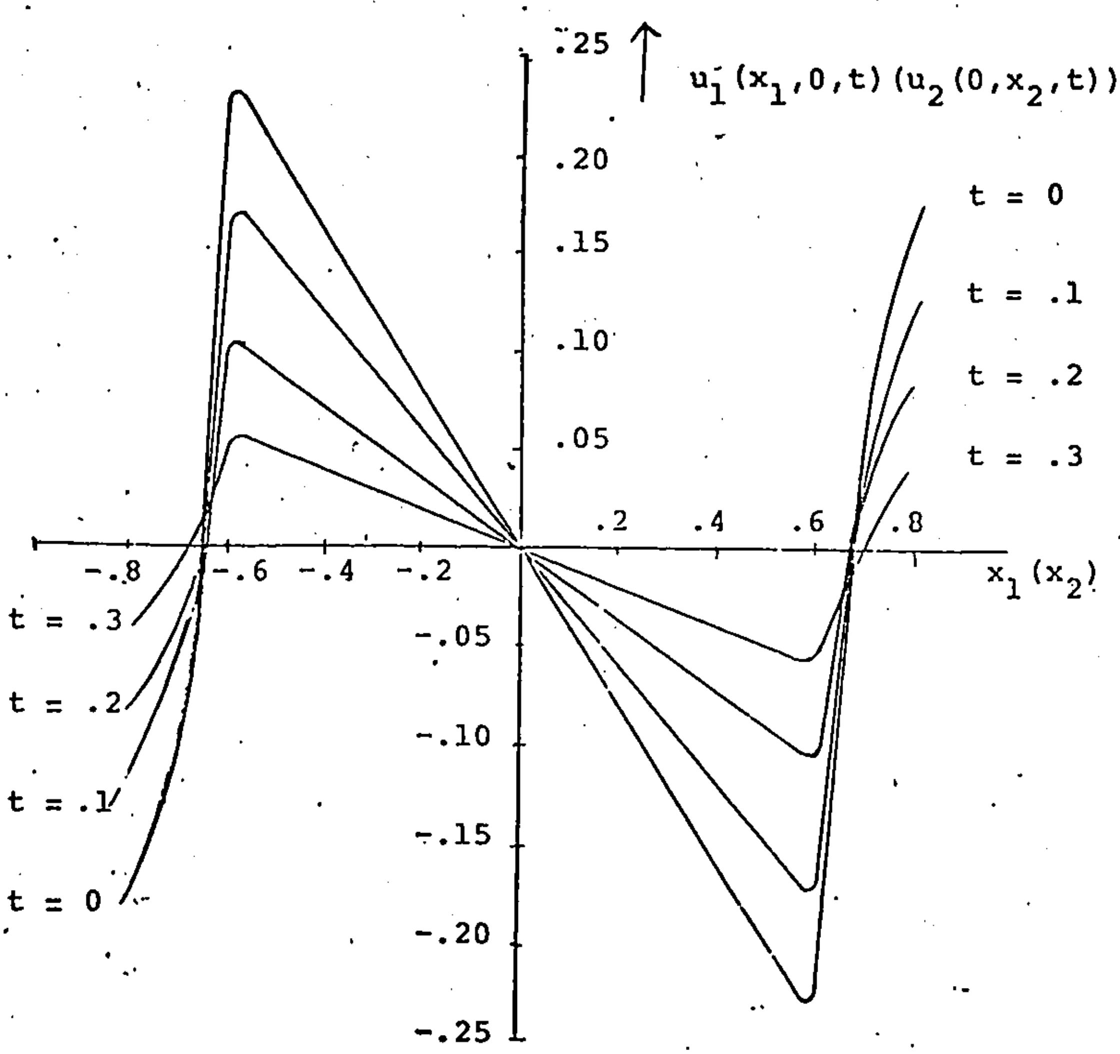


Figure 4. A Section of the Optimal Control

$t = 0$	$x_1 = -.667$	$x_1 = -.333$	$x_1 = 0$	$x_1 = .333$	$x_2 = .667$
$x_2 = .667$	.3483	.2185	.1776	.2204	.3483
$x_2 = .333$	.2204	.0843	.0421	.0843	.2185
$x_2 = 0$	.1776	.0421	.0000	.0421	.1776
$x_2 = -.333$	.2185	.0843	.0421	.0843	.2204
$x_2 = -.667$	.3483	.2204	.1776	.2185	.3485

TABLE 1. Expected Future Cost $\phi^0(x_1,x_2,t)$ at $t = 0$

$t = 0$	$x_1 = -.667$	$x_1 = -.333$	$x_1 = 0$	$x_1 = .333$	$x_2 = .667$
$x_2 = .667$	.2541	.2341	.2623	.2318	.2541
$x_2 = .333$	.2318	.2391	.2676	.2391	.2341
$x_2 = 0$	.2623	.2676	.2934	.2676	.2623
$x_2 = -.333$	.2341	.2391	.2676	.2391	.2318
$x_2 = -.667$	.2541	.2318	.2623	.2341	.2541

TABLE II. The Density $\psi^0(x_1,x_2,t)$ at $t = .4$

Time	Expected Future Cost	Confinement Probability
.0	.1212	1.0000
.1	.0916	0.9998
.2	.0612	0.9993
.3	.0306	0.9985
.4	0	0.9973

TABLE III. Expected Future Cost = $E\{\int_t^T f_0(\xi^0(\theta),\theta,u^0(\xi(\theta),\theta))d\theta\}$,
Confinement Probability = $\text{Prob}\{\xi^0(t) \in B\}$.

general slower when compared to algorithm A. When the step size is too large (50) or too small (.1), convergence is very slow.

A section of the optimal control is shown in figure 4. There is a region in B, beyond which the confinement of ξ in B is no longer economical and the controller abandons the confinement of states and pushes the system out of B in order to speed up the exit time. As will be shown in the next section, if a terminal cost is imposed into the problem, this effect will disappear.

In the linear quadratic problem without control constraints, it is known that the optimal control depends linearly on ϕ_x, and ϕ_x is linear in x. Consequently as long as the state is near the origin, so that the optimal control does not exceed the control constraints, the control is linear in x. This is evident in figure 4.

The system ϕ^0 of (2.2.59) corresponding to the optimal control u^0 is tabulated in table I for t = 0. The function ϕ^0 has been defined as

(2.2.68) $$\phi(x,t) \equiv E\{\int_t^T f_0(\xi(\theta),\theta,u^0(\xi(\theta),\theta))d\theta \mid \xi(t) = x\},$$

i.e., $\phi(x,t)$ is the conditional expected cost corresponding to the optimal control, given the information that the process ξ takes the value x at time t.

Let ξ^0 denote the process ξ corresponding to u^0. Then the solution ψ^0 of the adjoint system (2.2.6) corresponding to the optimal control u^0 is the density of ξ^0 i.e., ψ^0 satisfies

(2.2.69) $\int_E \psi^0(x,t)dx = \text{Prob}\{\xi^0(t) \in E\},\ E \subset B.$

The values for ψ^0 is tabulated in table II for $t = 0.4$.

From tables I and II, it is observed that there are certain symmetries in the functions ϕ^0 and ψ^0. In fact it is not difficult to show analytically that

(2.2.70) $\phi^0(x_1,x_2,t) = \phi^0(-x_1,-x_2,t) = \phi^0(-x_2,x_1,t),$

(2.2.71) $\psi^0(x_1,x_2,t) = \psi^0(-x_1,x_2,t) = \psi^0(-x_2,x_1,t),$

and

(2.2.72) $$\begin{aligned} &u_i^0(x_1,x_2,t) = -u_i^0(-x_1,-x_2,t) = u_i^0(x_2,-x_1,t),\ i = 1,2, \\ &u_i^0(x_1,x_2,t) = u_2^0(-x_2,x_1,t) \end{aligned}$$

for $(x_1,x_2,t) \in \bar{B} \times \bar{I}$.

The integral

(2.2.73) $\int_B \phi^0(x,t)\psi^0(x,t)dx = E\int_t^T f_0(\xi^0(\theta),\theta,u^0(\xi^0(\theta),\theta))d\theta$

represents the expected future cost at t corresponding to the optimal control. It provides an estimate of the future cost without the knowledge of the state $\xi(t)$. Table III shows that this expected future cost decreases with time as the probability of exit increases.

The confinment probability (2.2.69) with E = B gives the probability that the process ξ^0 is confined in B at time t. The confinment probability gives an idea of the flow of the process ξ^0 out of the desirable domain B. The way the confinment probability decreases with time is shown in Table III.

2.3 APPLICATION IN SATELLITE ATTITUDE CONTROL PROBLEM

In this section we first consider an application of the gradient method from subsection 2.2.2 to the satellite attitude control problem. Numerical results are presented. Then in the final subsection 2.3.4, the control problem is reformulated. Using calculus of variations, and assuming small noise intensities, we obtain a closed form solution to the satellite attitude control problem.

Satellite attitude control refers to the control of the angular velocity components about the body axes of the satellite. Consider a satellite which is stationary with respect to an observer on earth (such a satellite is called geo-stationary). Due to solar pressure and meteorite bombardment, an external torque is created, forcing the satellite to wobble. This wobbling (or precession, or nutation) interferes with the operation of the satellite. In order to minimize the amount of wobbling, the satellite may be equipped with gas jets which provide reaction torques to counteract the external torque.

The notations used in this section are listed below:

Notations

I_x, I_y, I_z	Moment of inertia of satellite in the x, y, and z axes of the satellite body frame.
p, q, r	Components of the angular momentum vector of the satellite in the x,y,z axes, respectively.
ω_0	A constant, $\omega_0 = 2\pi$ rad per 24 hrs.
C	Moment of inertia of the flywheel mounted in the y axis.
Ω	Angular velocity of the flywheel in the y axis.
T_x, T_y, T_z	Components of the disturbance torque in the x,y, and z axes respectively.
u_1, u_2	Jet controls in the x and z axes respectively.
σ_1^2, σ_2^2	Variances of T_x/I_x, T_z/I_z which are modeled as Gaussian white noises.
$I = [0,T]$	Time interval for firing control jets.
B	Tolerable set of satellite momenta (p,r).
u_{ip}, u_{ir}	Partial derivatives of u_i with respect to p and r, respectively.

2.3.1 PROBLEM FORMULATION

The dynamics of the angular momenta of the satellite subject to random disturbance torques T_x, T_y and T_z is modeled as follows [25]:

$$
\begin{aligned}
(2.3.1)\qquad & I_x\dot{p} + (q-\omega_0)r(I_z-I_y) = C\Omega r + T_x \\
& I_y\dot{q} + pr(I_x-I_z) = T_y \\
& I_z\dot{r} + (q-\omega_0)p(I_y-I_x) = -C\Omega p + T_z
\end{aligned}
$$

or equivalently,

$$
\begin{aligned}
(2.3.2)\qquad & \dot{p} = a_1 r + a_2 qr + n_1 \\
& \dot{q} = b_2 pr + n_2 \\
& \dot{r} = c_1 p + c_2 pq + n_3
\end{aligned}
$$

where

$$
\begin{aligned}
& a_1 = [\omega_0(I_z-I_y) + C\Omega]/I_x \\
& a_2 = (I_y-I_z)/I_x,\ b_2 = (I_z-I_x)/I_y \\
& c_1 = [\omega_0(I_y-I_x) - C\Omega]/I_z \\
& c_2 = (I_x-I_y)/I_z
\end{aligned}
$$

and $n_1 = T_x/I_x$, $n_2 = T_y/I_y$, $n_3 = T_z/I_z$ are independent white noise components with variances σ_1^2, σ_2^2, σ_3^2 respectively.

The system (2.3.2) can be written in the standard notation of stochastic differential equations:

$$\text{(2.3.3)} \qquad \begin{aligned} dp &= (a_1+a_2q)r\,dt + \sigma_1 dw_1 \\ dq &= b_2pr\,dt + \sigma_2 dw_2 \\ dr &= (c_1+c_2q)p\,dt + \sigma_3 dw_3 \end{aligned}$$

where w_1, w_2 and w_3 are independent standard (unit variance) Wiener processes.

For a satellite that is symmetric with respect to the y axis, $b_2 = 0$ $(I_x = I_z)$. Assuming that the satellite is spin stabilized along the y axis, the spin rate q is large and is expected to remain constant. (If σ_2 is large, this hypothesis does not hold.) The fractional mean square deviation of q from q_0 (the initial spin) is given by

$$E[(q(t) - q_0)/q_0]^2 = E\left[\int_0^t (\sigma_2/q_0)\,dw_2(\theta)\right]^2 = (\sigma_2/q_0)^2 t.$$

This quantity is negligible if σ_2 is sufficently small. For example, if $\sigma_2 = 10^{-3}$ and $q_0 = 100$ rad/s, the above mean square deviation is about 0.01 at the end of 3 years. In this situation, the system (2.3.3) can be approximated by a two-dimensional system:

$$\text{(2.3.4)} \qquad \begin{aligned} dp &= ar + \sigma_1 dw_1 \\ dr &= -ap + \sigma_3 dw_3 \end{aligned}$$

where $a \triangleq a_1 + a_2q_0 = -(c_1+c_2q_0)$

For a spin-stabilized satellite, precession is a common problem and requires control. When jet control is used as controls, the system (2.3.4) can be modified as

$$\text{(2.3.5)} \qquad \begin{aligned} dp &= ardt + u_1 dt + \sigma_1 dw_1 \\ dr &= apdt + u_2 dt + \sigma_3 dw_3 \end{aligned}$$

where u_1 and u_2 are the controls proportional to the torques generated by the jets placed in the x and z axes of the satellite, respectively.

Let $B \equiv \{(p,r): p_1 \leq p \leq p_2,\ r_1 \leq r \leq r_2\}$ be the set of tolerable angular momenta and π_0 the distribution of the initial state $(p(0), r(0))$ on B. Let T denote the first time the process (p,r) hits the boundary ∂B of B and let $[0,T]$ be a fixed time interval during which controls are applied. The problem is to find a feedback control law $u^0(t,p,r) = [u_1^0(t,p,r),\ u_2^0(t,p,r)] \triangleq u(t)$ so that

$$\text{(2.3.6)} \qquad \begin{aligned} J(u) = E\{\textstyle\int_0^{T\wedge\tau} [\alpha_1 p^2(t) + \alpha_2 r^2(t) + \beta_1 u_1^2(t) \\ + \beta_2 u_2^2(t)\ dt + L(p(T\wedge\tau),\ r(T\wedge\tau))\} \end{aligned}$$

is minimum under the constraints $|u_1(t)| \leq C_1$, $|u_2(t)| \leq C_2$, $\alpha_1, \alpha_2, \beta_1, \beta_2 > 0$. Denote the set of admissible controls by $\mathcal{U}$. Minimizing J is equivalent to minimizing the fuel cost and the deviations of the satellite from the zero state.

2.3.2 APPLICATION OF GRADIENT METHOD

By lemma 2.1.1, the stochastic optimization problem (2.3.5), (2.3.6) can be converted into an equivalent deterministic optimization problem. The resulting deterministic problem involves a parabolic partial differential equation, with first boundary conditions, specifically,

$$(2.3.7)\quad \begin{cases} \phi_t + \tfrac{1}{2}\sigma_1^2\phi_{pp} + \tfrac{1}{2}\sigma_3^2\phi_{rr} + (u_1+ar)\phi_p + (u_2-ap)\phi_r \\ + \alpha_1 p^2 + \alpha_2 r^2 + \beta_1 u_1^2 + \beta_2 u_2^2 = 0, \\ \text{for } (t,p,r) \in [0,T] \times B \\ \phi(t,p,r) = L(p,r), \text{ for } (p,r) \in \partial B \\ \phi(T,p,r) = 0, \text{ for } (p,r) \in B \end{cases}$$

$$(2.3.8)\quad \min_{u\in\mathcal{U}} \int_\Omega \phi(0,p,r)\pi_0(p,r)\,dp\,dr$$

where

$$(2.3.9)\quad \phi(t,p,r) \triangleq E\{\int_t^{\tau\wedge T}[\alpha_1 p^2(t) + \alpha_2 r^2(t) + \beta_1 u_1^2(t) + \beta_2 u_2^2(t)]\,dt + L(p(\tau), r(\tau))\,|\,p(t) = p,\ r(t) = r\}$$

is the conditional future cost at time t given the states at t. The problem (2.3.7), (2.3.8) can be solved with the optimality conditions in theorem 2.1.3.

From (2.1.8), the adjoint equation is given by

$$(2.3.10)\quad \begin{cases} -\psi_t + \tfrac{1}{2}\sigma_1^2\psi_{pp} + \tfrac{1}{2}\sigma_3^2\psi_{rr} - (u_1^0+ar)\psi_p - (u_2^0-ap)\psi_r \\ \quad - (u_{1p}^0+u_{2r}^0)\psi = 0, \text{ for } (t,p,r) \in [0,T] \times B \\ \psi(t,p,r) = 0, \text{ for } (p,r) \in \partial B \\ \psi(0,p,r) = \psi_0(p,r), \text{ for } (p,r) \in B,\ \pi_0(A) = \int_A \psi_0(p,r)\,dp\,dr \end{cases}$$

where $\psi(t,p,r)$ is the probability density of the stochastic process $[p(t), r(t)] = (p,r)$ at time t corresponding to the control law u^0, such that for any set $A \subset R^2$

$$\Pr\{[p(t), r(t)] \in A\} = \int_A \psi(t,p,r)\,dp\,dr.$$

In order that u^0 be an optimal control, it follows from (2.1.13) that the following integral inequality must hold for every control $u = (u_1, u_2)$ in the admissible class $\mathcal{U}$:

$$(2.3.11)\quad \int_{[0,T]\times B} (\phi_p^{u^0} + 2\beta_1 u_1^0)\psi(u_1-u_1^0) + (\phi_r + 2\beta_2 u_2^0)\psi(u_2-u_2^0) \cdot dt\,dp\,dr \leq 0$$

where ψ is the solution of (2.3.10) and ϕ^{u^0} is the solution of (2.3.7) corresponding to the optimal control u^0. From the optimality condition in the pointwise form (2.1.14),

$$(2.3.12)\quad \begin{aligned} &[\phi_p^{u^0}(t,p,r) + 2\beta_1 u_1^0(t,p,r)]\psi(t,p,r)u_1(t,p,r) \\ &+ [\phi_r^{u^0}(t,p,r) + 2\beta_2 u_2^0(t,p,r)]\psi(t,p,r)u_2(t,p,r) \\ &\leq [\phi_p^{u^0}(t,p,r) + 2\beta_1 u_1^0(t,p,r)]\psi(t,p,r)u_1^0(t,p,r) \\ &+ [\phi_r^{u^0}(t,p,r) + 2\beta_2 u_2^0(t,p,r)]\psi(t,p,r)u_2^0(t,p,r) \end{aligned}$$

holds for every point $(t,p,r) \in [0,T] \times B$ and every admissible control $u \in U$.

From section 2.2, the two inequalities (2.3.11) and (2.3.12) give rise to two different numerical approaches for obtaining the optimal control. Maximizing the left-hand side of the inequality (2.3.12) with respect to the values of u_1 and u_2 within the control constraints $|u_1| \leq C_1$, $|u_2| \leq C_2$, we obtain the following expressions for the optimal controls:

$$(2.3.13) \quad \begin{cases} u_1^0 = \mathrm{Sat}_{C_1}(-\phi_p^{u^0}/4\beta_1) \\ \\ u_2^0 = \mathrm{Sat}_{C_2}(-\phi_r^{u^0}/4\beta_2) \end{cases}$$

where

$$\mathrm{Sat}_k(y) \triangleq \begin{cases} k & \text{if} \quad y > k \\ y & \text{if} \quad |y| \leq k \\ -k & \text{if} \quad y < -k \end{cases}$$

Substituting u^0 of (2.3.13) into (2.3.7) and its associated adjoint equation (2.3.10), we arrive at a two-point boundary value problem involving two coupled nonlinear parabolic partial differential equations. As noted before, this approach is difficult to handle numerically. Instead we will consider the "gradient" method as presented in subsection 2.2.2.

Using (2.3.11), the gradients G_1 and G_2 defined as

$$(2.3.14)\qquad \begin{aligned} G_1(t,p,r,u_1,\phi_p,\psi) &\triangleq (\phi_p+2\beta_1u_1)\psi \\ G_2(t,p,r,u_2,\phi_r,\psi) &\triangleq (\phi_r+2\beta_2u_2)\psi \end{aligned}$$

can be used to obtain extremal controls, as in Algorithm A. The description of this procedure is repeated below.

Denote by u^n, $n = 1,2,\ldots$, an admissible control function such that $u^n(t,p,r) = (u_1^n(t,p,r),\ u_2^n(t,p,r))$. Let ϕ^n, ψ^n be the solutions of (2.3.7) and (2.3.10) corresponding to $u = u^n$ and $u^o = u^n$, respectively. Let

$$G^n = (G_1^n, G_2^n) \triangleq (G_1(t,p,r,u_1^n,\phi_p^n,\psi^n),\ G_2(t,p,r,u_2^n,\phi_r^n,\psi^n)).$$

Algorithm A is now used to construct the sequence of controls $\{u^n\}$ that converges to an extremal control. Specifically, we have the following procedure:

Algorithm

Guess an initial control $u^1(t,p,r)$ and perform the following steps for the n^{th} iteration:

Step 1: Truncate $u_1^n(t,p,r)$, $u_2^n(t,p,r)$ so that they take values within the control constraint sets. Denote the resulting controls again by u_1^n, u_2^n.

Step 2: Substitute u_1^n, u_2^n into (2.3.7) and (2.3.10), and solve the two uncoupled linear parabolic equations.

Step 3: Calculate the cost function $J(u^n)$ corresponding to u_1^n, u_2^n. If $|J(u^n) - J(u^{n-1})| < \varepsilon$, for some preassigned small number ε, iteration stops.

Step 4: Calculate G^n corresponding to u^n using (2.3.14)and evaluate the step sizes $\rho^n = (\rho_1^n, \rho_2^n)$ given by

$$(2.3.15)\quad \begin{cases} \rho_1^n = \int_{[0,T]\times B}(G_1^n)^2 dpdrdt/2\beta_1 \int_{[0,T]\times B}(G_1^n)^2 \psi^n dtdpdr \\ \rho_2^n = \int_{[0,T]\times B}(G_2^n)^2 dpdrdt/2\beta_2 \int_{[0,T]\times B}(G_2^n)^2 \psi^n dtdpdr. \end{cases}$$

Step 5: Set $u_1^{n+1} = u_1^n - \rho_1^n G^n$, $u_2^{n+1} = u_2^n - \rho_2^n G^n$ and start the next iteration by going to step 1.

Numerically, the most involved part of the above procedure is computing the solutions of the parabolic equations (2.3.7) and (2.3.10), which is discussed in appendix A. In the following subsection the numerical results of the satellite attitude control problem are presented.

2.3.3 NUMERICAL RESULTS

The satellite attitude control problem (2.3.5), (2.3.6) is solved using the alogrithm described in the last subsection with the data $I_x = I_z = 650$ slug$\cdot$ft^2, $I_y = 1000$ slug$\cdot$ft^2, $\sigma_1 = \sigma_2 = 10^{-5}$ rad/s, $|u_1|, |u_2| \leqslant 10^{-3}$ rad/s^2, $C\Omega = 14$ slug$\cdot$ft^2/s, $\omega_0 = 7.29 \times 10^{-5}$ rad/s, $|p|, |r| \leqslant 10^{-3}$ rad/s, initial distribution $\psi(0,p,r)$ of the states is assumed uniform, $\alpha_1 = \alpha_2 = \beta_1 = \beta_2 = 1$ s^4/rad^2, and

$$L(p,r) = \begin{cases} 0.6, & \text{for } (p,r) \in \partial B \\ 0, & \text{elsewhere.} \end{cases}$$

Using these values and multiplying p, r, u_1, u_2 by 10^3 so that the domain of interest is $\{(p,r): |p|, |r| \leqslant 1 \text{ rad/s}\}$ and the control constraints are $|u_1|, |u_2| \leqslant 1$ rad/s^2, the control problem is solved using the algorithm presented with the initial control guesses $u_1 = u_2 \equiv 1$ rad/s^2. The minimum (cost) is attained in 9 iterations as shown in figure 5 (curve 1).

For nonlinear problems it is difficult to compute the optimal ρ^n. In this case we guess initial step sizes ρ_1^1, ρ_2^1, and subsequent step sizes are reduced by a suitable factor if the cost function does not improve as compared to the previous iteration. Using this method, convergence of

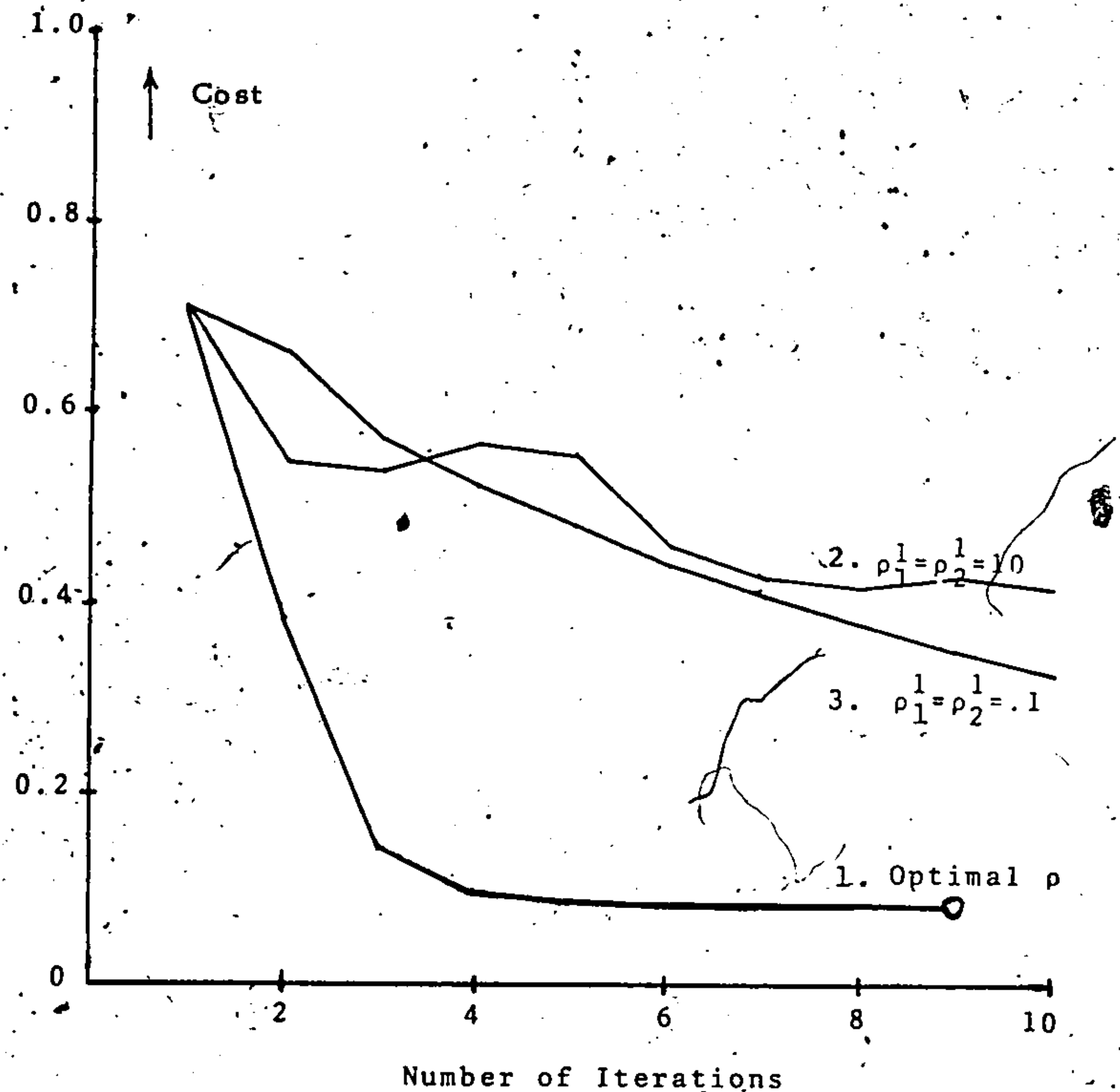


Figure 5. Rate of Convergence

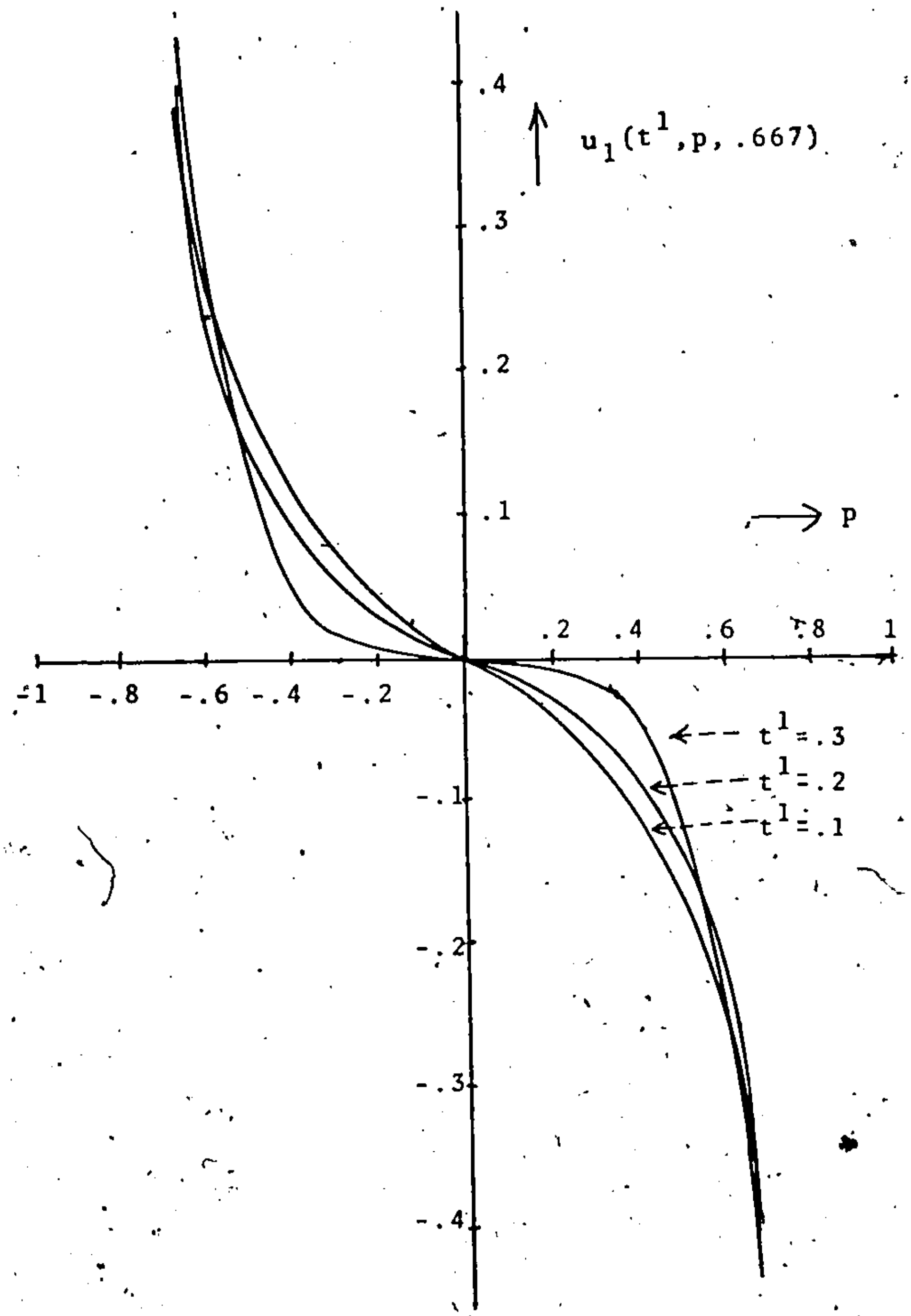


Figure 6. Optimal Control

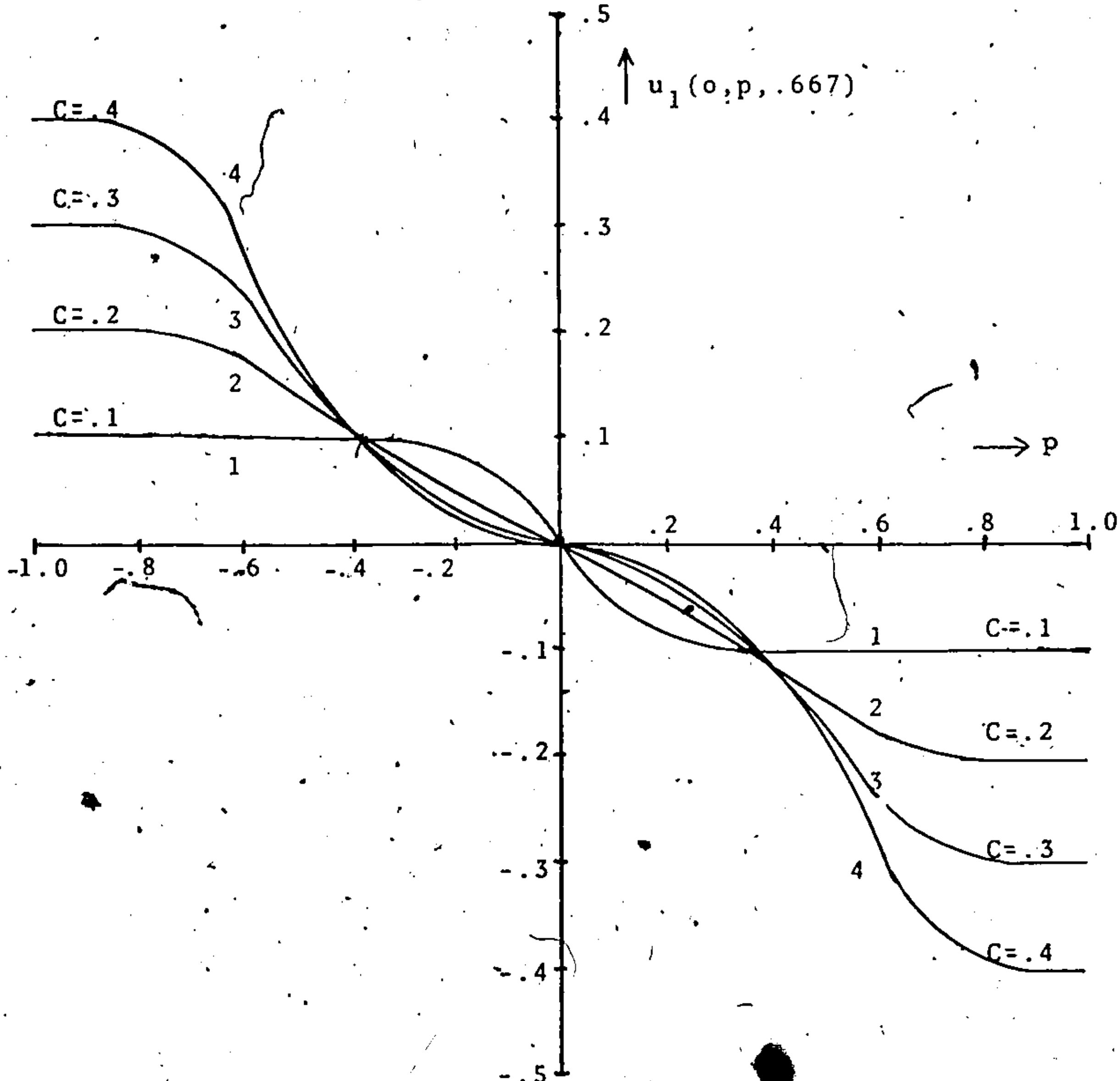


Figure 7. Optimal controls with varying control constraints C.

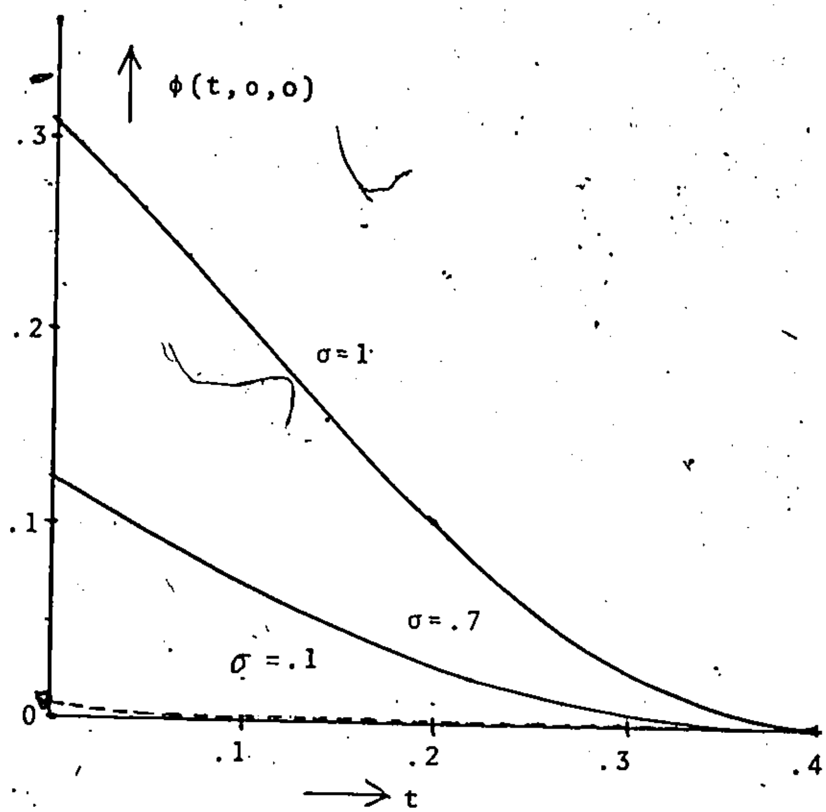


Figure 8. Conditional Expected Cost Corresponding to difference noise variances $\sigma = \sigma_1 = \sigma_2$

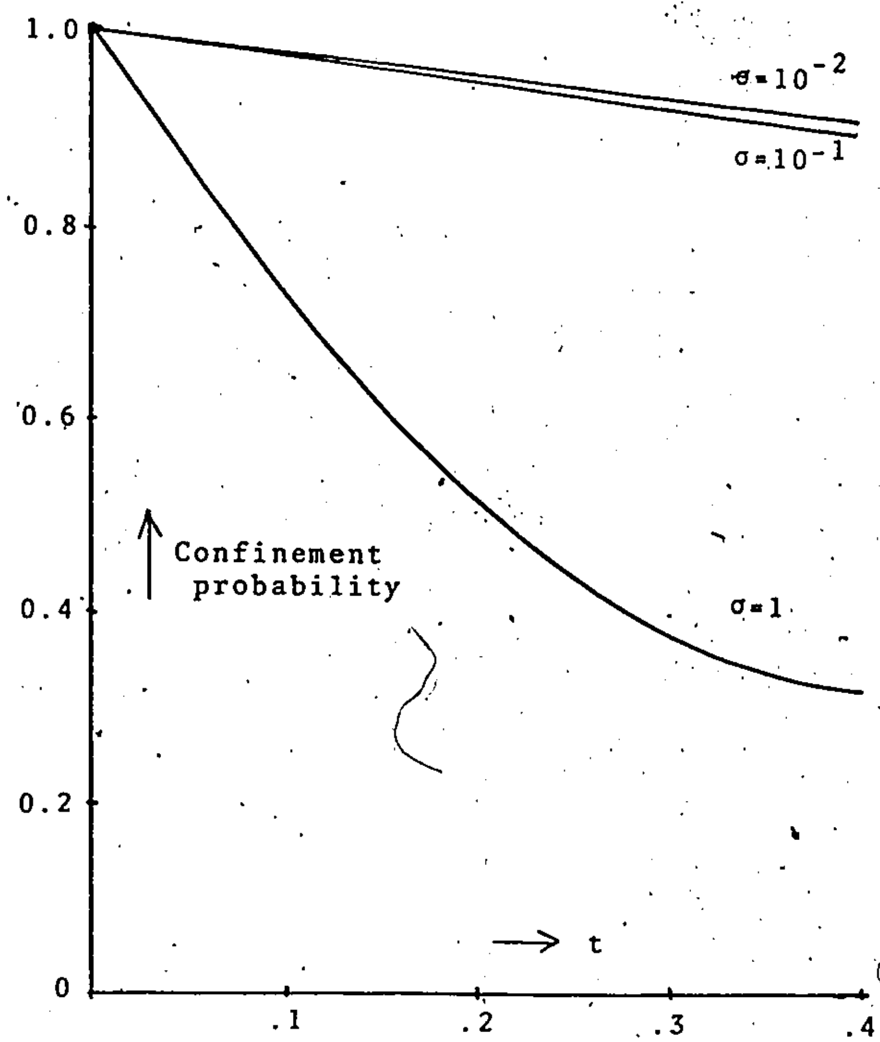


Figure 9. Confinement probability corresponding to different noise variances $\sigma = \sigma_1 = \sigma_2$

the cost function starting with i) $\rho_1^1 = \rho_2^1 = 10$ and ii) $\rho_1^1 = \rho_2^2 = 0.1$ is also shown in figure 5 (curves 2 and 3). It is observed that when ρ_1^1, ρ_2^1 are either too large or too small, there is no guarantee for cost improvement in every iteration and convergence is slower.

The extremal control obtained by this algorithm using the initial guess for the controls $u_1 = u_2 \equiv 1$ rad/s^2 has been found to be optimal and has been verified to be independent of the initial guess. For the sake of illustration a section of the optimal control taken at $t = t^1$, $r = 2/3$, and considered as a function of p is shown in figure 6 for $t^1 = 0.1, 0.2, 0.3$. The optimal control $u_1 (u_2)$ is symmetric with respect to the axis $r = 0$ ($p = 0$), but is skew symmetric (symmetric with one sign changed) with respect to the axis $p = 0$ ($r = 0$); this verifies the property (2.2.72) of the optimal control. The optimal control also confirms the intuition that $u_1 \leqslant 0$ ($u_2 \leqslant 0$) for $p \geqslant 0$ ($r \geqslant 0$) and $u_1 \geqslant 0$ ($u_2 \geqslant 0$) for $p \leqslant 0$ ($r \leqslant 0$).

Keeping all parameters fixed, the problem (2.3.5), (2.3.6) is solved using various control constraints. The optimal controls thus obtained are saturation functions as predicted by (2.3.13) and shows the same kind of symmetry as before. Numerical values of u_1^0, u_2^0 for control constraints $C_1 = C_2 = 0.1, 0.2, 0.3, 0.4$ are shown in figure 7.

In the formulation of the minimization problem (2.3.5), (2.3.6) it is assumed that the noise variance (or the noise intensity is fixed). However, in practical situations, the noise level may change according to the environment. Figure 8 shows the conditional expected costs

$$E\{\int_t^{\tau\wedge T}[p^2+r^2+u_1^2+u_2^2]dt + L[p(\tau\wedge T), r(\tau\wedge T)]\,|p(t) = r(t) = 0\} \equiv \phi(t,0,0)$$

for noise variances $\sigma = (\sigma_1,\sigma_3) = (1,1)$, $(0.7,0.7)$, and $(0.1,0.1)$.

The set B, as defined earlier, is determined from the tolerable limits of precession as set by the designer. The optimal controls found are those that minimize the fuel cost and the deviation of the satellite momenta from the origin. Clearly, the optimal control laws depend on the choice of the cost function. It is of interest to investigate the effectiveness of these optimal controls in confining the momenta to within the tolerable set B. The confinement probability at time t is given by $\int_B \psi(t,p,r)dpdr$. Figure 9 gives the confinement probability using the optimal controls corresponding to different noise parameters. It is observed that the system moves out of B rapidly in a highly noisy environment, as expected. If, however, the confinement probability itself is to be optimized, then the result would be different.

In the above optimal control problem we have imposed a non zero boundary cost L; a boundary cost was not imposed in the problem of subsection 2.2.4. This main difference between the two problems is reflected in the behaviour of their respective optimal controls in the region near the boundary. In this region, if there is no terminal cost, the optimal control would push the system out of B to speed up the exit time, thus minimizing the cost function. In the same region, if there is a terminal cost, the optimal control would not force the system to exit (figures 4 and 7).

The results of the two problems, however, have certain simularities; the corresponding minimum costs differs approximately by the boundary cost, and the patterns of convergence are strikingly similar (figures 1 and 5). This may indicate that, for the satellite control problem, algorithm A is 'insensitive' to the boundary costs.

2.3.4 AN ALTERNATIVE APPROACH TO THE SATELLITE ATTITUDE CONTROL PROBLEM

In this subsection, the satellite attitude control problem is reformulated, and the resulting control will be obtained in closed form. The reformulated problem involves a three dimensional attitude dynamics (state variables p,q,r) and a free end time in the cost function (instead of the fixed end time T). An additional assumption is made—the noise variance σ^2 is assumed small. With this assumption, it is possible to obtain a closed form approximate solution for the optimal control. The procedure of solution involves solving the corresponding deterministic problem. The resulting analytic solution is used to approximate the optimal control in the reformulated problem. When the assumption of 'small noise intensity' is satisfied, this method is much more computationally efficient because an analytic expression for the approximate optimal control is obtainable.

For clarity of presentation, the notation used here is slightly different from those of the last section:

I_1, I_2, I_3, Moments of inertia of satellite in the x, y and z axes.

x_1, x_2, x_3, $x_1 = p$, $x_2 = q - q_0$, $x_3 = r$; p,q,r are angular momentum components of the satellite.

x Angular momentum vector, $x = [x_1, x_2, x_3]'$.

D	Target set.
ω_0	A constant, $\omega_0 = 2\pi$ rad. per 24 hrs.
C	Moment of inertia of the flywheel mounted in the y axis.
Ω	Angular velocity of the flywheel in the y axis.
$u_1, u_2, u_3,$	Jet controls in the x, y and z axes respectively.
u	Jet control vector, $u = [u_1, u_2, u_3]'$.
$n_1, n_2, n_3,$	Uncorrelated white noise with unit variance.
σ	A constant reflecting the intensity of noise disturbances.
$\lvert\cdot\rvert$	Enclidean norm.

The three dimensional satellite attitude dynamics is, from (2.3.2),

$$
(2.3.16)\quad \begin{cases} \dot{x}_1 = a_1 x_3 + a_2(x_2 + q_0)x_3 + u_1 + 2\sigma n_1 \\ \dot{x}_2 = b_2 x_1 x_3 + u_2 + 2\sigma n_2 \\ \dot{x}_3 = c_1 x_1 + c_2 x_1 (x_2 + q_0) + u_3 + 2\sigma n_3 \\ x(0) = x_0 \text{ with distribution } \pi(x). \end{cases}
$$

Where for a satellite having $I_1 = I_3$

$$
(2.3.17)\quad \begin{cases} a_1 = [\omega_0(I_3 - I_2) + C\Omega]/I_1, \\ a_2 = (I_2 - I_3)/I_1,\ b_2 = (I_3 - I_1)/I_2 = 0, \\ c_1 = [\omega_0(I_2 - I_1) - C\Omega]/I_3, \\ c_2 = (I_1 - I_2)/I_3. \end{cases}
$$

Let the target set D, representing a set of tolerable angular momenta, be defined by

$$(2.3.18)\quad D \underline{\Delta} \{x \in R^3 \,|\, |x| \leq \epsilon,\ 0 < \epsilon < |x_0|\}$$

Denote by τ the first time the state $x(t)$ hits D.

Initially the satellite has been perturbed to the state $x(0) = x_0$. It is required to design an optimal feedback controller $u(t,x)$, $(t,x) \in R^4$, to drive the initial state x_0 to the target D so as to minimize the cost function

$$(2.3.19)\quad J^{\sigma}(u) = E \int_0^{\tau} |x(t)|^2 + |u(t,x(t))|^2 \, dt.$$

The problem (2.3.16), (2.3.19) will be refered to as problem 1.

Deterministic Optimal Control

We first solve a deterministic version of problem 1. The results will be used later to obtain an approximate expression for optimal control of problem 1.

To define the deterministic problem, we drop the stochastic quantities from (2.3.16). The resulting deterministic body rate dynamics is then

$$(2.3.20)\quad \begin{cases} \dot{x}_1 = a_1 x_3 + a_2 (x_2 + q_0) x_3 + u_1, \\ \dot{x}_2 = b_2 x_1 x_3 + u_2, \\ \dot{x}_3 = c_1 x_1 + c_2 x_1 (x_2 + q_0) + u_3, \\ x(0) = x_0,\ x_0 \text{ fixed.} \end{cases}$$

where the constants a_1, a_2, b_2, c_1 and c_2 are as defined in (2.3.17). These constants have the properties

(2.3.21) $\quad a_1 + c_1 = 0$

(2.3.22) $\quad a_2 + b_2 + c_2 = 0, \quad b_2 = 0$

It is required to design a controller $u(t)$ that would drive the states from $x(0) = x_0$ to the region D (as defined in (2.3.18)) and minimizing the cost function

(2.3.23) $\quad I(u) = \int_0^\tau \alpha|x(t)|^2 + \beta|u(t)|^2 \, dt$

where $\alpha, \beta > 0$ and τ is the first time the trajectory hits D. Such a control will be called an optimal control. For convenience, the problem (2.3.20), (2.3.23) will be called problem 2.

An application of Pontryagin's Maximum Principle will lead to a two point boundary value problem. Here we use the special structure (2.3.21), (2.3.22) of the dynamics (2.3.20) and consider the optimal control problem directly as a problem of calculus of variations. The optimal controller will be obtained in closed form.

Define $\rho(t) = |x(t)|$. Then from (2.3.20), (2.3.21) and (2.3.22) it follows that $\rho(t)\dot{\rho}(t) = x(t)\dot{x}(t) = u \cdot x$ or

(2.3.24) $\quad u(t) \cdot \dfrac{x(t)}{|x(t)|} = \dot{\rho}(t).$

Define $\theta : [0,\infty) \to [0,\pi]$ such that $\theta(t)$ = angle between **the** vectors $x(t)$ and $u(t)$.

By (2.3.24), the cost integral can be written as

(2.3.25) $$\int_0^\tau \alpha\rho^2(t) + \frac{\beta\dot{\rho}^2(t)}{\cos^2\theta(t)}\,dt,$$

and problem 2 is reduced to choosing τ, ρ and θ such that (2.3.25) is minimized. Let the integral (2.3.25) attain its minimum at $\tau = \tau^*$, $\rho = \rho^*$ and $\theta = \theta^*$. Then it is easy to show that either $\theta^*(t) \equiv 0$, or $\theta^*(t) \equiv \pi$. As observed from (2.3.24), $\theta^*(t) \equiv 0$ will result in a montonic increase of the magnitude $\rho(t)$ of the body rates with respect to t. Therefore we must have

(2.3.26) $$\theta^*(t) \equiv \pi,$$

and the integral (2.3.25) reduces to

(2.3.27) $$\int_0^\tau \alpha\rho^2(t) + \beta\dot{\rho}^2(t)\,dt$$

For any fixed τ, ρ^* has to satisfy the Euler Lagrange equation given by

(2.3.28) $$\ddot{\rho}^* - \frac{\alpha}{\beta}\rho^* = 0,\ \rho^*(0) = |x_0|,\ \rho^*(\tau) = \epsilon$$

The solution to (2.3.28) is

(2.3.29) $$\rho^*(t) = Ae^{-\omega t} + Be^{\omega t}$$

where

(2.3.30) $$\begin{cases} \omega = (\alpha/\beta)^{\frac{1}{2}} \\ A = (|x_0|e^{\omega\tau} - \epsilon)/(e^{\omega\tau} - e^{-\omega\tau}), \\ B = (\epsilon - |x_0|e^{-\omega\tau})/(e^{\omega\tau} - e^{-\omega\tau}). \end{cases}$$

The cost integral (2.3.25) corresponding to $\theta = \theta^*$, $\rho = \rho^*$ is

$$(2.3.31) \qquad (\alpha\beta)^{\frac{1}{2}}\{B^2(e^{2\omega\tau} - 1) - A^2(e^{-2\omega\tau} - 1)\}.$$

Minimizing (2.3.31) with respect to τ we obtain

$$(2.3.32) \qquad \tau^* = (\beta/\alpha)^{\frac{1}{2}} \cosh^{-1}\{\tfrac{1}{2}(\frac{\epsilon}{x_0} + \frac{x_0}{\epsilon})\}.$$

From (2.3.29), (2.3.30) and (2.3.32) it follows that

$$(2.3.33) \qquad \rho^*(t) = |x_0|e^{-\omega t}.$$

Letting $\rho = \rho^*$ in (2.3.24), we obtain the optimal control for problem 2:

$$u^*(t) = \dot{\rho}^*(t)\frac{x(t)}{\rho^*(t)}$$

or equivalently

$$(2.3.34) \qquad u^*(t) = -\omega x(t), \quad t \in [0,\tau^*].$$

The optimal cost of problem 2 is then

$$(2.3.35) \qquad I(u^*) = (\alpha\beta)^{\frac{1}{2}}(|x_0|^2 - \epsilon^2).$$

This completes the solution of problem 2.

The optimal control u^* as given by (2.3.34) is a proportional control and is therefore easy to implement. It is observed that for each $t \in [0,\tau^*]$, the control vector $u^*(t)$ always acts in the opposite direction as the state vector $x(t)$, and $|u^*(t)|$ is given by

(2.3.36) $\quad |u^*(t)| = \omega|x_0|e^{-\omega t}, \; t \in [0,\tau^*].$

The parts of the cost due to the control and due to the states are equal, in fact

(2.3.37) $\quad \int_0^{\tau^*} \beta|u^*(t)|^2 \, dt = (\alpha\beta)^{\frac{1}{2}}(|x_0|^2 - \epsilon^2)/2,$

(2.3.38) $\quad \int_0^{\tau^*} \alpha|x(t)|^2 \, dt = (\alpha\beta)^{\frac{1}{2}}(|x_0|^2 - \epsilon^2)/2.$

If the optimal control is applied at time t when the state is x(t), then the cost will be

(2.3.39) $\quad \phi^0(t,x) = \min_u \int_t^{\tau^*} \alpha|x(\theta)|^2 + \beta|u(\theta)|^2 \, d\theta$

(2.3.40) $\quad \phi^0(t,x) = (\alpha\beta)^{\frac{1}{2}}(|x(t)|^2 - \epsilon^2).$

The expression (2.3.40) will be used below to compute the cost function of the stochastic problem (problem 1).

It is observed that from (2.3.32) that τ^*, the time the optimal trajectory x hits the target D, does not depend on the constants in the dynamics (2.3.20). This is due to the special properties (2.3.21) and (2.3.22) of these constants.

As observed from (2.3.32), when $\epsilon = 0$, the optimal control will take infinite time to drive the system to the origin. However the corresponding cost remains finite and is equal to $(\alpha\beta)^{\frac{1}{2}}|x_0|^2$.

If a positive constant K is added to the cost integrand in (2.3.23), both the optimal control and the optimal trajectory remains the same. Hence the addition of K in the cost function will not speed up the final time τ^* but only serves to increase the value of the cost.

It is interesting to note that although the dynamics is nonlinear, the optimal control turns out to be linear in the state.

The Stochastic Problem

We now consider the stochastic problem (problem 1), in which white noise terms appear in the body rate dynamics. Corresponding to problem 1, we define the future cost $\phi^\sigma(t,x)$ by

$$\text{(2.3.41)} \qquad \phi^\sigma(t,x) = \min_u E_{t,x} \int_t^\tau |x(s)|^2 + |u(s,x(s))|^2 \, ds,$$

where $E_{t,x}(\cdot)$ denotes condition expectation given the information $x(t) = x$. Then from section 2.1, ϕ^σ is a solution to the following deterministic optimization problem,

$$\text{(2.3.42)} \quad \begin{cases} v_t + \sigma^2 \sum_{i,j=1}^{3} v_{x_i x_j} + [a_1 x_3 + a_2 x_2 x_3 + u_1(t,x)] v_{x_1} \\ \quad + u_2(t,x) v_{x_2} + [c_1 x_1 + c_2 x_1 x_2 + u_3(t,x)] v_{x_3} \\ \quad + \alpha |x|^2 + \beta |u(t,x)|^2 = 0, \\ \min_u \int_{R^3} v(t,x) \, d\pi(x). \end{cases}$$

The problem (2.3.42) involves a parabolic partial differential equation and can be solved by the gradient methods of section 2.2.

We consider here a simpler approach using Fleming's small noise approximation [22]. A description of this technique has been given by Holland [27]. Let ϕ^0 be as defined in (2.3.39), i.e. ϕ^0 equals to the expected future cost for the deterministic problem (problem 1). Further let x^0 denotes the optimal trajectory of problem 1. Then, using Fleming's method, the expected future cost for problem 1, as defined in (2.3.41) is given by

$$(2.3.43)\quad \phi^\sigma(t,x) = \phi^0(t,x) + \sigma^2\gamma(t,x) + 0(\sigma^2),\quad \lim_{h\to 0}\frac{0(h)}{h} = 0$$

where

$$(2.3.44)\quad \gamma(t,x) = \sum_{i=1}^{3}\int_t^{\tau^*} \phi^0_{x_i x_i}(t,x^0(t))\,dt,$$

and

$$(2.3.45)\quad \phi^\sigma_{x_i}(t,x) = \phi^0_{x_i}(t,x^0) + \sigma^2\gamma_{x_i}(t,x^0) + 0(\sigma^2),$$

and the optimal control $u(t,x)$ for problem 1 satisfies

$$(2.3.46)\quad \min_{u\in R^3}\{\alpha|x^0|^2 + \beta|u|^2 + f_i(x^0,u)\cdot\phi^\sigma_{x_i}\}$$
$$= \alpha|x^0|^2 + \beta|u^\sigma|^2 + f_i(x^0,u^\sigma)\cdot\phi^\sigma_{x_i},$$

where

$$(2.3.47)\begin{cases} f_1(x,u) = a_1x_3 + a_2x_2x_3 + u_1, \\ f_2(x,u) = b_2x_1x_3 + u_2, \\ f_3(x,u) = c_1x_1 + c_2x_1x_2 + u_3. \end{cases}$$

The expected future cost as calculated from (2.3.43) and (2.3.44) is given by

$$(2.3.48)\qquad \phi^{\sigma}(t,x) = (\alpha\beta)^{\frac{1}{2}}(|x|^2 - \epsilon^2) - 6\sigma^2\beta\cosh^{-1}\{\tfrac{1}{2}(\frac{\epsilon}{|x|} + \frac{|x|}{\epsilon})\}$$

$$- 6\sigma^2(\alpha\beta)^{\frac{1}{2}}t + 0(\sigma^2)$$

The corresponding optimal control is found from (2.3.46) to be

$$(2.3.49)\qquad u^{\sigma}(t,x) \simeq -(\omega + \frac{3\sigma^2}{|x|^2})\ x.$$

Let us compare the results of the deterministic problem (problem 2) and the stochastic problem (problem 1). In both cases the (approximate) optimal controls act in the opposite direction as the state vector. However the stochastic control u^{σ} has a larger magnitude than the deterministic optimal control u^*,

$$(2.3.50)\qquad |u^*(t)| = \omega|x(t)|,$$

$$(2.3.51)\qquad |u^{\sigma}(t,x)| = \omega|x| + 3\sigma^2/|x|.$$

The difference in the control magnitudes increases with the noise parameter σ. Due to the presence of white noise disturbance, the stochastic control becomes more powerful near the target so as to speed up the time to hit the target D and thus minimizing the value of the cost function.

The initial costs for problems 2 and 1 are given by

$$(2.3.52)\qquad \phi^0(0,x) = (\alpha\beta)^{\frac{1}{2}}(|x|^2 - \epsilon^2),$$

$$(2.3.53) \qquad \phi^{\sigma}(0,x) = (\alpha\beta)^{\frac{1}{2}}(|x|^2 - \epsilon^2) - 6\sigma^2\beta \cosh^{-1}\{\tfrac{1}{2}(\tfrac{\epsilon}{x} - \tfrac{x}{\epsilon})\}.$$

The cost for the stochastic case is larger and, as the noise parameter decreases to zero, the two costs become identical. As $\epsilon \to 0$, the cost for the stochastic case becomes infinite as observed from (2.3.48); the reason is that it is impossible to drive the states in a noisy environment to a target consisting of a single point in the state space.

2.4 SUGGESTIONS FOR FURTHER STUDY

Two gradient algorithms for computing the optimal control of completely observable stochastic differential systems have been presented in section 2.2. Algorithm A with suboptimal step size has in general the advantage of faster convergence whereas algorithm B has the advantage of solving only one parabolic equation per iteration. The author would recommend algorithm A when suboptimal step sizes are computable, and algorithm B otherwise. The reason is that suboptimal step sizes would save computation time. Further, if suboptimal sizes are not computatable, then various step sizes have to be tested; in this case it would be more economical to use algorithm B, which requires less computation time per iteration.

Another approach for solving the optimal control problem is the dynamic programming method. The resulting optimality condition is a non-linear parabolic equation (details in the next chapter). Bellman has suggested a quasilinearization method which solves the non-linear parabolic equation by successively solving a family of linear parabolic equations [9] [21]. It is noted that the gradient method has an advantage over the quasilinearization method —— successive improvements in the control are guided by the gradient of the cost function.

Optimality conditions of the stochastic optimal control problems based on partially observable states have been given by several authors, e.g. [5] [20]. It would be interesting to investigate the computational aspects of these optimality conditions.

In section 2.3, a satellite attitude control problem has been solved by the gradient method of section 2.2. The problem is then reformulated in three dimension and an optimal control is obtained in closed form. However in some satellites, the jet controls are in the form of a sequence of pulses. The optimal timing and magnitudes of these pulses would be an interesting control problem.

CHAPTER 3

OPTIMAL CONTROL OF STOCHASTIC DIFFERENTIAL SYSTEMS WITH JUMP PARAMETERS

CHAPTER 3 OPTIMAL CONTROL OF STOCHASTIC DIFFERENTIAL SYSTEMS WITH JUMP PARAMETERS

In this chapter a particular type of jump parameters, namely, first order Markov chain, is introduced into the stochastic differential system. In section 3.1 the resulting optimal control problem is investigated by applying the dynamic programming approach to a machine maintenance problem. Dynamic programming equations and backward Kolmogorov equations are derived. To illustrate the dynamic programming approach, closed form solutions are presented for two sample problems: a special case of the machine maintenance problem, and an investor's problem. In section 3.2 we consider the variational approach. Optimality conditions similar to those of chapter 2 are developed. The optimality conditions are applied to the special case of a linear quadratic Gaussian problem; the resulting equations are more general than those given by Sworder [54] and Wonham [66]. A gradient method similar to that of subsection 2.2.2 is proposed. In section 3.3, dynamic programming equations for the problem from section 2.2 are presented. The numerical aspects of the dynamic programming approach and variational approach are briefly discussed. Some suggestions for further studies are then made.

3.1 DYNAMIC PROGRAMMING APPROACH

In this section we consider the dynamic programming approach as applied to optimal control of stochastic differential systems with jump parameters. To gain some physical insights, the subject is investigated through a machine maintenance problem, which is formulated in subsection 3.1.1. Dynamic programming equations are derived to compute the optimal control and the value function. The classical backward equation is generalized to cover the stochastic problem with jump parameters. Derivations of the dynamic programming equation and the backward equation are presented in subsection 3.1.2. In subsection 3.1.3 we consider the special case in which white noise disturbance is absent. Closed form solutions are obtained for a machine maintenance problem. To illustrate the case in which white noise is present, we consider in subsection 3.1.4 an investor's problem, which also has a closed form solution.

The convention of repeated indices is used throughout this chapter.

3.1.1 PROBLEM FORMULATION — A MACHINE MAINTENANCE PROBLEM

Suppose a plant has n production units (e.g. machines). Each unit is either idle, or under repair, or in operation.

If any unit is in operation, the plant manager has to determine an appropriate amount of preventive maintenance for that unit. If any unit is down, he has to decide whether to repair it, and he also has to decide on the repair rate if required. The operating conditions of all units are described by a continuous parameter Markov chain y with m states. The number m corresponds to the number of all possible combinations of individual machine status (operational or down).

We now characterise the Markov chain y. Define the following quantities:

$C \equiv$ a compact subset of the real line R,

$U \equiv$ a compact subset of R,

$I \equiv [0,T]$, a time interval,

$M \equiv \{1,\ldots,m\}$, $m < \infty$,

$N \equiv \{1,\ldots,n\}$, $n < \infty$.

The failure rates $\lambda_j : I \to C$, $j \in N$ are defined as

(3.1.1) $$\lambda_j(t) \equiv \lim_{\Delta t \downarrow 0} \frac{1}{\Delta t} \text{Prob}\{j\text{th unit down at } t + \Delta t \mid j\text{th unit operational at } t\}$$

The maintenance policy $\mu : I \times M \times N \to U$ is defined by

(3.1.2) $$\mu(t,i,k) \equiv \lim_{\Delta t \downarrow 0} \frac{1}{\Delta t} \text{Prob}\{k\text{th unit operational at } t + \Delta t \mid k\text{th unit down at } t, y(t) = i\}$$

For convenience we also use the notation

$$\mu(t,y(t)) = [\mu(t,y(t),1),\ldots, \mu(t,y(t), n)]'.$$

The transition probability matrix P^{μ} corresponding to the maintenance policy μ is defined as

$$(3.1.3)\quad \begin{cases} P^{\mu}(s,t) = \{P^{\mu}_{ij}(s,t),\ i,\ j = 1,\ldots,m\}, & t \geq s, \\ P^{\mu}_{ij}(s,t) = \text{Prob}\{y(t) = j \mid y(s) = i,\ \mu \text{ fixed}\}. \end{cases}$$

A property of P^{μ} is that it satisfies the following forward equation

$$(3.1.4)\quad \begin{cases} \frac{d}{dt}P^{\mu}(s,t) = P^{\mu}(s,t)Q^{\mu}(t), \\ P^{\mu}(s,s) = \overline{I} \equiv \text{identity matrix}, \end{cases}$$

where the differential transition matrix Q is given by

$$(3.1.5)\quad \begin{cases} Q^{\mu}(t) = \{q^{\mu}_{ij}(t)\} \\ q^{\mu}_{ij}(t) = \begin{cases} \lim\limits_{\Delta t \downarrow 0} \frac{1}{\Delta t} P^{\mu}_{ij}(t,t+\Delta t), & i \neq j, \\ \lim\limits_{\Delta t \downarrow 0} \frac{1}{\Delta t} (P^{\mu}_{ij}(t,t+\Delta t) - 1). \end{cases} \end{cases}$$

The system output ξ (n - vector) is assumed perturbed by an r - dimensional standard Wiener process w,

$$(3.1.6)\quad d\xi(t) = f(t,\xi(t),\ y(t))\ dt + \sigma(t)\ dw(t)\ ,$$

where $f: I \times R^{n} \times M \to R^{n}$, $\sigma: I \to R^{n\times r}$. We will consider the cost function

$$(3.1.7)\quad J(\mu) = E\{\int_0^T L(t,\xi(t),y(t),\mu(t,y(t)))dt + G(\xi(T),y(T),\mu(T,y(T)))\}$$

Where $L : I \times R^n \times M \times U^N \to R$,

$$L(t,\xi(t),y(t),\mu(t,y(t))) = L_1(t,\xi(t),y(t)) + L_2(t,\mu(t,y(t)),y(t))$$

$L_1 : I \times R^n \times M \to R$,

$L_2 : I \times U^N \times M \to R$, and

$G : R^n \times M \times U^N \to R$

The optimal maintenance problem is: Subject to the dynamic constraint (3.1.6), find a feedback maintenance policy $\mu^0 : I \times M \times N \to U$ such that μ^0 is measurable on I and that μ^0 minimizes the cost function (3.1.7).

As usually done in the dynamic programming approach, we will develop an equation for the value function $\phi(t,x,k)$. The value function is intepreted as the expected value of the future cost accumulated from t to T under the optimal policy given the present states $y(t) = k$ and $\xi(t) = x$. Mathematically we define $\phi : I \times R^n \times M \to R$ by

$$\phi(t,x,k) = \operatorname*{Inf}_{\mu} E_{t,x,k} \{\int_t^T L(\theta,\xi(\theta),y(\theta),\mu(\theta,y(\theta)))d\theta + G(\xi(T),y(T),\mu(T,y(T)))\}. \tag{3.1.8}$$

3.1.2 DYNAMIC PROGRAMMING METHOD AND BACKWARD EQUATION

This subsection contains the derivation of the dynamic programming equation and the backward equation for the maintenance problem of the last subsection. Solving the dynamic programming equation we can obtain the value function and the optimal maintenance policy. Solving the backward equations, we can obtain the expected future cost of various cost components.

Now we show that the value function ϕ satisfies the dynamic programming equation

$$(3.1.9)\quad \begin{cases} \dfrac{\partial\phi}{\partial t} + a_{ij}\phi_{x_i x_j} + \min\limits_{\mu}\{b_i\phi_{x_i} + Q^{\mu}\phi + F^{\mu}\} = 0 \\ \phi(T,x) = V(x) \end{cases}$$

where

$$(3.1.10)\quad \begin{cases} \phi(t,x) = [\phi(t,x,1),\ldots,\phi(t,x,m)]' \\ a_{ij}(t) = \tfrac{1}{2}(\sigma\,\sigma')_{ij}(t),\ i,j = 1,\ldots,n, \\ b_i(t,x) = \mathrm{diag}\,(f(t,x,1),\ldots,f(t,x,m)), \\ F^{\mu}(t,x) = [L(t,x,1,\mu(t,1)),\ldots,\ L(t,x,m,\mu(t,m))]', \\ V(x) = [G(x,1,\mu(T,1)),\ldots,\ G(x,m,\mu(T,m))]'. \end{cases}$$

The minimization in the non-linear parabolic system (3.1.9) is defined componentwise, i.e.,

$$(3.1.11)\quad \min_{\mu}\{Y\} = \left[\min_{\mu(\cdot,1)} Y_1, \min_{\mu(\cdot,2)} Y_2, \ldots, \min_{\mu(\cdot,m)} Y_m\right]',$$

$$Y = [Y_1,\ldots,Y_m]'.$$

The proof of (3.1.9) is now presented.

<u>Lemma 3.1.1</u> Let ξ satisfy (3.1.6) and let Q^{μ} be the differential transition matrix of y corresponding to a fixed policy μ. Let $h : I \times R^n \times M \to R$ belong to $C^{1,2}$*, then

---

* $C^{1,2}$ denotes the class of scalar functions g defined on $I \times R^n \times M$ with continuous first partial in $t \in I$ and continuous second partial in $x_n \in R^n$, except for a set of Lebesque measure zero in $I \times R^n$.

$$\lim_{\Delta t \downarrow 0} E_{t,x,k} \frac{1}{\Delta t}\{h(t+\Delta t,\xi(t+\Delta t),y(t+\Delta t)) - h(t,\xi(t),y(t))\}$$

$$(3.1.12) \qquad = h_t(t,x,k) + a_{ij}(t)h_{x_i x_j}(t,x,k) + f_i(t,x,k)h_{x_i}(t,x,k)$$

$$+ q^{\mu}_{kj}(t)h(t,x,j).$$

where f_i denotes the i^{th} component of f.

Proof: The left side of (3.1.12) can be written as

$$\lim_{\Delta t \downarrow 0} E_{t,x,k}\left\{\frac{h(t+\Delta t,\xi(t+\Delta t),y(t+\Delta t)) - h(t+\Delta t,\xi(t+\Delta t),y(t))}{\Delta t}\right\}$$

$$(3.1.13) \qquad + \lim_{\Delta t \downarrow 0} E_{t,x,k}\left\{\frac{h(t+\Delta t,\xi(t+\Delta t),y(t)) - h(t,\xi(t),y(t))}{\Delta t}\right\}$$

$$\equiv A + B,$$

where A and B denote the individual terms on the left side of (3.1.13). The term A can be expressed as follows:

$$(3.1.14) \qquad A = \lim_{\Delta t \downarrow 0} E_{t,x,k} \frac{1}{\Delta t}\{\mathrm{Prob}[y(t+\Delta t) = j|y(t) = y]\cdot$$

$$\cdot h(t+\Delta t,\xi(t+\Delta t),j) - h(t+\Delta t,\xi(t+\Delta t),y(t))\},$$

But $\quad \mathrm{Prob}\{y(t+\Delta t) = j|y(t) = y\}h(t+\Delta t,\xi(t+\Delta t),j)$

$$(3.1.15) \qquad = \sum_{j \neq y}[q_{yj}(t)\Delta t + o(\Delta t)]h(t+\Delta t,\xi(t+\Delta t),j)$$

$$+ [1 + q_{yy}\Delta t + o(\Delta t)]h(t+\Delta t,\xi(t+\Delta t),y)$$

$$= q_{yj}(t)\cdot\Delta t\cdot h(t+\Delta t,\xi(t+\Delta t),j) + h(t+\Delta t,\xi(t+\Delta t)y) + o(\Delta t)$$

Using (3.1.15) into (3.1.14), we obtain the expression A as

$$(3.1.16)\qquad A = \lim_{\Delta t\downarrow 0} E_{t,x,k}\,\frac{1}{\Delta t}\{q_{yj}(t)\Delta t\cdot h(t+\Delta t,\xi(t+\Delta t),j) + o(\Delta t)\}$$

$$= E_{t,x,k}\{q_{yj}(t)h(t,\xi(t),j)\}$$

$$= q_{kj}(t)h(t,x,j).$$

The expression B can be written as

$$(3.1.17)\qquad B = \lim_{\Delta t\downarrow 0} E_{t,x,k}\{\frac{h_t(t,\xi(t),y(t))\Delta t}{\Delta t} + \frac{h_{x_i}(t,\xi(t),y(t))\Delta\xi_i}{\Delta t}$$

$$+ \frac{h_{x_i x_j}(t,\xi(t),y(t))\Delta\xi_i\Delta\xi_j}{2\Delta t} + \frac{o(\Delta t\cdot\Delta\xi_i)}{\Delta t} - \frac{o(\Delta t)}{\Delta t}$$

$$+ \frac{o(\Delta\xi_i\cdot\Delta\xi_j)}{\Delta t}\}$$

where $\Delta\xi_i$ denotes $\xi_i(t+\Delta t) - \xi_i(t)$. Some of the terms in (3.1.17) are evaluated as follows*,

$$(3.1.18)\qquad \lim_{\Delta t\downarrow 0} E_{t,x,k}\,\frac{h_{x_i}(t,\xi(t),y(t))\Delta\xi_i}{\Delta t}$$

$$= \lim_{\Delta t\downarrow 0} E_{t,x,k}\, h_{x_i}(t,\xi(t),y(t))\,\frac{[f_i(t,\xi(t),y(t))\Delta t + \sigma_{ip}(t)(w_p(t+\Delta t) - w_p(t))]}{\Delta t}$$

$$= f_i(t,x,k)h_{x_i}(t,x,k)$$

$$(3.1.19)\qquad \lim_{\Delta t\downarrow 0} E_{t,x,k}\,\frac{1}{2\Delta t}\,h_{x_i x_j}(t,\xi(t),y(t))\cdot$$

$$\{f_i(t,\xi(t),y(t))\Delta t + \sigma_{ip}(t)(w_p(t+\Delta t) - w_p(t))\}\cdot$$

$$\{f_j(t,\xi(t),y(t))\Delta t + \sigma_{js}(t)(w_s(t+\Delta t) - w_s(t))\}$$

$$= a_{ij}(t)h_{x_i x_j}(t,x,k)$$

---

* The convention of repeated indices are used in (3.1.18) (3.1.19).

Using (3.1.18) and (3.1.19) into (3.1.17) we obtain

(3.1.20) $$B = h_t(t,x,k) + f_i(t,x,k)h_{x_i}(t,x,k) + a_{ij}(t)h_{x_ix_j}(t,x,k).$$

The equation (3.1.12) now follows from (3.1.16) and (3.1.20) This completes the proof.

The above lemma is now used to prove a 'verification theorem' which verifies that the optimal expected cost ϕ (defined by (3.1.8)) satisfies the dynamic programming equation (3.1.9). This verification theorem also verifies that the repair policy obtained by minimizing the bracket in (3.1.9) is an optimal policy for the problem (3.1.6), (3.1.7). Let $L^u(t,\xi(t),y(t))$ denote $L(t,\xi(t),y(t),u(t,y(t)))$ where ξ satisfies (3.1.6) corresponding to the policy μ. We now present the verification theorem.

<u>Theorem 3.1.2</u> Let $\tilde{\xi}$ satisfy (3.1.6) corresponding to any arbitrary fixed policy $\tilde{\mu}$. If $W : I \times R^n \to R^m$ satisfies

(3.1.21) $$\begin{cases} W_t + a_{ij}W_{x_ix_j} + \min\limits_{\mu}\{b_iW_{x_i} + Q^{\mu}W + F^{\mu}\} = 0, & \text{a.e. in } I \times R^n \\ W(T,x) = V(x), & V(x) \text{ as defined in (3.1.10)}, \end{cases}$$

then,

(3.1.22) $$W(t,x,k) \equiv W_k(t,x) \leq E_{t,x,k}\{\int_t^T L(\theta,\tilde{\xi}(\theta),y(\theta),\tilde{\mu}(\theta,y(\theta)))dt$$
$$+ G(\tilde{\xi}(T),y(T),\mu(T,y(T))), \text{ a.e. in } I \times R^n.$$

Moreover if μ^* satisfies

(3.1.23) $\quad Q^{\mu^*}W + F^{\mu^*} = \min_{\mu}\{Q^{\mu}W + F^{\mu}\}$, a.e. in $I \times R^n$

for every $(t,x) \in I \times R^n$, then

(3.1.24) $$W(t,x,k) = E_{t,x,k}\{\int_t^T L^{\mu^*}(\theta,\xi^*(\theta),y(\theta))d\theta + G(\xi^*(T),y(T),\mu^*(T,y(T)))\} \text{ a.e. in } I \times R^n$$

where ξ^* is the solution of (3.1.6) corresponding to $\mu = \mu^*$. Thus μ^* is optimal.

Proof Let W be a solution of (3.1.21) and let $\tilde{\mu}(t)$ be a fixed repair policy. Then

(3.1.25) $$\begin{aligned} &W(T,x,k) - W(t,x,k) \\ &= E_{t,x,k}\{W(T,\tilde{\xi}(T),y(T)) - W(t,\tilde{\xi}(t),y(t))\} \\ &= E_{t,x,k}\int_t^T dW(\theta,\tilde{\xi}(\theta),y(\theta)). \end{aligned}$$

By lemma 3.1.1 we have

(3.1.26) $$\begin{aligned} &W(T,x,k) - W(t,x,k) \\ &\geq E_{t,x,k}\int_t^T\{W_t(\theta,\tilde{\xi}(\theta),y(\theta)) + a_{ij}(\theta)W_{x_ix_j}(\theta,\tilde{\xi}(\theta),y(\theta)) \\ &+ f_i(\theta,\tilde{\xi}(\theta),y(\theta))W_{x_i}(\theta,\tilde{\xi}(\theta),y(\theta)) \\ &+ q^{\tilde{\mu}}_{y(\theta),j}(\theta)W(\theta,\tilde{\xi}(\theta),j)\}d\theta. \end{aligned}$$

Since W satisfies (3.1.21) we have

(3.1.27) $$\begin{aligned} &W_t(t,x,k) + a_{ij}(t)W_{x_ix_j}(t,x,k) + f_i(t,x,k)W_{x_i}(t,x,k) \\ &+ q^{\tilde{\mu}}_{kj}(t,x)W(t,x_{ij}) + F^{\tilde{\mu}}_k(t,x) \geq 0, \end{aligned}$$

$F^{\tilde{\mu}}$ as defined in (3.1.10).

Combining (3.1.26) and (3.1.27) we have

(3.1.28) $W(T,x,k) - W(t,x,k) \geqslant -E_{t,x,k} \int_t^T L^{\tilde{\mu}}(\theta,\tilde{\xi}(\theta),y(\theta))d\theta,$

or equivalently,

(3.1.29) $W(t,x,k) \leqslant E_{t,x,k} \{\int_t^T L^{\tilde{\mu}}(\theta,\tilde{\xi}(\theta),y(\theta))d\theta + G(\tilde{\xi}(T),y(T))\}$

This proves (3.1.21). To prove (3.1.24) we note that if μ^* satisfies (3.1.23) then we have equality in (3.1.29). This completes the proof.

Backward Equations

In the rest of this section, we derive equations for various quantities corresponding to a fixed repair policy μ. The resulting equations are generalizations of the classical backward equations [23].

Theorem 3.1.1 Let ξ satisfy (3.1.6) and let Q^μ be the differential transition matrix of y corresponding to a fixed policy μ. Let $f_0 : I \times R^n \times M \times U^N \to R$ and define $v : I \times R^n \times M \to R$ by

(3.1.30) $v(t,x,k) = E_{t,x,k} \int_t^T f_0(t,\xi(\theta),y(\theta),\mu(\theta,\xi(\theta),y(\theta)))d\theta,$

and define

(3.1.31) $V(t,x) = [v(t,x,1),\ldots,v(t,x,m)]'.$

If $v \in C^{1,2}$, then

(3.1.32) $$\begin{cases} V_t + a_{ij}(t)V_{x_i x_j} + b_i(t,x)V_{x_i} + Q^\mu(t)V + F_0(t,x) = 0, \\ \qquad\qquad\qquad \text{a.e. in } I \times R^n \\ V(T,x) = 0 \end{cases}$$

where $F_0(t,x) \equiv [f_0(t,x,1,\mu(t,x,1)),\ldots,f_0(t,x,m,\mu(t,x,m))]'$, and a_{ij}, b_i are as defined in (3.1.10).

<u>Proof</u> For convenience let

(3.1.33) $$A = E_{t,x,k} \lim_{\Delta t \downarrow 0} \frac{v(t+\Delta t,\xi(t+\Delta t),y(t+\Delta t)) - v(t,\xi(t),y(t))}{\Delta t}$$

By lemma 3.1.1, we have

(3.1.34) $$A = v_t(t,x,k) + a_{ij}(t)v_{x_ix_j}(t,x,k) + f_i(t,x,k)v_{x_i}(t,x,k) + q^{\mu}_{ij}(t)v(t,x,j).$$

From(3.1.33), the expression A can also be written as

$$A = E_{t,x,k} \lim_{\Delta t \downarrow 0} \frac{1}{\Delta t} \cdot$$
$$\cdot \{E_{t+\Delta t,\xi(t+\Delta t),y(t+\Delta t)} \int_{t+\Delta t}^{T} f_0(\theta,\xi(\theta),y(\theta),\mu(t,\xi(\theta),y(\theta))d\theta$$
$$- E_{t,\xi(t),y(t)} \int_t^T f_0(\theta,\xi(\theta),y(\theta),\mu(\theta,\xi(\theta),y(\theta)))d\theta\}$$
$$= E_{t,x,k} \lim_{\Delta t \downarrow 0} \frac{1}{\Delta t} E_{t,\xi(t),y(t)} \{\int_{t+\Delta t}^{T} f_0(\theta,\xi(\theta),y(\theta),\mu(\theta,\xi(\theta),y(\theta)))d\theta$$
$$- \int_t^T f_0(\theta,\xi(\theta),y(\theta),\mu(\theta,\xi(\theta),y(\theta)))d\theta\}$$
$$= -E_{t,x,k} \lim_{\Delta t \downarrow 0} \frac{1}{\Delta t} \int_t^{t+\Delta t} f_0(\theta,\xi(\theta),y(\theta),\mu(\theta,\xi(\theta),y(\theta)))d\theta.$$

By Lebesgue density lemma [14] we have

$$A = -E_{t,x,k} f_0(t,\xi(t),y(t),\mu(t,\xi(t),y(t)))$$

or

(3.1.35) $$A = -f_0(t,x,k,\mu(t,x,k))$$

The backward equation (3.1.32) now follows from (3.1.34) and (3.1.35). This completes the proof.

3.1.3 AN EXAMPLE FOR THE CASE WITHOUT WHITE NOISE —— MACHINE MAINTENANCE PROBLEM

Dynamic programming equation and backward equation have been derived in the last subsection. In the present subsection similar results are presented for the maintenance problem without white noise disturbance. An example will be worked out.

The plant dynamics (3.1.6) contains the Weiner process $w(t)$ which corresponds to white noise in physical applioations. We have shown in the last subsection that the dynamic programming equation is a parabolic partial differential equation.

When white noise is not present in the plant modelling, the dynamics is given by

(3.1.36) $$d\xi(t) = f(t,\xi(t),y(t))dt.$$

The resulting dynamic programming equation is a hyperbolic partial differential equation as shown below.

Problem: Subject to the dynamic constraint (3.1.36), find a feedback maintenance policy $\mu^0 : I \times M \times N \to U$ such that μ^0 is measurable on I and that μ^0 minimizes the cost function (3.1.7).

The dynamic programming equation and backward equation for the above problem can be obtained formally by letting $\sigma(t) = 0$ in the last section. The results are given below.

Let $\xi^{\tilde{\mu}}$ satisfy (3.1.36) corresponding to the policy $\tilde{\mu}$, if $W : I \times R^n \to R^m$ satisfies

$$(3.1.37) \quad \begin{cases} W_t + \min_\mu \{b_i W_{x_i} + Q^\mu W + F^\mu\} = 0, \text{ a.e. in } I \times R^n \\ W(T,x) = V(x), \quad V(x) \text{ as defined in (3.2.2),} \end{cases}$$

then

$$(3.1.38) \quad W(t,x,k) \equiv W_k(t,x) \leq E_{t,x,k}\{\int_t^T L(\theta,\tilde{\xi}(\theta),y(\theta),\tilde{\mu}(\theta,y(\theta))dt$$

$$+ G(\tilde{\xi}(T),y(T)) \text{ , a.e. in } I \times R^n.$$

Moreover if μ^* satisfies

$$(3.1.39) \quad Q^{\mu^*} W + F^{\mu^*} = \min_\mu \{Q^\mu W + F^\mu\}$$

for a.e. $(t,x) \in I \times R^n$, then

$$(3.1.40) \quad W(t,x,k) = E_{t,x,k} \{\int_t^T L^{\mu^*}(\theta,\xi^*(\theta),y(\theta))d\theta$$

$$+ G(\xi^*(T),y(T),\mu(T,y(T)))\},$$

where ξ^* is the solution of (3.1.36) corresponding to $\mu = \mu^*$. Further, let $f_0 : I \times R^n \times M \times U^N \to R$ and define $v : I \times R^n \times M \to R$ and V by

(3.1.41) $v(t,x,k) = E_{t,x,k} \int_t^T f_0(\theta,\xi^*(\theta),y(\theta),\mu^*(\theta,\xi^*(\theta),y(\theta)))d\theta.$

(3.1.42) $v(t,x) = (v(t,x,1),\ldots,v(t,x,m))'.$

If $v \in C^{1,2}$, then it follows that

$$(3.1.43)\quad \begin{cases} V_t + b_i(t,x)V_{x_i} + Q^{\mu^*}(t)V + F_0(t,x) = 0, \text{ a.e. in } I \times R^n \\ V(T,x) = 0 \end{cases}$$

where $F_0(t,x) \equiv [f_0(t,x,1,\mu^*(t,x,1)),\ldots,f_0(t,x,m,\mu^*(t,x,m))]'$ and b_i is as defined in (3.1.10).

In the rest of this subsection we work out a simple example of an optimal machine maintenance problem by an application of the results above. Closed form solutions are obtained. Only a single machine is involved in the following example; therefore we will write, for brevity, $\mu(t,i)$ instead of $\mu(t,i,k)$ $(k \equiv 1)$.

EXAMPLE

The production rate ξ of a single machine when it is operating $(y(t) = 1)$ and down $(y(t) = 2)$ are c and 0 respectively. ξ is constant between successive transitions of y, i.e.,

(3.1.44) $d\xi(t) = 0$, for a.e. $t \in [0,T]$

The differential transition matrix of y is assumed to be

$$(3.1.45)\quad Q^{\mu}(t) = \begin{bmatrix} -\lambda - a\mu(t,1) & \lambda + a\mu(t,1) \\ \mu(t,2) & -\mu(t,2) \end{bmatrix},$$

where for each $t \in [0,T]$, the maintenance effort $\mu(t,1)$ and $\mu(t,2)$ takes values in $[0,\mu_{1m}]$ and $[0,\mu_{2m}]$ respectively. We want to find maintenance policies $\mu_1(t) \equiv \mu(t,1)$ and $\mu_2(t) \equiv \mu(t,2)$ to minimize the cost function

$$J^{\mu} = E_{0,x_0,y_0}\{\int_0^T [c_1(\xi(t) - D)^2 + c_2\mu(t,y(t))]dt + G(y(T))\} \tag{3.1.46}$$

where $x_0 = \xi(0)$, $y_0 = y(0)$ and D are given constants and

$$G(y(T)) = \begin{cases} 0, & \text{if } y(T) = 1, \\ g, & \text{if } y(T) = 2, \end{cases} \tag{3.1.47}$$

$g \geqslant 0$. For simplicity let $a = 0$ and $D = c$, $c_2 \geqslant g$ (the case $c_2 < g$ can be treated accordingly).

The dynamic programming equation is

$$\begin{cases} \phi_t^1 + \min\limits_{\mu_1} \{-\lambda\phi^1 + \lambda\phi^2 + c_2\mu_1\} = 0 \\ \phi_t^2 + \min\limits_{\mu_2} \{\mu_2\phi^1 - \mu_2\phi^2 + c_1c^2 + c_2\mu_2\} = 0 \\ \phi^1(T) = 0, \quad \phi^2(T) = g, \end{cases} \tag{3.1.48}$$

where $\phi = [\phi^1,\phi^2]$ is the expected future cost, since ϕ from (3.1.48) depends only on t, we have,

$$(3.1.49)\quad \begin{cases} \phi^1(t) = E_{t,0,1}\{\int_t^T c_1(\xi(\theta) - D)^2 + c_2\mu(\theta,y(\theta))d\theta + G(y(T))\} \\ \phi^2(t) = E_{t,0,2}\{\int_t^T c_1(\xi(\theta)-D)^2 + c_2\mu(\theta,y(\theta))d\theta + G(y(T))\} \end{cases}$$

The bang-bang type optimal policy obtained from the minimizations appearing in (3.1.48) is

$$(3.1.50)\quad \begin{cases} \mu_1^* = 0, \\ \mu_2^* = \frac{\mu_{2m}}{2}[1 - \text{sign}\ (c_2 - \phi^2 + \phi^1)] \end{cases}$$

The solution to the problem (3.1.44) (3.1.45) (3.1.46) depends on the values of various constants. There are two possible cases

case (i) $\quad c_2 \geqslant g, \qquad c_2 \geqslant c_1c^2/\lambda,$

case (ii) $\quad c_2 \geqslant g, \qquad c_2 < c_1c^2/\lambda.$

In case (i), $\phi^2 - \phi^1$ can be obtained from (3.1.48) to be less than c_2 for $t \in I$. Therefore from (3.1.50), $\mu_1^* = \mu_2^* = 0$. Let us consider case (ii). Here the labor rate c_2 is smaller (of case (i)); and we assume that the terminal cost g is large enough so that

$$(g - c_1c^2/\lambda)\ e^{-\lambda T} > c_2 - c_1c^2/\lambda$$

(otherwise $\mu_1^* = \mu_2^* = 0$). From (3.1.48) and (3.1.50) we obtain for $t \in [0,T]$

$$(3.1.51)\quad \begin{cases} \mu_1^*(t) = 0 \\ \mu_2^*(t) = \begin{cases} \mu_{2m} \text{ if } t \in [0,\tau] \\ 0 \text{ otherwise} \end{cases} \end{cases}$$

where τ is a 'switching time',

$$(3.1.52)\quad \tau = T - \frac{1}{\lambda}\log_e \frac{c_1 c^2 - \lambda g}{c_1 c^2 - \lambda c_2}$$

It is observed that τ increases with the terminal cost g so that the terminal cost can be minimized. On the other hand τ decreases with labor rate c_2 so that the labor cost can be kept low. Also τ does not depend on the value of μ_{2m}.

The solution to the dynamic programming equation (3.1.48) is

$$(3.1.53)\quad \phi^1(t) = \begin{cases} K_1 + K_2 t + K_3 e^{-(\lambda+\mu_{2m})(\tau-t)}, \text{ for } t \in [0,\tau), \\ K_4 + K_5 t - (K_4 + K_5 T)e^{-\lambda(T-t)}, \text{ for } t \in [\tau,T], \end{cases}$$

$$(3.1.54)\quad \phi^2(t) = \begin{cases} \phi^1(t) - K_2/\lambda + K_6 e^{-(\lambda+\mu_{2m})(\tau-t)}, \text{ for } t \in [0,\tau), \\ \phi^1(t) - K_5/\lambda + (K_4 + K_5 T)e^{-\lambda(T-t)}, \text{ for } t \in [\tau,T]. \end{cases}$$

where K_i, $i = 1,\ldots,6$ are constants given by

$$
(3.1.55)\quad
\begin{cases}
K_1 = g - c_2 + c_1c^2T - \dfrac{(c_1c^2-c_2\lambda)(\lambda+\mu_{2m}^2\tau+\lambda\mu_{2m}\tau)}{(\tau+\mu_{2m})^2}, \\[2ex]
K_2 = -\lambda(c_1c^2+c_2\mu_{2m})/(\lambda+\mu_{2m}), \\[1ex]
K_3 = \lambda(c_1c^2-c_2\lambda)/(\lambda+\mu_{2m})^2,\ K_4 = c_1c^2(T-\frac{1}{\lambda}) + g, \\[1ex]
K_5 = -c_1c^2,\ K_6 = K_3(\lambda+\mu_{2m})/\lambda.
\end{cases}
$$

The expected future costs ϕ^1 and ϕ^2 given by (3.1.53) and (3.1.54) are positive and monotonically decreasing in time. For each t, $\phi^2(t) \geqslant \phi^1(t)$. This demonstrates the higher expected cost associated with the machine failure, If the final time $T \to \infty$ (therefore $\tau \to \infty$), then

$$
(3.1.56)\quad
\begin{cases}
\lim\limits_{T\to\infty} \dfrac{1}{T-t}\,\phi^1(t) = \dfrac{\lambda}{\lambda+\mu_{2m}}\,(c_1c^2 + c_2\mu_{2m}), \\[3ex]
\lim\limits_{T\to\infty} \dfrac{1}{T-t}\,\phi^2(t) = \dfrac{\lambda}{\lambda+\mu_{2m}}\,(c_1c^2 + c_2\mu_{2m}),
\end{cases}
$$

are the long term cost per unit time which does not depend on time and the present value of y. The quantity $\lambda/(\lambda+\mu_{2m})$ in (3.1.56) is the percentage of failure time in the long run as will be shown in (3.1.62) below.

From (3.1.4), the transition probability for y satisfies

$$
(3.1.57)\quad \begin{cases} \dfrac{d}{dt} P^{\mu}(s,t) = P^{\mu}(s,t)Q^{\mu}(t) \\ P^{\mu}(s,s) = I \\ Q^{\mu}(t) = \begin{cases} \begin{bmatrix} -\lambda & \lambda \\ \mu_{2m} & -\mu_{2m} \end{bmatrix} & \text{for } t \in [0,\tau) \\ \begin{bmatrix} -\lambda & \lambda \\ 0 & 0 \end{bmatrix} & \text{for } t \in [\tau,T] \end{cases} \end{cases}
$$

Solving (3.1.57) we have

(i) For $0 < s,\ t < \tau$,

$$
(3.1.58)\quad \begin{cases} P_{11}(s,t) = \dfrac{\mu_{2m}}{\lambda+\mu_{2m}} + \dfrac{\lambda}{\lambda+\mu_{2m}} e^{-(\lambda+\mu_{2m})(t-s)}, \\ P_{22}(s,t) = \dfrac{\lambda}{\lambda+\mu_{2m}} - \dfrac{\mu_{2m}}{\lambda+\mu_{2m}} e^{-(\lambda+\mu_{2m})(t-s)}, \end{cases}
$$

(ii) For $0 \leqslant s \leqslant \tau \leqslant t \leqslant T$,

$$
(3.1.59)\quad \begin{cases} P_{11}(s,t) = \left[\dfrac{\mu_{2m}}{\lambda+\mu_{2m}} + \dfrac{\lambda}{\lambda+\mu_{2m}} e^{-(\lambda+\mu_{2m})(\tau-s)}\right] e^{-\lambda(t-\tau)}, \\ P_{22}(s,t) = 1 - \dfrac{\mu_{2m}}{\lambda+\mu_{2m}} e^{-\lambda(t-\tau)}\left(1-e^{-(\lambda+\mu_{2m})(\tau-s)}\right), \end{cases}
$$

(iii) For $\tau \leqslant s,\ t \leqslant T$,

$$
(3.1.60)\quad \begin{cases} P_{11}(s,t) = e^{-\lambda(t-s)}, \\ P_{22}(s,t) = 1, \end{cases}
$$

where $P_{12} = 1 - P_{11}$ and $P_{21} = 1 - P_{22}$.

It is difficult to calculate the expected down time or operating time of the system from (3.1.58)——(3.1.60). However these quantities can be calculated easily using the backward equation (3.1.43). Define the expected operating time by

$$(3.1.61) \qquad P^k(t,x) = E_{t,x,k} \int_t^T \delta(y(\theta),1)dt,$$

$$\delta(i,j) = \begin{cases} 1 \text{ if } i = j, \\ 0 \text{ if } i \neq j. \end{cases}$$

Then from (3.1.43) P^k, $k = 1,2$, satisfy

$$(3.1.62) \quad \begin{cases} P_t^1 - \lambda(P^1 - P^2) + 1 = 0, \\ P_t^2 + \mu_2^*(P^1 - P^2) = 0, \\ P^1(T,x) = P^2(T,x) = 0. \end{cases}$$

The solution to (3.1.62) is

$$(3.1.63) \qquad P^1(t,x) = \begin{cases} K_7 + K_8(\tau - t) + K_9 e^{-(\lambda+\mu_{2m})(\tau-t)}, \ t \in [0,\tau), \\ \frac{1}{\lambda}(1 - e^{-\lambda(T-t)}), \ t \in [\tau,T]. \end{cases}$$

$$(3.1.64) \qquad P^2(t,x) = \begin{cases} K_8(\tau - t) + K_{10}\left[1 - e^{-(\lambda+\mu_{2m})(\tau-t)}\right], \ t \in [0,\tau), \\ 0, \ t \in [\tau,T]. \end{cases}$$

where

$$(3.1.65)\quad \begin{cases} K_7 = \dfrac{1}{\lambda+\mu_{2m}} + \dfrac{\mu_{2m}}{\lambda(\lambda+\mu_{2m})}\left[\dfrac{\mu_{2m}}{\lambda+\mu_{2m}} - e^{-\lambda(T-\tau)}\right], \\ K_8 = \dfrac{\mu_{2m}}{\lambda+\mu_{2m}},\ K_9 = \dfrac{1}{\lambda+\mu_{2m}}\left[\dfrac{\mu_{2m}}{\lambda+\mu_{2m}} - e^{-\lambda(T-\tau)}\right], \\ K_{10} = \mu_{2m}K_9/\lambda. \end{cases}$$

The expected down time

$$\tilde{P}^k(t,x) = E_{t,x,k} \int_t^T \delta(y(\theta),2)d\theta,\ k = 1,2$$

can be calculated from (3.1.63) and (3.1.64) in fact

$$\tilde{P}^k(t,x) = E_{t,x,k} \int_t^T [1 - \delta(y(\theta),1)]d\theta$$

or

$$(3.1.66)\quad \tilde{P}^k(t,x) = T - t - P^k(t,x).$$

When $T \to \infty$ (therefore $\tau \to \infty$), the fraction of down time is

$$(3.1.67)\quad \lim_{T\to\infty} \frac{\tilde{P}^1(t,x)}{T-t} = \lim_{T\to\infty} \frac{\tilde{P}^2(t,x)}{T-t} = \frac{\lambda}{\lambda+\mu_{2m}},$$

i.e. the system will be down for $\lambda/(\lambda+\mu_{2m}) \times 100\%$ of the remaining time. Intuitively we can write

$$(3.1.68)\quad \lim_{\substack{T\to\infty \\ \tau\to\infty}} \frac{1}{T-t}\phi^k(t) = \text{(Fraction of down time} \times \text{Cost per unit time when system is down)} - \text{(Fraction of operating time} \times \text{Cost per unit time when system is operating)}$$

$$= \frac{\lambda}{\lambda+\mu_{2m}}(c_1c^2+c_2\mu_{2m}),\ k = 1,2$$

The expression (3.1.68) is the same as that in (3.1.56) obtained earlier.

Finally we calculate the expected maintenance cost $R^k(t,x)$, $k = 1,2$, defined by

(3.1.69) $$R^k(t,x) = E_{t,x,k} \int_t^T c_2\mu^*(\theta,y(\theta))d\theta, \quad k = 1,2.$$

$R^k(t,x)$ satisfies the backward equation (3.1.43) given by

(3.1.70) $$\begin{cases} R_t^1 - \lambda(R^1-R^2) = 0, \\ R_t^2 + \mu_2^*(R^1-R^2) + c_2\mu_2^* = 0 \\ R^1(t,x) = R^2(t,x) = 0 \end{cases}$$

Solving (3.1.70) we obtain

(3.1.71) $$R^1(t,x) = \begin{cases} K_{11}(\tau-t) - K_{12}[1-e^{-(\lambda+\mu_{2m})(\tau-t)}], & t \in [0,\tau) \\ 0, & t \in [\tau,T] \end{cases}$$

(3.1.72) $$R^2(t,x) = \begin{cases} K_{11}(\tau-t) + K_{13}[1-e^{-(\lambda+\mu_{2m})(\tau-t)}], & t \in [0,\tau) \\ 0, & t \in [\tau,T] \end{cases}$$

where

(3.1.73) $$\begin{cases} K_{11} = c_2\lambda\mu_{2m}/(\lambda+\mu_{2m}), \quad K_{12} = K_{11}/(\lambda+\mu_{2m}), \\ K_{13} = \mu K_{12}/\lambda. \end{cases}$$

This concludes the example on the optimal maintenance problem for a single machine. Closed form solutions have been obtained because of the simplicity of the situation.

In the next subsection, we present an example of an investor's problem; in addition to a jump parameter, white noise is also present in the dynamics. Part of the optimal investment policy will be obtained in closed form; and the other part can be obtained by solving a system of ordinary differential equations.

3.1.4 AN EXAMPLE FOR THE CASE WITH WHITE NOISE — AN INVESTOR'S PROBLEM

Suppose an investor wants to put his capital into two kind of investments. One kind provides a relatively steady rate of return (e.g. long-term deposit), the other yields a higher return in the long run but has a fluctuating rate of return. The investor wants to find a strategy of investment that allows him to consume his wealth at a reasonably fast rate for as long as possible.

The mathematical problem is now formulated. The two types of investments are:

a) A 'risk-free' type having the price dynamics

$$dP_1 = rP_1\, dt, \tag{3.1.73}$$

where $r(t)$ is a homogeneous Markov chain with differential transition matrix $Q = \{q_{ij}\}$ and the finite state space $\{r_1, r_2, \ldots, r_N\}$,

b) A 'risky' type having the price dynamics

(3.1.74) $$dP_2 = \alpha dt + \sigma dw,$$

where $\alpha, \sigma > 0$ and w is a standard Wiener process. It is assumed that the investor can switch his investment from one type to another ~~without~~ the penalty of transaction expenses.

Denote by $\xi(t)$ his 'wealth' at t. The controls at his disposal are:

(i) u_1, the fraction of investment in the risky asset, takes values in $[0,1]$, and

(ii) u_2, the consumption rate, takes values in $[0,\infty)$.

From (3.1.73) and (3.1.74), his wealth is governed by the following 'budget equation'.

(3.1.75) $$d\xi = [(1-u_1)\xi r + u_1\xi\alpha - u_2]dt + \sigma u_1 \xi dw,$$

where $t \in [0,T]$ and T is a constant.

The performance index to be maximized is

(3.1.76) $$J = E\int_0^{\tau\wedge T} e^{-\rho t} [u_2(\theta)]^\gamma d\theta$$

where $\rho > 0$ and $\gamma \in (0,1)$ and $\tau \equiv \min\{t\in [0,T] \mid \xi(t) < 0\}$.

The controls are allowed to depend on time t, the wealth at t, and the state $r(t)$, i.e.,

$$u_1 = u_1(t,\xi(t),r(t)),$$

$$u_2 = u_2(t,\xi(t),r(t)).$$

The optimal control is assumed to exist and is denoted by $u_1^*(t,x,r_k)$ and $u_2^*(t,x,r_k)$ where $\xi(t) = x$ and $r(t) = r_k$.

Let $W^k(t,x) \equiv W(t,x,r_k)$ be the value function, i.e.,

$$(3.1.77) \quad W^k(t,x) = \sup_{u_1,u_2} E\{\int_t^{\tau\wedge T} e^{-\rho t}[u_2(\theta)]^\gamma d\theta \,|\, \xi(t)=x,\ r(t)=r_k\}.$$

To simplify the notation let $u_i^k(t,x) \equiv u_i(t,x,r_k)$, $i = 1,2$. The dynamic programming equation is obtained in a similar way as in subsection 3.1.2:

$$(3.1.78) \begin{cases} W_t^k + \max\limits_{\substack{0 \leq u_1 \leq 1 \\ u_2 \in [0,\infty)}} \{\tfrac{1}{2}\sigma^2 (u_1^k)^2 x^2 W_{xx}^k + [(1-u_1^k)xr^k + u_1^k x\alpha \\ \quad - u_2^k] W_x^k + e^{-\rho t}(u_2^k)^\gamma\} + q_{kj} W^j = 0, \\ W^k(t,x)|_{t=T} = 0,\ x \in [0,\infty), \\ W^k(t,x)|_{x=0} = 0,\ t \in [0,\infty). \end{cases}$$

For future references, we partition the state space of r into three subsets:

$$(3.1.79) \begin{cases} \{r_1,\ldots,r_{N_1}\} = \{r_k | r_k < \alpha - \sigma^2(1-\gamma)\}, \\ \{r_{N_1+1},\ldots,r_{N_1+N_2}\} = \{r_k | \alpha - \sigma^2(1-\gamma) \leq r_k \leq \alpha\}, \\ \{r_{N_1+N_2+1},\ldots,r_N\} = \{r_k | r_k > \alpha\}. \end{cases}$$

To obtain solution of the dynamic programming equation (3.1.78), we can attempt the trial solution,

$$(3.1.80) \quad W^k(t,x) = g_k(t)x^\gamma, \quad k = 1,2,\ldots,N$$

After substituting this trial solution into (3.1.78), simplifying and rearranging terms, we obtain an equation of the form

$$G_k(t)x^\gamma = 0 \qquad k = 1,\ldots,N$$

which holds for all $x \geqslant 0$. Thus $G_k(t) = 0$ for $t \in [0,T]$ and $k = 1,\ldots,N$. It follows from $G_k(t) \equiv 0$ that

$$(3.1.81) \quad \begin{cases} \dot{g}_k + \nu_k g_k + (1-\gamma)(e^{\rho t} g_k^\gamma)^{\frac{1}{r-1}} + q_{kj} g_j = 0, \\ g_k(T) = 0 \end{cases}$$

where

$$(3.1.82) \quad \nu_k = \begin{cases} -\frac{1}{2}\sigma^2\gamma(1-\gamma), & k < N_1, \\ \dfrac{\gamma(r_k-\alpha)^2}{\sigma^2(\gamma-1)} + r_k\gamma, & k = N_1+1,\ldots,N_1+N_2, \\ \gamma r_k, & k > N_1+N_2, \end{cases}$$

The resulting optimal controls are

$$(3.1.83) \begin{cases} u_1^{k*}(t) = \begin{cases} 1, & k \leqslant N_1, \\ \dfrac{\alpha - r_k}{\sigma^2(1-\gamma)}, & k = N_1+1,\ldots,N_1+N_2, \\ 0, & k > N_1+N_2 \end{cases} \\ \\ u_2^{k*}(t) = [e^{\rho t} g_k(t)]^{\frac{1}{\gamma-1}} x, \quad k = 1,2,\ldots,N. \end{cases}$$

Denoting $[e^{\rho t}g_k(t)]^{\frac{1}{1-\gamma}}$ by $h_k(t)$. Then from (3.1.81)we obtain a system of ordinary differential equations for h_k

$$(3.1.84)\quad \begin{cases} \dot{h}_k + \beta_k h_k + \frac{q_{kj}}{1-\gamma} h_j^{1-\gamma} h_k + 1 = 0, \\ h_k(T) = 0,\ k = 1,2,\ldots,N, \end{cases}$$

$$(3.1.85)\quad \beta_k = \frac{\nu_k - \rho}{1-\gamma},\ k = 1,2,\ldots,N,$$

and $u_1^{k*}(t)$ is given as in (3.1.83) and

$$(3.1.86)\quad u_2^{k*}(t) = h_k^{-1}x,\ k = 1,2,\ldots,N.$$

The optimal controls are now obtainable by solving the system (3.1.84) of ordinary differential equations. The corresponding performance index is obtained from (3.1.80)

Remarks

(i) When the average rate of return of the risk-free assets is less than that of the risky asset, i.e., $r(t) < \alpha$, the optimal investment policy does not call for total investment in the risky assets. This demonstrates the risk-aversion property of the performance criterion. However, when $r(t) > \alpha$, the optimal policy is to invest totally in risk-free assets, independent of the differential transition matrix Q. In this case the optimal policy ignores the 'risks' involved in risk-free assets.

(ii) The optimal consumption rate given by (3.1.86) is a linear functional of x [41]. This is true for utility functions of the type

$$(3.1.87) \quad \frac{1-\gamma}{\gamma} e^{-\rho t} \left(\frac{\beta u_2(t)}{1-\gamma} + \eta\right)^{\gamma},$$

where $\gamma \neq 1$, $\beta > 0$, $\left(\frac{\beta u_2(t)}{1-\gamma} + \eta\right) > 0$; $\eta = 1$ if $\gamma = -\infty$.

(iii) The case for $\gamma = 0$ corresponds to maximizing the exit time τ. Using the trial solution (3.1.80) in the dynamic programming equation (3.1.78), we arrive at

$$(3.1.88) \quad \begin{cases} \dot{g}_k + q_{kj} g_j + e^{-\rho t} = 0, \\ g_k(T) = 0, \ k = 1,2,\ldots,N. \end{cases}$$

In this case the problem is singular and the dynamic programming approach does not yield any information on the optimal control.

If $\gamma > 1$ the maximization

$$(3.1.89) \quad \max_{u_2^k \geq 0} \{e^{-\rho t}(u_2^k)^{\gamma} - u_2^k \gamma x^{\gamma-1} g_k(t)\}$$

yields $u_2^k = \infty$. Although the consumption rate $u_2^k = \infty$ in

(3.1.89) is not practical, $u_2^k(t) = \delta(a)$ for fixed a is physically meaningful. It would be an interesting problem to consider $u_2^k(t)$ as a linear combination of Dirac delta functions whose position and magnitude are to be chosen.

Comments

In this section we have considered the dynamic programming approach for solving certain stochastic optimal control problems which involves jump parameters. This approach is applicable to more general problems; the dynamic programming equation for a more general problem will be formally presented in section 3.3.

In the two examples (optimal maintenance problem and investor's problem), the dynamic programming approach leads to closed form solutions. For general problems, however, one has to solve a system of non-linear parabolic differential equations. Such numerical problems have seldom been attempted, and the non-linearities are expected to cause numerical difficulties. In the following section, we derive optimality conditions for a general problem. These optimality conditions leads to a gradient method which involve systems of linear parabolic differential equations.

3.2 VARIATIONAL APPROACH

In the last section we have considered the dynamic programming approach for solving stochastic differential systems with jump parameters; in this section we will consider the variational appraoch. In a manner similar to that of section 2.2, it will be shown that the variational approach leads to a gradient method for computing the optimal control. It will be shown that the gradient method requires the numerical solution of a system of linear parabolic equations; the dynamic programming method, a system of non-linear parabolic equations.

In subsection 3.2.1, we formulate the control problem, which is essentially the problem considered in section 2.1, but with jump parameter included in the formulation. In subsection 3.2.2 we derive optimality conditions which consist of parabolic system, the associated adjoint (parabolic) system, and two inequalities, one in a pointwise form and the other in a integral form. A more general theorem on the existence of solutions to parabolic systems (covering the two parabolic systems in the optimality conditions) is proved. The optimality conditions are then applied in subsection 3.2.3 to a linear quadratic Gaussian problem. In this special case, the optimality conditions leads to a set of ordinary differential equations, whose solution is used to compute the optimal control and the optimal cost. In subsection 3.2.4, we consider the gradient method of computing the optimal control for the problem formulated in subsection 3.2.1.

3.2.1 PROBLEM FORMULATION

Let $\mathcal{N} = \{z_1, z_2, \ldots, z_N\}$ be a finite set of points in R^m; denote by $\{y(t),\ t \in [0,T]\}$ a homogeneous Markov chain with state space $\mathcal{N}$, initial distribution P_0, and transition matrix P ~~with~~ elements

$$P_{ik}(t) = \text{prob}\{y(t) = z_k \mid y(0) = z_i\}$$

$(i, k = 1,2,\ldots,N)$. Define

$$q_{ik} \underline{\Delta} \lim_{t \downarrow 0} \frac{P_{ik}(t)}{t}\ , \quad i \neq k$$

$$q_{ii} \underline{\Delta} \lim_{t \downarrow 0} \frac{P_{ii}(t) - 1}{t} = -\sum_{k \neq i} q_{ik}$$

It is known that P satisfies the differential equation

$$\frac{dP}{dt} = PQ\ ,\ P(0) = p^0 \text{ and } Q = \{q_{ik}\}$$

Denote the time interval $(0,T)$ by I. Let $w(t)$, $t \in I$ be the standard Wiener process with values in R^d. Suppose U is any closed convex set in R^r. Consider the stochastic differential system in R^n given by

$$(3.2.1)\quad \begin{cases} d\xi = f(\xi,t,y,u)dt + g(\xi,t,y)dw \\ \xi(0) = \xi_0, \qquad \text{with distribution } \pi \end{cases}$$

where u is any measurable function on I with values in U. Let B be an open bounded and connected set in R^n with boundary ∂B. Denote the cylinder $B \times (0,T)$ by Q_T and the lateral boundary $\partial B \times (0,T)$ by S, and let τ be the first time the process ξ hits the set ∂B. Define $\mathcal{U}$ to be the class of admissible control laws consisting of functions u defined on $\bar{B} \times I \times \mathcal{N}$ with values in U so that for each $z \in \mathcal{N}$, $u(\cdot,\cdot,z)$ is measurable on $B \times I$. For any function h defined on $B \times I \times \mathcal{N}$ we will write, for brevity, $h(t,x,k)$ for $h(t,x,z_k)$. Let $f_0 : R^n \times I \times \mathcal{N} \times U \to R^1$ so that for each arbitrary but fixed point $z \in \mathcal{N}$ and $u \in U$, f_0 is measurable in $(x,t) \in R^n \times I$, and for each $(x,t,z) \in R^n \times I \times \mathcal{N}$, f_0 is continuous in u on U.

The problem is to find a control law $u_0 \in \mathcal{U}$ so that the cost functional

$$(3.2.2) \qquad J(u) = E \int_0^{\tau\wedge T} f_0(\xi(t),t,y(t),u(t))dt$$

attains its minimum at $u(t) = u_0(\xi(t),t,y(t))$ subject to the dynamic constraints (3.2.1). Call this problem P_0.

3.2.2 OPTIMALITY CONDITIONS

The optimization problem P_0 of the last subsection can be transformed into an equivalent problem involving a system of partial differential equations of parabolic type with first boundary conditions. Define

$$(3.2.3) \qquad \phi(x,s,k) = E \{\int_s^{T\wedge\tau} f_0(\xi(\theta),\theta,y(\theta),u(\xi(t),\theta,y(\theta))d\theta \mid \xi(s) = x, y(s) = z_k\}$$

and denote by Φ and $F(u)$, the vectors $(\phi(\cdot,\cdot,1),\ldots,\phi(\cdot,\cdot,N))'$ and $(f_0(\cdot,\cdot,1,u),\ldots,f_0(\cdot,\cdot,N,u))'$. Redefine $\mathcal{U}$ to be the class of admissible controls consisting of functions $\{u(x,t)\}$ where $u(x,t) = (v_1(x,t,1),\ v_2(x,t,1),\ldots,v_r(x,t,1);\ldots;\ v_1(x,t,N),\ v_2(x,t,N),\ldots,\ v_r(x,t,N))'$ defined and measurable on Q_T with $v(x,t,k) = \{v_i(x,t,k),\ i = 1,\ldots,r\}$, $k = 1,\ldots,N$, taking values in U. In other words u is a measurable function on Q_T with values in U^N (N copies of U), which is contained in R^{rN}.

Using the convention of repeated indices for summation, we define the operator $A(u)$, for $u \in \mathcal{U}$, by

$$A(u)\Phi = A_{ij}\Phi_{x_i x_j} + B_i(u)\Phi_{x_i} + C\Phi \tag{3.2.4}$$

or in the divergence form,

$$A(u)\Phi = (A_{ij}\Phi_{x_i})_{x_j} - \bar{B}_i(u)\Phi_{x_i} + C\Phi \tag{3.2.5}$$

where $\bar{B}_i(u) \triangleq \partial A_{ij}/\partial x_j - B_i(u)$ and

$$\left\{\begin{aligned} &A_{ij} = \mathrm{diag}(a^1_{ij},\ a^2_{ij},\ldots,a^N_{ij}) \\ &\qquad a^k_{ij}(x,t) = \tfrac{1}{2}(gg')_{ij}(x,t,k) \\ &B_i(u) = \mathrm{diag}(f^1_i,\ f^2_i,\ldots,\ f^N_i) \\ &\qquad f^k_i(x,t) = f_i(x,t,k,v(x,t,k)) \\ &C = Q \end{aligned}\right. \tag{3.2.6}$$

Lemma 3.2.1

Suppose for each control $u \in \mathcal{U}$ the system (3.2.1) has a solution ξ. Then the optimization problem P_0 is equivalent to the following problem P_1:

$$(3.2.7) \qquad P_1 \begin{cases} \Phi_t + A(u)\Phi + F(u) = 0, & (x,t) \in B \times [0,T) \\ \Phi(x,t) = 0, & (x,t) \in \partial B \times (0,T) \\ \Phi(x,T) = 0, & x \in B \end{cases}$$

$$J(u) = \int_B (\Phi(x,0)\,,\ p^0)\,\pi(dx) = \min_{u \in \mathcal{U}}\ ; \quad p^0 = \mathrm{Prob}\{y(0) = Z_k\}$$

where Φp^0 denotes the scalar product of Φ and p^0 in R^N.

The proof of the above lemma follows from the backward equations of section 3.1.2.

By lemma 3.2.1 we have converted the original stochastic optimization problem P_0 into an optimal control problem which consists of a deterministic parabolic system with first boundary conditions. From this point on, we will consider only the problem P_1.

In order that the parabolic system (3.2.7) be mathematically meaningful, it is necessary to prove, under some appropriate conditions, the existence of solutions to (3.2.7). We now present sufficient conditions for the existence of weak solution of a more general problem, which covers (3.2.7) as a special case. The result is also applicable to the adjoint system (of (3.2.7)), which will also appear in the optimality conditions to be presented later.

Existence of Solutions of a Class of General Parabolic Systems and Associated Adjoints

The more general problem just mentioned is:

$$(3.2.8)\quad \begin{cases} \Phi_t + [A_{ij}(x,t)\Phi_{x_i} + A_j(u)(x,t)\Phi]_{x_j} + B_i(u)(x,t)\Phi_{x_i} \\ \quad + C(u)(x,t)\Phi + F(u)(x,t) = 0, \ (x,t) \in B \times [0,T) \\ \Phi(x,t)|_{S_T} = 0 \\ \Phi(x,T) = \Phi_T(x), \ x \in B . \end{cases}$$

The result will cover the system (3.2.7) and its associated adjoint problem. First, we introduce some notations.

The Euclidean inner product in R^N will be denoted by vw; the Euclidean norm by $|v|$; the integral $\int_B vw\, dxdt$ by $\langle v,w\rangle_B$ and the integral $\int_{Q_T} vw\, dxdt$ by $\langle v,w\rangle$. Matrix norms are defined as $|\cdot|$, where

$$(3.2.9)\qquad |A|^2 = \sum_{k,\ell=1}^{N} |A_{k\ell}|^2, \text{ and } |v_x|^2 = \sum_{k=1}^{N}\sum_{i=1}^{n}\left[\frac{\partial v_k}{\partial x_i}\right]^2$$

The symbol $R^N \otimes R^N$ will denote the space of $N \times N$ matrices and "'" will denote matrix transpose. The complement of a set B will be denoted by $\complement B$.

The notations for function spaces are as follows. $W_2^{1,0}(Q_T)$ is the Hilbert space of R^N-valued functions with scalar product

$$(\phi,\psi)_{W_2^{1,0}(Q_T)} = \int_{Q_T} \phi\psi + \phi_{x_i}\psi_{x_i}\, dx\, dt$$

$V_2(Q_T)$ will denote the Banach space consisting of all elements in $W_2^{1,0}(Q_T)$ with the norm defined by

$$|\phi|_{Q_T} = \operatorname*{ess\,sup}_{0 \leq t \leq T} ||\phi(\cdot,t)||_{2,B} + ||\phi_x||_{2,Q_T}.$$

where

$$||\phi_x||_{2,Q_T} = \left(\int_{Q_T} |\phi_x|^2 dx\, dt\right)^{\frac{1}{2}}$$

$L_{q,r}(Q_T)$ is the Banach space consisting of all R^N-valued functions on Q_T with a finite norm

$$||\phi||_{q,r,Q_T} = \int_I \left(\int_B |\phi(x,t)|^q dx\right)^{r/q} dt^{\ 1/r}$$

where $q \geq 1$ and $r \geq 1$.

$W_2^1(B)$ is the Banach space consisting of all elements of $L_q(B,R^N)$ having first order derivatives that are square integrable with the norm

$$||\psi||_{W_2^1(B)} = \sum_{j=0}^{1} \langle\langle\psi\rangle\rangle_{2,B}^{(j)}$$

where

$$\langle\langle\psi\rangle\rangle_{2,B}^{(j)} = \sum_{(j)} ||D_x^j\psi||_{2,B}$$

The symbol D_x^j denotes any derivative of $\psi(x)$ with respect to x of order j, where $\sum_{(j)}$ denotes the summation over all possible derivatives of ψ of order j.

$\overset{0}{W}{}_2^1(B)$ is the subspace of $W_2^1(B)$ in which consists of the set of functions that are infinitely differentiable.

The following assumptions will be made on the coefficients of (3.2.8):

(i) $A_{ij} : B \times I \to R^N \otimes R^N$ satisfies the parabolicity condition

$$\nu \sum_{k=1}^{N} \sum_{i=1}^{n} (\xi_i^k)^2 \leq A_{ij}\xi_i\xi_j \leq \mu \sum_{k=1}^{N} \sum_{i=1}^{n} (\xi_i^k)^2$$

for all $\xi_i = (\xi_i^1, \xi_i^2, \ldots, \xi_i^N)'$, $i = 1,2,\ldots,n$, from R^N uniformly with respect to $(x,t) \in Q_T$ for some positive constants ν and μ.

(ii) $A_i(u)$, $B_i(u)$, $C(u) : B \times I \to R^N \otimes R^N$ are measurable functions with norms

$$|| \, |A_i(u)|^2; \; |B_i(u)|^2; \; |C(u)| \, ||_{q,r,Q_T} \leq \mu_1$$

for every $u \in \mathcal{U}$ and q,r satisfying

$$(3.2.10) \quad \begin{cases} 1/r + n/2q = 1 \\ q \in (n/2,\infty], \; r \in [1,\infty] \text{ for } n \geq 2 \\ q \in [1,\infty], \; r \in [1,2] \text{ for } n = 1 \end{cases}$$

(iii) The free term $F(u) : B \times I \to R^N$ has a finite norm

$$||F(u)||_{q_1,r_1,Q_T} \leq \mu_2$$

for every $u \in \mathcal{U}$ and

$$(3.2.11)\quad \begin{cases} 1/r_1 + n/2q_1 = 1 + n/4 \\ q_1 \in [2n/n+2,\ 2],\ r_1 \in [1,2] \text{ for } n \geq 3 \\ q_1 \in (1,2],\ r_1 \in [1,2) \text{ for } n = 2 \\ q_1 \in [1,2],\ r_1 \in [1,\ 4/3] \text{ for } n = 1 \end{cases}$$

A generalized (or weak) solution Φ of (3.2.8) is defined to be an element in $V_2(Q_T)$ that satisfies $\Phi(x,t)|_{S_T} = 0$ and that for any element $\eta(x,t) \in W_2^{1,1}(Q_T)$, $\eta(x,0) = 0$, the following integral identity holds,

$$(3.2.12)\quad \int_{Q_T} \{-\Phi\eta_t - A_{ij}\Phi_{x_i}\eta_{x_j} - A_j(u)\Phi\eta_{x_j} + B_j(u)\Phi_{x_j}\eta + C(u)\Phi\eta + F(u)\eta\}\, dx\, dt = -\int_B \Phi\eta\, dx|_{t=T}$$

For proving the existence of generalized solution for the parabolic system (3.2.8), we need an energy estimate:

<u>Lemma 3.2.2</u> (Energy estimate)

Suppose $\Phi \in V_2(Q_T)$, with $\Phi(x,t)|_{S_T} = 0$ satisfies for almost all t_1, $t_2 \in [0,T]$ the following inequality

$$(3.2.13)\quad \tfrac{1}{2}\int_B |\Phi(x,t)|^2 dx \Big|_{t=t_1}^{t=t_2} - \int_{t_1}^{t_2} \int_B (A_{ij}\Phi_{x_i} + A_j(u)\Phi)\Phi_{x_j} + (B_i(u)\Phi_{x_i} + C(u)\Phi + F(u))\Phi\, dx\, dt \geq 0$$

and A_{ij}, $A_j(u)$, $B_i(u)$, $C(u)$, $F(u)$ satisfy (i) - (iii). Then there exists a constant c depending on n, ν, μ, μ_1, and q so that

$$|\Phi|_{Q_T} \leq c\left[||\Phi(x,T)||_{2,B} + ||F(u)||_{2,Q_T}\right]$$

<u>Proof</u>:

Let $\tau \varepsilon [t_1,t_2]$. From (3.2.13)

$$\tfrac{1}{2}\int_B |\Phi(x,t)|^2 dx\Big|_{t=\tau}^{t=t_2} - \int_\tau^{t_2}\int_B A_{ij}\Phi_{x_i}\Phi_{x_j} dx\, dt + \int_\tau^{t_2}\int_B \{-A_j(u)\Phi\Phi_{x_j}$$

$$+ B_j(u)\Phi_{x_j}\Phi + C(u)\Phi\Phi + F(u)\Phi\}\, dx\, dt \geq 0$$

Using the assumption (i) and Cauchy's inequality in the above expression, we arrive at

$$\int_B |\Phi(x,\tau)|^2 dx + \nu\int_\tau^{t_2}\int_B |\Phi_x|^2 dx\, dt \leq \int_B |\Phi(x,t_2)|^2 dx$$

$$+ 2\int_\tau^{t_2}\int_B \{(\tfrac{1}{\nu} A_j(u)|^2 + \tfrac{1}{\nu}|B_j(u)|^2 + |C(u)|)|\Phi|^2 + |F(u)|\}dx\, dt$$

By taking essential supremum with respect to τ over $[t_1,t_2]$, we obtain

(3.2.14) $$\min(1,\nu)|\Phi|^2_{Q_{t_1,t_2}} \leq ||\Phi(\cdot,t_2)||^2_{2,B} + 2\int_{Q_{t_1,t_2}} \mathcal{D}|\Phi|^2 dx\, dt$$

$$+ 2\int_{Q_{t_1,t_2}} |F(u)\Phi|dx\, dt$$

where

$$\mathcal{D} = \frac{1}{\nu}\sum_{j=1}^{n}\left(|A_j(u)|^2 + |B_j(u)|^2\right) + |C(u)|$$

and

$$|\Phi|_{Q_{t_1,t_2}} = \operatorname*{ess\,sup}_{t_1\le\tau\le t_2} \|\Phi(\cdot,\tau)\|_{2,B} + \left[\int_{t_1}^{t_2}\int_B |\Phi_x|^2 dx\, dt\right]^{\frac{1}{2}}.$$

By applying Cauchy's inequality in (3.2.14) and choosing t_1 and t_2 such that $\min(1,\nu) > 2\beta^2 \|\mathcal{D}\|_{q,r,Q_{t_1,t_2}}$, the lemma follows ([37], lemma 2.1, p.139).

Theorem 3.2.3

If the conditions (i) - (iii) are satisfied and $\Phi_T(x) \in L_2(B,R^N)$, then the system (3.2.8) has a generalized solution in $V_2(Q_T)$.

Proof:

Choose a fundamental system of functions $\psi_k(x)$, $k = 1,2,\ldots$, in W_2^1 (B) so that ψ_k vanishes on S_T for each k and that they are orthonormal in $L_2(B,R^N)$. Let us denote

$$\Phi^m(x,t) = \sum_{k=1}^{m} C_k^m(t)\psi_k(x)$$

where $C_k^m(t) = \int_B \Phi^m(x,t)\psi_k(x)dx$ are determined by a system of ordinary differential equations

$$\frac{d}{dt} C_k^m(t) = \sum_{\ell=1}^{m} \{<A_{ij}(\psi_\ell)_{x_i}, (\psi_k)_{x_j}>_B + <A_j(u)\psi_\ell, (\psi_k)_{x_j}>_B$$

$$- <B_i(u)(\psi_\ell)_{x_i}, \psi_k>_B - <C(u)\psi_\ell, \psi_k>_B\}C^m(t) - <F(u), \psi_k>_B;$$

$$C_k^m(T) = \int_B \Phi_T^m(x)\psi_k(x)\, dx$$

The theorem can then be proved by verifying that the sequence

ϕ^m so constructed has a subsequence which converge to the generalized solution of (3.2.8) [37].

Consider the adjoint problem to (3.2.8) defined by

$$(3.2.15)\quad \begin{cases} -\psi_t + (A'_{ij}\psi_{x_j} - B'_i(u)\psi)_{x_i} - A'_j(u)\psi_{x_j} + C'(u)\psi = 0, \\ \qquad\qquad (x,t) \in B \times [0,T) \\ \psi(x,t)|_{S_T} = 0 \\ \psi(x,0) = \psi_0(x) \end{cases}$$

Similar to the definition of generalized solution of (3.2.8), we define a generalized or weak solution ψ of (3.2.15) to be an element in $V_2(Q_T)$ that satisfies $\psi(x,t)|_{S_T} = 0$ and that for any element $\gamma(x,t) \in W_2^{1,1}(Q_T)$, $\gamma(x,T) = 0$, the following integral identity holds,

$$\int_{Q_T}\{\gamma_t\psi - \gamma_{x_i}A'_{ij}\psi_{x_j} + \gamma_{x_i}B'_i(u)\psi - \gamma A'_j(u)\psi_{x_j} + \gamma C'(u)\psi\}dx\,dt = -\int_B \gamma\psi dx|_{t=0}$$

The following theorem follows from theorem 3.2.3 by simply reversing the flow of time.

<u>Theorem 3.2.4</u>

If the coefficients A_{ij}, $A_j(u)$, $B_i(u)$, $C(u)$ of the adjoint problem (3.2.15) satisfy the assumptions (i) - (iii), and that $\psi_0(x) \in L_2(B,R^N)$, then (3.2.15) has a generalized solution in $V_2(Q_T)$.

Denote the adjoint operator corresponding to $A(u)$ in (3.2.5) by $A^*(u)$, defined as

$$A^*(u)\psi = (A'_{ij}\psi_{x_j} + \bar{B}'_i(u)\psi)_{x_i} + C'\psi$$

where $\bar{B}_i(u) = \partial A_{ij}/\partial x_j - B_i(u)$. Consider the adjoint problem to (3.2.7)

$$(3.2.16)\quad \begin{cases} -\psi_t + A^*(u)\psi = 0, & (x,t) \in B \times (0,T] \\ \psi(x,t)|_{S_T} = 0 \\ \psi(x,0) = \psi_0(x) \underline{\Delta} P^0 q(x), & x \in B \end{cases}$$

where $\int_A q(x)dx = \pi(A)$. The following corollary is an immediate consequence of theorems 3.2.3 and 3.2.4.

Corollary 3.2.5

Suppose $\partial A_{ij}/\partial x_j$ is measurable, $i,j = 1,2,\ldots,n$ and belongs to $L_{q,r}(Q_T)$ with q, r satisfying (3.2.10) and A_{ij}, $B_i(u)$, C, $F(u)$ satisfy the assumptions (i) - (iii) for each fixed but arbitrary $u \in \mathcal{U}$. Then the system (3.2.7) and its adjoint (3.2.16) have weak solutions in $\overset{\circ}{V}_2(Q_T)$.

OPTIMALITY CONDITIONS

Optimality conditions for the problem P_1 (defined in lemma 3.2.1) will now be presented under a more general setting. Assuming that the coefficients and free term satisfy the hypothesis of theorem 3.2.3, we consider the following problem P_2:

$$(3.2.17)\quad \begin{cases} \Phi_t + (A_{ij}\Phi_{x_i} + A_j(u)\Phi)_{x_j} + B_i(u)\Phi_{x_i} + C(u)\Phi \\ \qquad + F(u) = 0, \ (x,t) \in B \times [0,T) \\ \Phi(x,t)|_{S_T} = 0 \\ \Phi(x,T) = \Phi_T(x), \ x \in B \\ J(u) = \int_B \Phi(u)(x,0) p^0 \pi(dx) = \min\limits_{u \in \mathcal{U}} \end{cases}$$

where $\Phi(u)$ denotes the generalized solution of the first boundary value problem (3.2.17) corresponding to $u \in \mathcal{U}$. The following lemma will be needed for the derivation of the optimality conditions.

Lemma 3.2.6

Suppose $A_j^{(k)}$, $B_i^{(k)}$, $C^{(k)}$, $F^{(k)}$, satisfying the assumptions (i) - (iii) independently of k, converge almost everywhere to $A_j(u)$, $B_i(u)$, $C(u)$, $F(u)$, respectively, on Q_T and there exist bounded measurable functions $G \in L_{q,r}(Q_T)$ and $H \in L_{q_1,r_1}(Q_T)$ such that $|A_j^{(k)}|$, $|B_i^{(k)}|$, $|C^{(k)}| \leq |G|$ and $|F^{(k)}| \leq H$ a.e. Further, suppose that $\Phi_T(x) \in L_2(B, R^N)$. Let the sequence $\Phi^k \in V_2(Q_T)$ be the generalized solution of the system (3.2.17) with coefficients $A_j^{(k)}$, $B_i^{(k)}$, $C^{(k)}$, $F^{(k)}$. Then there exists a subsequence of the sequence $\{\phi^k\}$, again denoted by $\{\phi^k\}$, and an element $\Phi \in V_2(Q_T)$ so that ϕ^k and $\phi^k_{x_i}$ converge weakly to Φ and Φ_{x_i}, respectively, in $L_2(Q_T)$ where $\Phi = \phi^u$ is the weak solution of (3.2.17).

Proof:

It follows from the energy inequality in lemma 3.2.2 that $|\phi^k|_{Q_T} \leq \beta$ uniformly for $k = 1,2,\ldots$, where β is independent of k. Therefore, we can select a common subsequence from $\{\phi^k\}$, again denoted by $\{\phi^k\}$, such that ϕ^k and $\phi^k_{x_i}$ converge weakly to Φ and Φ_{x_i}, respectively, in $L_2(Q_T)$. To show that Φ is a weak solution of (3.2.17), it suffices to consider $\eta \in \overset{0}{W}{}^{1,1}_2(Q_T)$ with $\eta(x,0) = 0$ and show that as $k \to \infty$,

$$(3.2.18) \quad \langle \Phi_t - \phi^k_t + (A_{ij}\Phi_{x_i} + A_j(u)\Phi)_{x_j} - (A_{ij}\phi^k_{x_i} + A^k_j\phi^k)_{x_j} + B_i(u)\Phi_{x_i} - B^k_i\phi^k_{x_i} + C(u)\Phi - C^k\phi^k + F(u) - F^k, \eta\rangle \to 0$$

This follows from dominated convergence theorem and the fact that if $f_k \to f_0$ weakly and $g_k \to g_0$ strongly, then $\langle f_k, g_k\rangle \to \langle f_0, g_0\rangle$. This completes the proof.

The proof of the necessary conditions for optimality of problem P_2 is based on approximating the system coefficients by their integral averages and the convergence properties of the resulting coefficients. We need the preparations below.

Following the method in Aronson [8] and Ahmed and Teo [3], [4], we let $F = 0$ in $\complement Q_T$ (complement of Q_T in $R^n \times I$) and $\phi_T = 0$ in $\complement B$ (complement of B in R^n). Let $A_j^{(m)}(u)$, $B_i^{(m)}(u)$, $C^{(m)}(u)$, $F^{(m)}(u)$ denote the integral averages of the corresponding coefficients formed with an averaging kernel whose support lies in $|x|^2 + t^2 < m^{-2}$ for integers $m \geq 1$ and define the operators $L^m(u)$ by

$$L^m(u)\phi = \phi_t + (A_{ij}\phi_{x_i} + A_j^{(m)}(u)\phi)_{x_j} + B_i^m(u)\phi_{x_i} + C^{(m)}(u)\phi$$

Let ϕ_T^m denote the integral average of ϕ_T formed with a kernel whose support lies in $|x| < m^{-1}$. Let $\{B^m\}$ be a sequence of open domains with smooth boundaries such that $\overline{B^m} \subset B^{m+1} \subset \overline{B^{m+1}} \subset B$ for all $m \geq 1$ and $\lim_{m\to\infty} B^m = B$. Let $\{g_m\}$ be a sequence of functions in $C_0^\infty(B^m)$ (the space of infinitely differentiable functions defined and having compact support in B^m) so that $g_m = 1$ on $\overline{B^{m-1}}$ and $0 \leq g_m(x) \leq 1$ for $x \in B^m \setminus B^{m-1}$. Then the sequence of boundary value problems

$$(3.2.19)\quad \begin{cases} L^m(u)\phi + F^m = 0, & (x,t) \in Q_T^m = B^m \times [0,T) \\ \phi(x,T) = \phi_T^m(x)g_m(x), & x \in B^m \\ \phi(x,t) = 0, & (x,t) \in \partial B^m \times [0,T) \end{cases}$$

with $F^m \equiv F^m(u)$ and B^m sufficiently smooth, has a classical solution $\phi^m \in C^1(\overline{Q^m})$. Details on smooth solutions can be found in, for example, Friedman ([24], p.65) and Ladyzenskaja, Solonikov, Uralcera ([37], p.320).

Remark 3.2.7

It follows from lemma 3.2.6 that Φ^m and $\Phi^m_{x_i}$ converge weakly to Φ and Φ_{x_i}, respectively in $L_2(Q_T)$ where Φ^m and Φ are weak solutions of the systems (3.2.19) and (3.2.17).

Define the adjoint problem of (3.2.7) by

$$(3.2.20)\quad \begin{cases} L^*(u_0)\psi = 0, & (x,t) \in B \times (0,T] \\ \psi(x,t)\big|_{S_T} = 0 \\ \psi(x,0) = \psi_0(x) \equiv P^0 q(x), & x \in B,\ \pi(A) = \int_A q(x)\,dx \end{cases}$$

where $L^*(u_0)$ is the formal adjoint of the operator

$$L(u_0)\Phi = \Phi_t + (A_{ij}\Phi_{x_i} + A_j(u_0)\Phi)_{x_j} + B_i(u_0)\Phi_{x_i} + C(u_0)\Phi$$

where u_0 is an optimal control for the problem P_2 (existence assumed). The existence of generalized solution of (3.2.20) follows from theorem 3.2.4. Similar to the integral averaged system (3.2.19), we define the following adjoint problems to (3.2.19),

$$(3.2.21)\quad \begin{cases} L^{*m}(u_0)\psi^m = 0, & (x,t) \in Q_T^m \\ \psi^m(x,t) = 0, & (x,t) \in \partial B^m \times (0,T] \\ \psi^m(x,0) = \psi_0^m(x) g_m(x), & x \in B^m \end{cases}$$

It follows that (3.2.21) has a unique classical solution $\psi^m \in C^1(\overline{Q^m})$.

Remark 3.2.8

By the same arguments as in remark 3.2.7, ψ^m and $\psi^m_{x_i}$ converge weakly to ψ and ψ_{x_i} respectively, in $L_2(Q_T)$ where ψ is the weak solution of (3.2.20).

Let $\Phi(u)$ denote the generalized solution of (3.2.17) corresponding to an admissible control u , and let ψ denote the generalized solution of (3.2.20) corresponding to the optimal control u_0. We now state another lemma which will be needed in deriving the optimality conditions.

Lemma 3.2.9

Suppose the coefficients and the free term in (3.2.17) satisfy the assumptions (i) - (iii) and that there exist functions $G \in L_{q,r}(\Omega_T)$ and $H \in L_{q_1,r_1}(Q_T)$ so that $|A_i(u)|$, $|B_i(u)|$, $|C(u)| \leq G$ a.e. and $|F(u)| \leq H$ a.e. uniformly in $\mathcal{U}$. Let $\Phi_T \in L_2(B,R^N)$. Then

(3.2.22) $$\lim_{m\to\infty} \langle L(u_0)(\Phi(u_0) - \Phi(u)), \psi^m\rangle = -\int_B (\Phi(u_0) - \Phi(u))\psi dx|_{t=0}$$

for any $u \in \mathcal{U}$.

Proof:

We will follow a similar procedure as given in ([3], p.986). Let $\Phi^r(u)$, $\Phi^r(u_0)$ be the solutions of problem (3.2.19) with $m = r$ and $\Phi(u)$, $\Phi(u_0)$ the weak solutions of the problem (3.2.17) both corresponding to the controls u and u_0, respectively. The following equalities can be easily verified.

$$(1)\quad \langle L(u_0)(\Phi(u_0) - \Phi(u)),\ \psi^m\rangle = \langle L(u_0)(\Phi(u_0) - \Phi(u) - \Phi^r(u_0) + \Phi^r(u)),\ \psi^m\rangle + \langle L(u_0)(\Phi^r(u_0) - \Phi^r(u)),\ \psi^m\rangle$$

$$(2)\quad \langle L(u_0)(\Phi^r(u_0) - \Phi^r(u)),\ \psi^m\rangle = \langle \Phi^r(u_0) - \Phi^r(u),\ L^*(u_0)\psi^m\rangle + \int_B (\Phi^r(u_0) - \Phi^r(u))\psi^m dx\Big|_{t=0}^{t=T}$$

$$(3)\quad \langle \Phi^r(u_0) - \Phi^r(u),\ L^*(u_0)\psi^m\rangle = \langle \Phi^r(u_0) - \Phi^r(u), (L^*(u_0) - L^{*m}(u_0))\psi^m\rangle + \langle \Phi^r(u_0) - \Phi^r(u),\ L^{*m}(u_0)\psi^m\rangle$$

$$(4)\quad \langle \Phi^r(u_0) - \Phi^r(u),\ L^{*m}(u_0)\psi\rangle = -\int_B (\Phi^r(u_0) - \Phi^r(u))\psi^m dx\Big|_{t=T}$$

From (1) - (4) we obtain:

$$(3.2.23)\quad \langle L(u_0)(\Phi(u_0) - \Phi(u)), \psi^m\rangle = \langle L(u_0)(\Phi(u_0) - \Phi(u) - \Phi^r(u_0) + \Phi^r(u)), \psi^m\rangle + \langle \Phi^r(u_0) - \Phi^r(u), (L^*(u_0) - L^{*m}(u_0))\psi^m\rangle - \int_B (\Phi^r(u_0) - \Phi^r(u))\psi^m dx\Big|_{t=0}$$

By remark 3.2.18, $\Phi^r(u_0)$, $\Phi^r_{x_i}(u_0)$ converge weakly to $\Phi(u_0)$, $\Phi_{x_i}(u_0)$, respectively, in $L_2(\Omega_T)$. Using this fact and the boundedness properties of the coefficients and the free term (from the hypothesis of this lemma), one can show that the first term on the right hand side vanishes. By virtue of the convergence of $\Phi^r(u_0)$, $\Phi^r_{x_i}(u_0)$ mentioned above, together with the a.e. convergence of the integral averaged coefficients on Q_T and the dominated convergence theorem, the second term on the right hand side of (3.2.23) vanishes as $r, m \to \infty$. In a similar way, the first two terms on the right side of (3.2.23) vanishes as $m \to \infty$ and then $r \to \infty$, i.e., the order of taking limits is immaterial. It is also easily shown that as $r, m \to \infty$,

$$\int_B (\Phi^r(u) - \Phi^r(u_0))\psi^m \, dx \to \int_B (\Phi(u) - \Phi(u_0))\psi dt$$

and the proof is complete.

The following assumptions will be made on the Gateaux derivatives (defined in (2.1.12)) of the coefficients and free term of (3.2.17).

(iv) The strong Gateaux differentials $A_{ju}(v)$, $B_{ju}(v)$, $C_u(v)$, $F_u(v)$ exist at every point $u \in \mathcal{U}$. These differentials belong to the same function space as A_j, B_j, C, F, respectively (see assumptions i, ii, iii) for every $v \in \mathcal{U}$.

With these preparations, we now present the optimality conditions of the control problem P_2.

Theorem 3.2.10

Suppose that the coefficients A_{ij}, $A_j(u)$, $B_j(u)$, $C(u)$, $F(u)$ in the systems (3.2.17) and (3.2.20) satisfy the hypothesis of lemma 3.2.9 and that their strong Gateaux derivatives satisfy the assumption (iv). Let $\Phi_T \in L_2(B, R^N)$. Then in order that $u_0 \in \mathcal{U}$ be an optimal control, it is necessary that there exists a $\psi \in V_2(Q_T)$ so that

$$(3.2.24) \quad \int_{Q_T} \{A_{ju_0}(v-u_0)\Phi(u_0)\psi_{x_j} - B_{iu_0}(v-u_0)\Phi_{x_i}(u_0)\psi - C_{u_0}(v-u_0)\Phi(u_0)\psi - F_{u_0}(v-u_0)\psi\} dx\, dt \leq 0$$

for all $v \in \mathcal{U}$.

Proof:

Multiplying the first equation in (3.2.17) by $\eta \in \overset{0}{W}{}_2^{1,1}(Q_T)$, where $\eta|_{S_T} = 0$, and integrating by parts, we obtain

$$\int_B \Phi_T(x)\eta(x,T)dx + \int_{Q_T} (-\Phi(u)\eta_t - A_{ij}\Phi_{x_i}(u)\eta_{x_j} - A_j(u)\Phi(u)\eta_{x_j} + B_j(u)\Phi_{x_j}(u)\eta + C(u)\Phi(u)\eta + F(u)\eta)dx\, dt = \int_B \Phi(u)(x,0)\eta(x,0)dx$$

Using the above formula, one can show that

$$\langle L(u_0)\Phi(u_0) - L(u_0 + \epsilon v)\Phi(u_0 + \epsilon v), \psi^m\rangle$$
$$= \langle F(u_0 + \epsilon v) - F(u_0), \psi^m\rangle$$

Therefore,

$$(3.2.25)\quad \langle L(u_0)(\Phi(u_0) - \Phi(u_0 + \epsilon v)), \psi^m\rangle + \langle (L(u_0) - L(u_0 + \epsilon v))\Phi(u_0 + \epsilon v), \psi^m\rangle = \langle F(u_0 + \epsilon v) - F(u_0), \psi^m\rangle$$

Letting $m \to \infty$, the first term on the left hand side converges to $\int_B (\Phi(u_0 + \epsilon v) - \Phi(u_0))\psi dx|_{t=0}$ (≥ 0) by lemma 3.2.9. Therefore, taking limits as $m \to \infty$ in the last equation, we arrive at

$$\langle (L(u_0) - L(u_0 + \epsilon v))\Phi(u_0 + \epsilon v) + F(u_0) - F(u_0 + \epsilon v), \psi\rangle \leq 0$$

Dividing both sides by $\epsilon \in (0,1]$, we obtain

$$\int_{Q_T} \left\{ \left[\frac{A_j(u_0 + \epsilon v) - A_j(u_0)}{\epsilon}\right] \Phi(u_0 + \epsilon v)\psi_{x_j} - \left[\frac{B_i(u_0 + \epsilon v) - B_i(u_0)}{\epsilon}\right] \Phi_{x_i}(u_0 + \epsilon v)\psi - \left[\frac{C(u_0 + \epsilon v) - C(u_0)}{\epsilon}\right] \Phi(u_0 + \epsilon v)\psi - \left[\frac{F(u_0 + \epsilon v) - F(u_0)}{\epsilon}\right] \psi \right\} dx\, dt \leq 0$$

Note that the integrands here (except the last one) involve a matrix operation followed by a scalar multiplication in R^N. Next we take the limit in the last inequality as $\epsilon \downarrow 0$. Since the strong Gateaux differentials exist and are linear in $\mathcal{U}$ (assumption iv), it follows from the fact "$\langle f_k, g_k\rangle \to \langle f_0, g_0\rangle$ whenever $f_k \to f_0$ strongly and $g_k \to g_0$ weakly (in $L_2(Q_T)$)" that (3.2.24) holds for every $u \in \mathcal{U}$. Indeed, using the inequality 1.12 in Ladyzenskaja et al. ([37], p.137), one can verify that

$$\left[\frac{B_i'(u_0 + \epsilon v) - B_i'(u_0)}{\epsilon} \right] \psi \to B_{ju_0}'(v)\psi$$

strongly in $L_2(Q_T, R^n)$. Further, we know that $\Phi_{x_i}(u_0 + \epsilon v) \to \Phi_{x_i}(u_0)$ weakly in $L_2(Q_T, R^N)$. Therefore the third term in the above inequality converges to the corresponding term in the inequality (3.2.24) as $\epsilon \downarrow 0$. Similar conclusions hold for the other terms. This completes the proof.

Note that for a fixed u_0 the integrand in (3.2.24) is linear in $(v - u_0)$. Thus, there exists a function $G : Q_T \times U^N \times R^N \times (R^N \otimes R^N) \times R^N \times (R^N \otimes R^N) \to R^{rN}$ such that the left hand side of (3.2.24) is equal to

$$\int_{Q_T} G(x,t,u_0,\Phi(u_0), \Phi_x(u_0), \psi(u_0), \psi_x(u_0)) \cdot (v - u_0)\,dx\,dt$$

Using theorem 3.2.10, we can obtain the following pointwise necessary condition.

Corollary 3.2.11

Suppose the hypotheses of theorem 3.2.10 are satisfied, then for almost every point $(x,t) \in Q_T$ corresponding to the optimal control u_0

$$(3.2.26)\quad G(x,t,u_0(x,t),\Phi(u_0)(x,t),\Phi_x(u_0)(x,t),\psi(u_0)(x,t),\psi_x(u_0)(x,t)) \cdot (v - u_0(x,t)) \le 0$$

for all $v \in U^N$.

Proof:

Let (x^0,t^0) be a point in Q_T. Suppose A is a measurable subset of Q_T containing (x^0,t^0) and A is contracting to $\{(x^0,t^0)\}$ as $|A| \to 0$. Define a new control

$$v(x,t) = \begin{cases} v, & (x,t) \in A \\ u_0, & \text{elsewhere.} \end{cases}$$

Then the pointwise optimality condition (3.2.26) is obtained by dividing (3.2.24) by $|A|$, replacing v by the newly defined control, and applying Lebesgue's density lemma * [14].

---

* Let h be a Lebesgue integrable function defined on Q_T and let $A \subset Q_T$ be any measurable set containing $(x^0,t^0) \in Q_T$ and contracting to the one point set $\{(x^0,t^0)\}$. Then

$$\lim_{|A|\to 0} \frac{1}{|A|} \int_A h(x,t)\,dx\,dt = h(x^0,t^0) \quad \text{a.e. in } Q_T$$

As in section 2.2, the necessary condition (3.2.24) leads to a gradient method for computing the optimal control. The algorithm is presented later in this section. The pointwise necessary condition (3.2.26) gives an expression for the optimal control u_0 in terms of Φ, Φ_x, ψ, ψ_x. In the linear quadratic problem considered in the following section, (3,2,26) gives rise to ordinary differential equations.

Remark 3.2.12

Problem P_0 can be modified to the case $B = R^n$ and $J(u) = \int_0^T f_0(\xi,t,y,u)dt$ as in [3]. The optimality conditions remain the same as in (3.2.24) and (3.2.26) where now $\Phi(u_0)$ and ψ are solutions of the corresponding Cauchy problems.

The optimality condition for the original optimization problem P_1 as state in lemma 3.2.1 follows from theorem 3.2.10. This is presented below.

Theorem 3.2.13

Suppose the following assumptions for the system (3.2.7) are satisfied, i.e.,

(1) A_{ij} satisfies the assumption (i).

(2) For each $u \in \mathcal{U}$, $\bar{B}_i(u) \triangleq \partial A_{ij}/\partial x_j - B_i(u)$ and $F(u)$ are measurable functions on Q_T with values in $R^N \otimes R^N$ and R^N, respectively. Further, there exist functions $G \in L_{q,r}(Q_T)$ and $H \in L_{q_1,r_1}(Q_T)$ so that $|\bar{B}_i(u)| \leq G$ and $|F(u)| \leq H$ a.e. uniformly in $\mathcal{U}$.

(3) The Gateaux differentials $\bar{B}_{iu}(v)$ and $F_u(v)$ exist for every $u \in \mathcal{U}$ and are linear in $v \in \mathcal{U}$. These differentials belong to the same function space as $\bar{B}_i(u)$ and $F(u)$.

(4) $\Phi_T \in L_2(B)$.

Let ψ be the weak solution of the adjoint system (3.2.16) corresponding to the control u_0. Then in order that u_0 be the optimal control, it is necessary that

$$(3.2.27) \quad \int_{Q_T} [\bar{B}_{iu_0}(v - u_0)\Phi_{x_i}(u_0) - F_{u_0}(v - u_0)]\psi \, dx \, dt \leq 0$$

for all $v \in \mathcal{U}$. Further, for each regular point $(x,t) \in Q_T$ corresponding to the measurable function represented by the integrand in (3.2.27),

$$(3.2.28) \quad [\bar{B}_{iu_0}(x,t)(v - u_0(x,t))\Phi_{x_i}(u_0)(x,t) - F_{u_0}(x,t) \cdot (v - u_0(x,t))]\psi(x,t) \leq 0$$

for all $v \in U^N$.

3.2.3 APPICATION OF OPTIMALITY CONDITIONS TO LINEAR QUADRATIC GAUSSIAN PROBLEMS

In this subsection we apply the necessary condition (3.2.28) to a linear quadratic regulator problem and discuss some applications. Consider the completely controllable linear dynamical system

$$(3.2.29)\quad dx(t) = \left[A(t)x(t) + B(t)u(t) + C(t)\right]dt + D(t)dw(t),\ t \in [0,T]$$

where $\{A(t), B(t), C(t), D(t)\}$ takes the value $\{A^m(t), B^m(t), C^m(t), D^m(t)\}$ when a homogeneous Markov chain $y(t)$ takes the value m, $m = 1,2,\ldots,N$. $y(t)$ has the differential transition matrix $Q = \{q_{ij}\}$. The problem P_3 will be to choose a control $u(t,x,y)$ so as to minimize the quadratic cost criteria,

$$(3.2.30)\quad J(u) = E \int_0^T [x'Rx + u'Su + L(y)]\ dt$$

where $R \geq 0$, $S > 0$ depend on y and t, and $L(y)$ is a scalar function.

Corollary 3.2.14

The optimal control in the problem P_3 is given by

$$u^m = -S^{m^{-1}} B^{m'} \left[\tfrac{1}{2}H^m + P^m x\right]$$

and the conditional expected cost is given by

$$\phi^m = E\ \{\int_t^T [x'Rx + u'Su + L(y)]\,dx\ d\tau\,|\,x(t) = x, y(t) = m\}$$

$$= p^m + H^{m'}x + x'P^m x$$

where p^m, H^m, P^m satisfy the following scalar, vector and matrix Riccati differential equations,

$$(3.2.31)\begin{cases} \dot{p}^m = -\,\mathrm{trace}\,(D^{m'}P^mD^m) - C^{m'}H^m + \tfrac{1}{4}H^mB^mS^{m^{-1}}B^{m'}H^m \\ \qquad - q_{mi}p^i - L(m) \\ \dot{H}^m = (P^mB^mS^{m^{-1}}B^{m'} - A^{m'})H^m - 2P^{m'}C^m - q_{mi}H^i \\ \dot{P}^m = -P^mA^m - A^{m'}P^m + P^mB^mS^{m^{-1}}B^{m'}P^m - R^m - q_{mi}P^i \end{cases}$$

Proof:

By remarks 3.2.12, we can apply the optimality condition (3.2.28). Noting that $\psi(x,t) > 0$ for every $(x,t) \in R^N \times I$, we obtain

$$(3.2.32) \quad u^m = -\tfrac{1}{2}S^{m^{-1}}B^{m'}\phi_x^m, \quad m = 1,2,\ldots,N$$

Assuming a quadratic form

$$(3.2.33) \quad \phi^m(x,t) = p^m(t) + H^m(t)x + x'P^m(t)x .$$

Substituting it into (3.2.7) and grouping terms with the same power in x, we arrive at (3.2.31). The expression for the optimal control u^m is obtained from (3.2.32) and (3.2.33).

The result of a linear quadratic regulator problem considered by Wonham ([64], p.192) follows as a special case of the above corollary. A housing policy problem was considered by Kazangey and Sworder ([28], pp.1120-1128). Another problem of control of plant dynamics, repair and inventory was also considered by the same authors ([57], p.342), who used the dynamic programming approach. The differential equations they derive can be obtained as special cases from (3.2.31).

3.2.4 GRADIENT METHOD

For computation procedures, the same remark as in the case with no jump parameters (section 2.2) can be made. The pointwise optimality condition (3.2.26) can be used to obtain an expression of an extremal control in terms of Φ, Φ_x, ψ and ψ_x. The resulting computation problem is then converted into a two point boundary value problem, which is difficult to handle numerically. The integral-type optimality condition (3.2.24) leads to a gradient method. The procedures of analysis is the same as that of subsection 2.2.2, and is not repeated here. We will only state the gradient algorithm:

Algorithm:

Provide an initial guess u^1 for the optimal control and ρ^0 for the step size. In the i-th iteration, perform the following steps:

Step 1. Solve equations (3.2.17) and (3.2.20) to obtain $\Phi(u^i)$ and $\psi(u^i)$, $i = 1,2,\ldots$

Step 2. Compute $J(u^i) = \int_B \Phi(u^i)(x,0)\pi(x)dx$. If $|J(u^i) - J(u^{i-1})| < \epsilon$ for some predetermined small positive number ϵ, no more iterations will be required.

Step 3. If $J(u^{i-1}) > J(u^i)$, set $\rho^i = \rho^{i-1}$; otherwise set $\rho^i = \rho^{i-1}/C$ $(C > 1)$. Obtain u^{i-1} by

$$u^{i-1} = u^i - \rho^i G^i$$

where $G^i \triangleq G(x,t,u^i,\phi(u^i),\phi_x(u^i),\psi(u^i),\psi_x(u^i))$ and G is as defined in subsection 3.2.2.

Step 4. If u^{i-1} takes value outside the control constraint set U^N, truncate u^{i-1} as in subsection 2.2.2 so that it lies within U^N. Repeat step 1 with the new control $u^{i-1} \in \mathcal{U}$.

Note that an algorithm parallel to algorithm B of subsection 2.2.3 cannot be obtained in a similar manner because ψ is now a vector-valued function and the arguments in the proof of theorem 2.2.7 no longer holds.

The above gradient algorithm requires considerable memory and computation time. The main reason is the numerical solution of the two <u>systems</u> of parabolic equations (3.2.17) and (3.2.20). The computational aspects of this gradient method and the dynamic programming approach are briefly discussed in the following subsection.

3.3 SUGGESTIONS FOR FURTHER STUDY

In this chapter, we have derived optimality conditions for certain problems of optimal control of stochastic differential systems with jump parameters. Both the dynamic programming approach and the variational approach have been examined to a certain extent.

A detailed numerical comparison between the two approaches would be an interesting problem for further study. To discuss some possible approaches to such a comparison, we repeat below the set of optimality conditions (3.2.7), (3 2.16) and (3.2.27):

$$(3.3.1)\quad \begin{cases} \Phi_t + A_{ij}\Phi_{x_i x_j} + B_i(u)\Phi_{x_i} + C\Phi + F(u) = 0, & (x,t) \in B \times [0,T) \\ \Phi(x,t) = 0, & (x,t) \in \partial B \times (0,T) \\ \Phi(x,T) = 0, & x \in B \end{cases}$$

$$(3.3.2)\quad \begin{cases} -\psi_t + (A'_{ij}\psi_{x_j} + \overline{B}'_i(u)\psi)_{x_i} + C'\psi = 0, & (x,t) \in B \times (0,T] \\ \psi(x,t) = 0, & (x,t) \in \partial B \times (0,T) \\ \psi(x,0) = p^0 q(x), & x \in B \end{cases}$$

$$(3.3.3)\quad \int_{Q_T} \overline{B}_{iu_0}(v-u_0)\Phi_{x_i}(u_0) - F_{u_0}(v-u)\psi \, dx\, dt \leq 0$$

where $\bar{B}_i(u) = \partial A_{ij}/\partial x_i - B_i(u)$ and $\int_A q(x)dx = \pi(A)$ for $A \subset Q_T$.
At each iteration of the gradient algorithm, we solve the linear parabolic systems (3.3.1)-(3.3.2), and use the local gradient computed from (3.3.3) to modify the control.

If the dynamic programming approach is used, the optimality condition for the value function Φ can be obtained formally (as in section 3.1) as follows:

$$(3.3.4)\quad \begin{cases} \Phi_t + A_{ij}\Phi_{x_i x_j} + \min\limits_{u\in U^N} \{B_i(u)\Phi_{x_i} + C\Phi + F(u)\} = 0, \\ \qquad\qquad (x,t) \in B \times [0,T) \\ \Phi(x,t) = 0, \ (x,t) \in \partial B \times (0,T) \\ \Phi(x,T) = 0, \ x \in B \end{cases}$$

The corresponding optimal control u_0 has to satisfy

$$(3.3.5)\quad B_i(u_0)\Phi_{x_i} + C\Phi + F(u_0) = \min_U \{B_i(u)\Phi_{x_i} + C\Phi + F(u)\}$$

for a.e. $(x,t) \in Q_T$. A direct way to solve the dynamic programming equation (3.3.4) is therefore to express u_0 in terms of Φ and Φ_{x_i} and substitute u in (3.3.4) by the resulting expression for u_0; the dynamic programming equation (now a non-linear parabolic system of equations) can be solved for Φ, and the optimal control can be obtained from the earlier expression for u_0. A disadvantage of the above procedure for solving (3.3.4) is the numerical solution of the non-linear system of parabolic equations.

It may be possible to obtain the optimal control by solving a family of linear parabolic systems in the following manner. Define u_k, ϕ_k, $k = 1,2,\ldots,$ as follows: u_1 is arbitrary, and u_k and ϕ_k are given by

$$(3.3.6) \quad B_i(u_{k+1})(\phi_k)_{x_i} + C\phi_k + F(u_{k+1}) = \min_{u \in U^N} \{B_i(u)(\phi_k)_{x_i} + C\phi_k + F(u)\}$$

for a.e. $(x,t) \in Q$. For each $k = 1,2,\ldots,$ ϕ_k satisfies the linear parabolic system

$$(3.3.7) \quad \begin{cases} \phi_t + A_{ij}\phi_{x_i x_j} + B_i(u_k)\phi_{x_i} + C\phi + F(u_k) = 0, & (x,t) \in B \times [0,T) \\ \phi(x,t) = 0, & (x,t) \in \partial B \times (0,T) \\ \phi(x,T) = 0, & x \in B \end{cases}$$

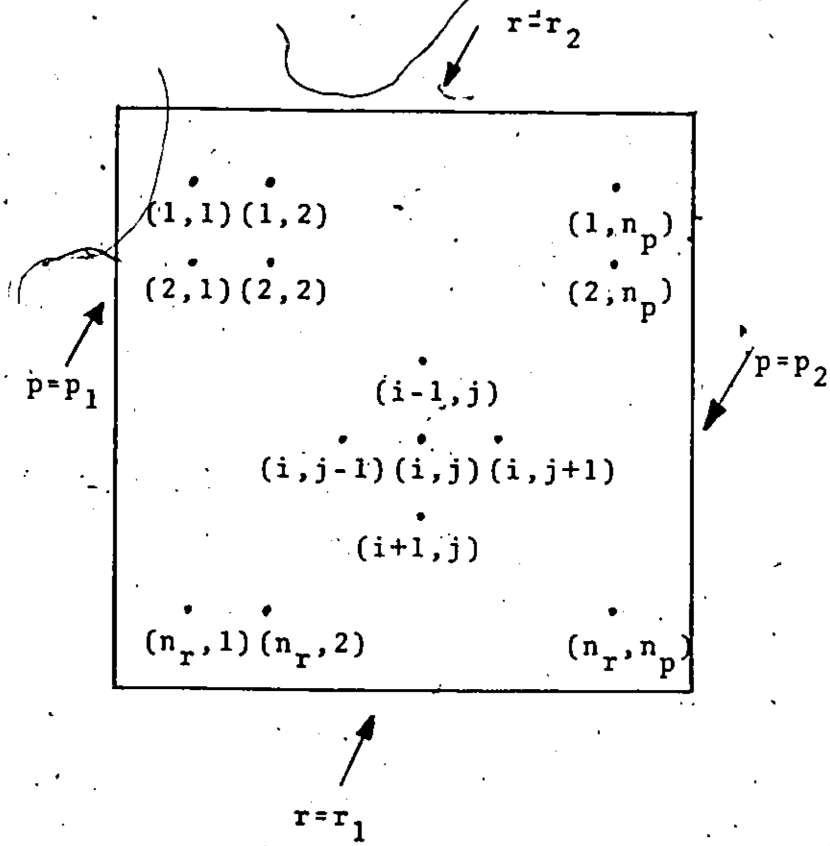
This procedure is analogous to the quasilinearization technique proposed by Bellman ([9], p.498). When the system coefficients do not depend on the Markov chain ((3.3.4) reduces to a single parabolic equation); in this special case Fleming [21] has shown that $\phi_1 \geq \phi_2 \geq \ldots$ and the limit as $k \to \infty$ is the desired function ϕ in (3.3.4). It would be interesting to establish a similar result for the case where Markov chain is involved.

If the quasilinearization method is applicable for (3.3.4), then a further problem would be to compare the effectiveness of the quasilinearization method and the gradient method. The gradient method has the advantage that successive controls are modified by the gradient of the cost function, and thus convergence should be faster. A study of the computation times required by each method in a variety of examples would be an interesting problem.

APPENDIX A

Numerical Solution of Partial Differential Equation

Numerically, the most involved part of the computation problem is step 2 in which parabolic equations are solved. The tolerable set of satellite momenta B is partitioned by $(n_p \times n_r)$ evenly spaced grid points, each of distance h from its neighboring points (shown in diagram below). The time interval $[0,T]$ is also partitioned by n_T evenly spaced grid points such that the time increment of successive points is Δ. Define ϕ_{ij}, ϕ_{ij}^{Δ} to be the value of ϕ at the (i,j)th grid point (diagram) at time t and $t + \Delta$, respectively. Similarly, ψ_{ij} and ψ_{ij}^{Δ} are defined. Denote by $u_i^{\Delta/2}$, $i = 1,2$, to be the value of u_i at time $t + \Delta/2$. Similarly the partial



derivatives $u_{ip}^{\Delta/2}$ and $u_{ir}^{\Delta/2}$ are defined. Using the Crank - Nicolson method [40] for solving parabolic partial differential equations, we obtain the following iteration schemes for problems (2.3.7) and (2.3.12):

(A1)
$$
\begin{aligned}
&\phi_{ij}^{\Delta} + (\phi_{ij+1}^{\Delta} - 2\phi_{ij}^{\Delta} + \phi_{ij-1}^{\Delta})\Delta\sigma_1^2/4h^2 \\
&+ (\phi_{i-1j}^{\Delta} - 2\phi_{ij}^{\Delta} + \phi_{i+1j}^{\Delta})\Delta\sigma_3^2/4h^2 \\
&+ (\phi_{ij+1}^{\Delta} - \phi_{ij-1}^{\Delta})\Delta u_1^{\Delta/2}/4h + (\phi_{ij+1}^{\Delta} - \phi_{ij-1}^{\Delta})\Delta ar/4h \\
&+ (\phi_{i-1j}^{\Delta} - \phi_{i+1j}^{\Delta})\Delta u_2^{\Delta/2}/4h - (\phi_{i-1j}^{\Delta} - \phi_{i+1j}^{\Delta})\Delta ap/4h \\
&+ \Delta\alpha(p^2 + r^2) + \Delta\beta\left[(u_1^{\Delta/2})^2 + (u_2^{\Delta/2})^2\right] \\
&= \phi_{ij} - (\phi_{ij+1} - 2\phi_{ij} + \phi_{ij-1})\Delta\sigma_1^2/4h^2 \\
&- (\phi_{i-1j} - 2\phi_{ij} + \phi_{i+1j})\Delta\sigma_3^2/4h^2 \\
&- (\phi_{ij+1} - \phi_{ij-1})\Delta u_1^{\Delta/2}/4h + (\phi_{ij+1} - \phi_{ij-1})\Delta ap/4h
\end{aligned}
$$

(A2)
$$
\begin{aligned}
&\psi_{ij}^{\Delta} - (\psi_{ij-1}^{\Delta} - 2\psi_{ij}^{\Delta} + \psi_{ij+1}^{\Delta})\Delta\sigma_1^2/4h^2 \\
&- (\psi_{i-1j}^{\Delta} - 2\psi_{ij}^{\Delta} + \psi_{i+1j}^{\Delta})\Delta\sigma_3^2/4h^2 \\
&+ (\psi_{ij+1}^{\Delta} - \psi_{ij-1}^{\Delta})\Delta(u_1^{\Delta/2} + ar)/4h \\
&+ (\psi_{i-1j}^{\Delta} - \psi_{i+1j}^{\Delta})\Delta(u_2^{\Delta/2} - ap)/4h + \psi_{ij}^{\Delta}(u_{1p}^{\Delta/2} + u_{2r}^{\Delta/2})\Delta/2 \\
&= \psi_{ij} + (\psi_{ij-1} - 2\psi_{ij} + \psi_{ij+1})\Delta\sigma_1^2/4h^2 \\
&+ (\psi_{i-1j} - 2\psi_{ij} + \psi_{i+1j})\Delta\sigma_3^2/4h^2 \\
&- (\psi_{ij+1} - \psi_{ij-1})\Delta(u_1^{\Delta/2} + ar)/4h \\
&- (\psi_{i-1j} - \psi_{i+1j})\Delta(u_2^{\Delta/2} - ap)/4h - \psi_{ij}(u_{1p}^{\Delta/2} + u_{2r}^{\Delta/2})\Delta/2.
\end{aligned}
$$

Note that the parabolic equation (2.3.7) has a given terminal condition and (A 1) can be used to solve for ϕ backward in time since the left-hand side of (A 1) is known at each time grid point. Similarly, (2.3.12) has a given initial condition and (A 2) can be used to solve for ψ forward in time since the right-hand side of (A 2) is known at each time grid point. Equations (A 1) and (A 2) can be put into the following matrix algebraic equation with unknown $\Phi = \{\Phi_{ij}\}$ and $\Psi = \{\Psi_{ij}\}$:

$$\text{(A 3)} \qquad \sum_{i=1}^{n_1} A_i \Phi B_i = C,$$

$$\sum_{j=1}^{n_2} D_j \Psi E_j = F$$

where A_i, B_i, C, D_j, E_j, F are known square matrices. The solution of (A 3) is obtained by standard methods.

REFERENCES

1. N.U. Ahmed, Optimal Control of Stochastic Systems, Probabilistic Analysis and Related Topics, Edited by A.T. Bharuchareid Academic Press, 2(1979) 1-68.

2. N.U. Ahmed and K.L. Teo, An Existence Theorem on Optimal Control of Partially Observable Diffusions, SIAM J. Control, 12(1974) 351-355.

3. N.U. Ahmed and K.L. Teo, Necessary Conditions for Optimality of Cauchy Problems for Parabolic Differential Systems, SIAM J. Control, 13(1975) 981-993.

4. N.U. Ahmed and K.L. Teo, Optimal Control of Ito Differential Systems with Fixed Terminal Time, Adv. Appl. Prob., 17(1975) 154-178.

5. N.U. Ahmed and K.L. Teo, Optimal Control of Stochastic Ito-Differential Systems with Fixed Terminal Time, Adv. Appl. Prob., 7(1974) 1-25.

6. N.U. Ahmed and K.L. Teo, Optimal Control of Stochastic McShane Differential Systems, J. Appl. Prob., 11(1974) 302-309.

7. N.U. Ahmed and H.W. Wong, A Minimum Principle for Systems Governed by Ito Differential Equations with Markov Jump Parameters, Lecture notes in Pure and Appl. Math., Differential Games and Control Theory, Vol. 30, Edited by E.O. Roxin, P.T. Liu and R.L. Sternberg (1977) 265-294.

8. D.G. Aronson, Non-Negative Solutions of Linear Parabolic Equations, Ann. Scuola Norm. Sup. Pisa, 22(1968) 607-694.

9. Richard Bellman and R.E. Kalaba, Quasilinearization and Nonlinear Boundary-Value Problems, American Elsevier, 1965.

10. J.M. Bismut, An Introductory Approach to Duality in Optimal Stochastic Control, SIAM Review, 20(1978) 62-78.

11. J.M. Bismut, Conjugate Convex Functions in Optimal Stochastic Control, J. Mathematical Analysis and Applications, 44(1973) 384-404.

12. J.M. But, Growth and Optimal Intertemporal Allocation of Risks, J. Economic Theory, 10(1975) 239-257.

13. M.H.A. Davis and P. Varaiya, Dynamic Programming Conditions for Partially Observable Stochastic Systems, SIAM J. Control, 11(1973) 226-261.

14. N. Dunford and J.T. Schwartz, Linear Operators Part I, Interscience Publishers, New York, (1958) 216.

15. W.H. Fleming, Controlled Diffusions under Polynomial Growth Conditions, Control Theory and Calculus of Variations, Edited by A.V. Balakrishnan, Academic Press (1969) 209-234.

16. W.H. Fleming, Duality and a Priori Estimates in Markovian Optimization Problems, J. Math. Anal. Appl. 16(1966) 254-279.

17. W.H. Fleming, Dynamical Systems with Small Stochastic Terms, Techniques of Optimization, Edited by A.V. Balakrishnan, Academic Press (1972) 325-334.

18. W.H. Fleming, Optimal Control of Diffusion Processes, Functional Analysis and Optimization, Edited by E.R. Caianello, Academic Press (1966) 65-84.

19. W.H. Fleming, Optimal Continuous Parameter Stochastic Control, SIAM Review, 11(1969) 470-509.

20. W.H. Fleming, Optimal Control of Partially Observable Diffusions, SIAM J. Control, 6(1968) 194-213.

21. W.H. Fleming, Some Markovian Optimization Problems, J. Math. Mech., 12(1963) 131-140.

22. W.H. Fleming, Stochastic Control for Small Noise Intensity, SIAM J. Control, 9(1971) 473-517.

23. W.H. Fleming and R.W. Rishel, Deterministic and Stochastic Optimal Control, New York, Heidelberg, Berlin: Springer-Verlag, (1975) 128.

24. A. Friedman, Partial Differential Equations of Parabolic Type, Prentice-Hall, Englewood Cliffs, N.J. (1974).

25. A.L. Greensite, Analysis and Design of Space Vehicle Flight Control Systems. New York: Spartan Books, (1970) 385-394.

26. E. Hewitt and K. Stromberg, Real and Abstract Analysis, Springer-Verlag, New York (1965).

27. C.J. Holland, A Numerical Technique for Small Noise Stochastic Problems, JOTA, 13(1974) 74-93.

28. T. Kazangey and D.D. Sworder, Effective Federal Policies for Regulating Residential Housing, Proc. 1971 Summer Computer Simulation Conference, (1971) 1120-1128.

29. V.B. Kolmanovsky, Some Problems of Time Optimal Control for Stochastic Systems, Problems of Control and Information Theory 4(1974) 353-367.

30. N.N. Krasvoskii and E.A. Lidskii, Analytic Design of Controllers in Stochastic Systems with Velocity Limited Controlling Action, Appl. Math. Mech., 25(1961), 627-643.

31. H.J. Kushner, Near Optimal Control in the Presence of Small Stochastic Perturbations, J. Basic Engg, 87(1965) 103-108.

32. H.J. Kushner, Necessary Conditions for Continuous Parameter Stochastic Optimization Problems, SIAM J. Control, 10 (1972) 550-565.

33. H.J. Kushner, On the Existence of Optimal Stochastic Controls, J. SIAM Control, Ser. A., 3(1966) 463-474.

34. H.J. Kushner, On Stochastic Extremum Problems, Calculus, J. Math. Anal. Appl., 10(1965) 354-367.

35. H.J. Kushner, On the Stochastic Maximium Principle; Fixed Time of Control, J. Math. Anal. Appl., 11(1965) 78-92.

36. H.J. Kushner, Probability Methods for Approximations in Stochastic Control and for Elliptic Equations, Academic Press. (1977).

37. O.A. Ladyzenskaja, V.A. Solonnikov and N.N. Uralceva, Linear and Quasi-Linear Equations of Parabolic Type, Translation of Math. Monographs, Amer. Math. Society, Providence. (1968).

38. S.J. Leese, Convergence of Gradient Methods for Optimal Control Problems, Journal of Optimization Theory and Applications, 21(1977) 329-337.

39. E.A. Lidskii, Optimal Control of Systems with Random Properties, Appl. Math. Mech., 27(1963), 33-45.

40. G.I. Marchuk, Methods of Numerical Mathematics, New York, Heidelberg, Berlin: Springer-Verlag, (1975) 142-145.

41. R.C. Merton, Optimum Consumption and Portfolio Rules in a Continuous-Time Model, J. Econ. Theory, 3(1971) 373-413.

42. E.J. McShane and R.B. Warfield Jr., On Filippov's Implicit Functions Lemma, Proc. Amer. Math. Soc., 18(1967) 41-47.

43. R. Morton, On Optimal Control of Stationary Diffusion Processes with Inaccessible Boundaries and No Discounting, J. Appl. Prob. 8(1971) 551-560.

44. R. Morton, On the Optimal Control of Stationary Diffusion Processes with Natural Boundaries and Discount Cost, J. Appl. Prob. 8(1971) 561-572.

45. S.R. Pliska, A Diffusion Process Model for the Optimal Operation of a Reservoir System, J. Appl. Prob. 12(1975) 859-863.

46. D.W. Reid, On the Computational Methods for Optimal Control Problems of Distributed Parameter Systems, Ph.D. Thesis, School [illegible] Mathematics, University of New South Wales, Australia, March 1980.

47. D.W. Reid and K.L. Teo, First-Order Strong Variation Algorithm for A Class of Distributed Optimal Control Problems, to appear in Applied Mathematics and Optimization.

48. D.W. Reid and K.L. Teo, On a Computational Method for Stochastic Optimal Parameter Selection Problem, to appear.

49. R. Rishel, Dynamic Programming and Minimum Principles for Systems with Jump Markov Disturbances, SIAM J. Control, 13(1975) 338-371.

50. R.W. Rishel, Weak Solutions of a Partial Differential Equation of Dynamic Programming, SIAM J. Control, 9(1971) 519-528.

51. P. Robinson and J. Moore, Solution of the Stochastic Control Problem in Unbounded Domain, J. Franklin Institute, 295 (1973) 185-192.

52. D.W. Stroock and S.R.S. Varadhan, Diffusion Processes with Continuous Coefficients I, Comm. Pure Appl. Math., 22(1969) 345-400.

53. D.D. Sworder, Feedback Control of a Class of Linear Stochastic Systems, Proc. 1968, Joint Automatic Control Conf., (1968) 34-44.

54. D.D. Sworder, Feedback Control of a Class of Linear Systems with Jump Parameters, IEEE Trans. Automatic Control, 14(1969) 9-14.

55. D.D. Sworder, On the Control of Stochastic Systems, Int. J. Control, 6(1967) 179-188.

56. D.D. Sworder, On the Stochastic Maximium Principle, J. Math. Anal. Appl., 24(1968) 627-640.

57. D.D. Sworder and T. Kazangey, Optimal Control, Repair and Inventory Strategies for a Linear Stochastic System, IEEE Trans. Systems, Man, and Cybernetics, 2(1972) 342-347.

58. K.L. Teo, Optimal Control of Systems Governed by Ito Stochastic Differential Equations, Ph.D. Thesis, Department of Electrical Engineering, University of Ottawa, Canada, (1974).

59. K.L. Teo, D.W. Reid and I.E. Boyd, Stochastic Optimal Control Theory and Its Computational Methods, International Journal of System Sciences, 11(1980) 77-95.

60. K.L. Teo and D.W. Reid, Optimal Parameter Selection of Parabolic Systems, to appear in Mathematics of Operations Research.

61. H.W. Wong, Suboptimal Control of Satellite Momenta in a Noisy Environment, IEEE Trans. Aerospace and Electronic System, 14(1978) 368-371.

62. H.W. Wong and N.U. Ahmed, Gradient Method for Computing Optimal Control for Stochastic Ito Differential Equations, to appear.

63. H.W. Wong and N.U. Ahmed, Optimal Precession Control of Randomly Perturbed Spin-Stabilized Satellite, IEEE Trans. Aerospace and Electronic Systems, 13(1977) 172-178.

64. W.M. Wonham, On the Separation Theorem of Stochastic Control, SIAM J. Control, 6(1968) 321-326.

65. W.M. Wonham, Optimal Stochastic Control, Automatica, 5 (1969) 113-118.

66. W.M. Wonham, Random Differential Equations in Control Theory, Probabilistic Methods in Applied Mathematics, Vol. 2, edited by A.T. Bharucha-Reid, Academic Press (1970) 131-212.

67. Y. Yavin and M. Friedman, Computation of Optimal Controls of Stochastic Van der Pol Type Oscillation, J. Franklin Institute, 229(1975) 547-561.

VITA

NAME: Hon-Wing WONG

DATE OF BIRTH: March 29, 1949

EDUCATION:

I. HIGH SCHOOL: Ying Wa College, Hong Kong.

II. UNIVERSITIES: Sir George Williams University, Montreal, Canada; University of Ottawa, Ottawa, Canada.

III. Degrees: B.A.Sc. (1973) and M.A.Sc. (1974); both in Electrical Engineering and from University of Ottawa.