TIME OPTIMAL CONTROL IN STOCHASTIC SYSTEMS

A Dissertation

Presented to

the Faculty of the Graduate School

University of Missouri

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

by

Won-Kyoo Lee

May 1972

Richard H. Luecke   Dissertation Supervisor

ACKNOWLEDGEMENTS

The author wishes to express his sincere gratitude to his advisor, Dr. R. H. Luecke, for his valuable guidance, encouragement and constructive criticism throughout this work.

The author wishes to thank the Department of Chemical Engineering for making teaching assistantship available to him. The financial support of the National Science Foundation under Grant No. NSF GK 13617 is gratefully acknowledged.

Special thanks are due my wife, Mih-Song, for her patience, understanding, encouragement and assistance without which this work might never have been completed, and my son, Albert, for many lost hours.

ABSTRACT

Time optimal control of stochastic dynamic system is considered. It is assumed that noise-free observations are available at all times. An optimal admissible feedback control policy is formulated leading to minimization of the expectation of the length of time required to reach the desired terminal region.

Dynamic programming formalism leads to a second order nonlinear partial differential equation. The difference between the stochastic and deterministic equations is represented by a truncated power series and the optimal switching surface for a "bang-bang" controller is then computed through a direct search using repetitive simulations.

Numerical results for the location of the stochastic switching curves for a specific second order system are computed and discussed.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER I

INTRODUCTION

A mathematical description of the physical processes is required for application of modern control theory methods. Uncertainty or random properties can enter this model in several ways. For example, some process parameters may vary randomly with time. There also may be random fluctuations in the input to the process and the control inputs are themselves random process if the controlling scheme is based on a random policy. Furthermore, observations or measurements may be corrupted with some inherent error which also has random properties.

A system involving such random variables is a stochastic system. A stochastic system which evolves according to a rule which also involves variables or parameters under external control is called a stochastic control system. If these variables or parameters are determined so that the system behaves as well as possible as measured by some well-defined criterion, one has achieved optimal control of the stochastic system.

The conventional approach to the computation of optimal control for stochastic systems has been to replace the random variables by their expected values and computing the control policy using deterministic control methods. To allow for uncertainty, gross safety factors have been added to the deterministic results.

While the introduction of stochastic features into a description of the process cannot reduce the uncertainty, it can lead to the more precise statement about uncertainty and to the better decision making.

The control of dynamic process in the presence of random noise and disturbances is the subject of stochastic control theory, a branch of modern control theory which has been growing rapidly during the past years. Powerful tools for the generation of the optimal control law such as Pontryagin's Maximum Principle or the Hamilton-Jacobi equation of the calculus of variations do not apply in a straightforward manner to stochastic optimal control problems.

Optimal control theory for stochastic systems has evolved since 1960 largely within the framework of dynamic programming. The mathematical theory is not in completely satisfactory form. Results of practical interest have been obtained only in a few simple academic problems such as a linear regulator problem. Nevertheless considerable progress has been made in understanding the nature of stochastic optimal control problems.

The optimal control problems for stochastic systems are generally concerned with the minimization of an expected cost functional. Only in the linear quadratic case is the stochastic optimal control problem completely solved.

In many situations, such as control of batch chemical processes, it is time rather than control energy which is expensive. That is, in a time optimal control problem in stochastic systems, the criterion of performances is taken to be the expected value of time required to move the system from the given initial state to the desired terminal region using a constrained control variable. In this work it is desired to determine an admissible feedback control law which minimizes this performance criterion. It is assumed throughout that the state of system can be measured noise-free at all times, and that the amplitude of control inputs is constrained.

The dynamic programming approach leads to the problem of finding the solution to a nonlinear partial differential equation called the stochastic Hamilton-Jacobi equation. This equation differs from the corresponding equation for the deterministic case by the appearance of terms involving the second partial derivatives of the optimal expected value function. Analytic solutions for this equation are quite difficult to obtain. A number of procedures have been proposed for solving numerically the nonlinear partial differential equations.

The objective of the present work is to generate the optimal switching surface for the resulting stochastic bang-bang controller. For this purpose the difference between the stochastic and deterministic equations for switching surface is formulated, and then computed using repetitive simulations. This procedure is applied to a specific second order system for locating the optimal switching curve. It is used in conjunction with a stochastic switching strategy of multiple switchings, which is necessary to guarantee the system to be transferred to the desired terminal region. Numerical results are presented and discussed.

CHAPTER II

DESCRIPTION OF THE PROBLEM

Optimization is considered of a process whose time behavior can be described by a set of ordinary differential equations. The process is subject to additive random disturbances so that the process can be represented by a set of stochastic differential equations.

The basic problem is to find the optimal control policy $u^*(t)$ such that the system state can be moved from some given initial state, $x(t_o)$, to a desired terminal region in the minimum length of time.

In order to retain as much simplicity as possible, it is assumed throughout that noise-free observations of state vector, $x(t)$, are available at all times. The results from this study can be extended to observable systems with noise in the observations but this aspect of the problem will not be considered here.

In many deterministic problems, the desired terminal region consists of a single point. However, the input perturbation prevents the stochastic system from remaining arbitrarily close to any given point, and the length of time to reach a single point may not be finite. In other words, in the stochastic case a terminal region composed of a single point may not lead to a well-posed problem (16, 17). Thus, the terminal region for this problem is taken to be a compact convex set called the target set S.

A problem exists in maximizing the probability or the expectation of the length of time that the state will remain in the target set S once that set has been reached. That problem has been discussed by others (15, 30), and will not be addressed in this study. Our objective is to determine the path that the state should take outside the target set.

CHAPTER III

MATHEMATICAL FORMULATION OF STOCHASTIC OPTIMAL CONTROL

Stochastic Dynamical System

The system is subject to additive random disturbance which is assumed to be Brownian motion and can be described by the vector stochastic differential equation,

$$dx(t,\omega) = \{f(x(t,\omega),t) + B(t)u(x(t,\omega),t)\}dt + G(t)d\eta(t,\omega), \quad (3\text{-}1)$$

or

$$\dot{x}(t,\omega) = f(x(t,\omega),t) + B(t)u(x(t,\omega),t) + G(t)w(t,\omega), \quad (3\text{-}2)$$

$$x(t_o) = x_o = \text{fixed constant},$$

for all $t \geq t_o$,

where ω is an element of a probability space, Ω;

$x(t,\omega)$ is an n-vector state;

$u(x(t,\omega),t)$ is an m-vector control with values in a nonempty compact convex restraint set $U \subset R^m$;

$f(x(t,\omega),t)$ is an n-vector valued function;

$B(t)$ is an $n \times m$ matrix, $m \leq n$;

$G(t)$ is an $n \times p$ matrix, $p \leq n$;

$\eta(t,\omega)$ is a p-vector Wiener process (Brownian motion);

$w(t,\omega)$ is not defined mathematically but it is called a gaussian white noise in the engineering literature;

and

$$\eta(t + \Delta t) - \eta(t) = \int_t^{t+\Delta t} w(\tau)d\tau \equiv d\eta(t)$$

$$E\{\eta(t+\Delta t)-\eta(t)\} \equiv E\{d\eta(t)\} = 0^{*}$$
$$E\{d\eta(t)d\eta'(t)\} = Q(t)dt \qquad (3\text{-}3)$$
$$E\{w(t)\} = 0, \text{ for all } t \geq t_o$$
$$E\{w(t)w'(\tau)\} = Q(t)\delta(t-\tau), \text{ for } t,\ \tau \geq t_o$$

where E{ } indicates the mathematical expectation, and $Q(t)$ is a $p \times p$ non-negative definite matrix.

For existence, uniqueness and continuity of solution to equation (3-1) or (3-2), it is assumed that $f(x,t)$ and $u(x,t)$ satisfy a uniform Lipschitz condition in x, and that $B(t)$ and $G(t)$ are bounded and measurable. Since $d\eta(t)$ in equation (3-1) is not multiplied by a function of the state $x(t)$, no stochastic integrals are involved; that is, the integral $\int G(t)d\eta(t)$ can be interpreted as an ordinary Rieman-Stieltjes integral (10).

Formal Definition of Optimal Control Policy

Since the process is subject to additive gaussian white noise, the length of time to reach the target set is a random variable. This is why the objective must be formulated to give the control policy that minimizes the expectation of the length of time required to move the system from $x(t_o)$ to the target set. Thus the performance criterion becomes

$$J(x_o,u) = E\{\int_{t_o}^{t_s} dt \mid x(t_o) = x_o\} \qquad (3\text{-}4)$$

---

*From now on the dependence of processes on ω will be suppressed for notational simplicity.

where t_s is the time required to first reach the desired target set S, given the initial state, $x(t_o)$.

The control u(x,t) is constrained to a compact convex set U and is assumed to be piecewise continuous. The problem then is to determine an admissible feedback control law u(x,t) which minimizes J, given the constraint of equations (3-1) or (3-2).

Define V(x) to be the minimum of the performance criterion (the minimum expected time) from a given initial state x(t) at time t over a set of admissible controls, U:

$$V(x) \equiv \min_{u \varepsilon U} J(x,u) \tag{3-5}$$

Formal application of dynamic programming to this problem (3, 5, 6, 8, 16) yields a second order nonlinear partial differential equation that is satisfied by V(x):

$$\min_{u \varepsilon U} \{\tfrac{1}{2}\, \mathrm{tr}((\partial^2 V/\partial x^2)\ GQG') + (\partial V/\partial x)'(f + Bu) + 1\} = 0 \tag{3-6}$$

where $\partial/\partial x = (\partial/\partial x_1, \ldots\ldots, \partial/\partial x_n)'$;

$$\mathrm{tr}((\partial^2 V/\partial x^2)\ GQG') = \sum_{i,j} g_{ij}(\partial^2 V/\partial x_i \partial x_j);$$

$$(g_{ij}) = GQG'.$$

Equation (3-6) is to be solved outside the target set S with the boundary condition,

$$V(x) = 0 \text{ for } x \ \varepsilon\ \partial S, \text{ boundary of S.} \tag{3-7}$$

Equation (3-6) for stochastic optimization can be heuristically obtained by a formal Taylor's series expansion with added complication of including second order terms. Expansion is carried to the second power in Δx because $\eta(t)$ is a Brownian motion. It is generally required that $V(x)$ exists, and that $V(x)$ is twice continuously differentiable in x and continuously differentiable in t.

In many applications equation (3-6) is degenerate in that matrix (g_{ij}) is not of full rank. For any stochastic control system governed by a single nth order differential equation with gaussian white noise input, the corresponding partial differential equation is also degenerate. In that case the analytic theory yields few results on existence, uniqueness, or smoothness of solutions to equation (3-6), either numerical or analytical. That problem has been discussed by others (10, 25). They showed that the degenerate partial differential equation had a solution under mild assumptions. It is also shown that an optimal control law may be determined by minimizing equation (3-6) in terms of the partial derivatives of that solution.

Kushner (18) obtained equation (3-6) as a sufficient condition for the optimality of a stochastic control system. However, recently Rishel (25) showed that dynamic programming conditions gave both necessary and sufficient conditions for optimality in stochastic optimal control systems.

With the control vector u a member of the compact convex set $U = \{u: |u_i| \leq 1, i = 1, \ldots, m\}$, it can be shown that it follows from equation (3-6) that the optimal control u^* is given by

$$u^* = - \operatorname{sgn}\{B'(\partial V/\partial x)\} \tag{3-8}$$

where sgn{y} means the sign of y. Thus, the optimal control policy is of the "bang-bang" type with the switching surface being given by

$$B'(\partial V/\partial x) = 0 \tag{3-9}$$

Substituting equation (3-8) into equation (3-6), it follows that

$$\tfrac{1}{2}\, tr((\partial^2 V/\partial x^2)\ GQG') + (\partial V/\partial x)'f - |(\partial V/\partial x)'B| + 1 = 0 \tag{3-10}$$

Equation (3-10) holds outside the target set S with the boundary condition given in equation (3-7).

For existence of an optimal stochastic control one condition required by Kushner (17) is that the control laws satisfy local Lipschitz condition in x and t. That is, there is a K > 0 such that

$$||u(x + \Delta x,\ t + \Delta t) - u(x,t)|| \leq K\ (||\Delta x|| + |\Delta t|) \tag{3-11}$$

However, Benes (4) has proved that the admissible control laws do not have to be Lipschitz.

The formulation of equation (3-6) differs from that for the controller in deterministic system only by the first term of second order partial derivatives. Thus equation (3-6) approaches the corresponding deterministic equation, as it should, as the variance of the input noise Q approaches zero. The bang-bang optimal control policy for the deterministic system has a form identical to equation (3-8) where only the partial differential equation constraining the function V(x) is different. For the deterministic case, solutions, and sometimes even analytic solutions, of the partial differential equation can be found by indirect means; for the stochastic case with the second order derivatives, only numerical methods, however, are even conceivable.

CHAPTER IV

REVIEW OF PREVIOUS WORK

The use of dynamic programming formalism reduces the stochastic time optimal control problem to solving a second order nonlinear partial differential equation (stochastic Hamilton-Jacobi equation). The nonlinear partial differential equation can rarely be solved analytically.

Various methods for numerical approximation to the solution of equation (3-6) have been proposed. Knowledge of all boundary conditions are required for numerical solution of the partial differential equation (3-6). On the boundary of the target set S, $V(x) = 0$ since the minimum time to reach the set S from its boundary is zero. However, the values of $V(x)$ for large values of $||x||$ are not known. It has been cited (7, 22) that $V(x)$ approaches, at most, an algebraic function of the components of x for large values of $||x||$. However, formal justification for this conclusion was not given.

Most of the existing work on time optimal control in stochastic systems has been focused on solution to equation (3-6). Time optimal control problem in stochastic system was studied by Aoki (2) and Novosel'tsev (24). In those studies the continuous-time system was approximated by a discrete-time system, and a different switching curve for the same system was obtained by each author. This difference has not been resolved.

Several papers have been published on methods applicable to the solution to equation (3-6) if the noise level is low. In a perturbation technique (successive approximation method) the objective of this method is to solve approximately the stochastic optimal control problem using the solution of the corresponding deterministic problem. The solution for equation (3-6) is obtained as a convergent power series in Q, the variance of noise:

$$V_s(x) = V_D(x) + QV_1(x) + Q^2V_2(x) + \ldots..$$
$$(\partial V_s(x)/\partial x) = (\partial V_D(x)/\partial x) + Q(\partial V_1(x)/\partial x) + Q^2(\partial V_2(x)/\partial x) + \ldots.. \tag{4-1}$$

where $V_S(x)$ is a solution for the stochastic system and $V_D(x)$ the corresponding deterministic solution.

The coefficients $V_i(x)$ satisfy the equation found by differentiating equation (3-6) repeatedly with respect to Q and then setting Q = 0. It is required that $V_s(x)$ be continuous across a switching surface. Since the series solution, equation (4-1), is associated with a singular perturbation which results in a change of order of the perturbed differential equation, it is known (1) that such a series does not, in general, represent a solution which is uniformly valid for all x.

For linear time-optimal control problems one may seek approximations to the switching surface for the stochastic optimal control as perturbations of the corresponding surface for the deterministic optimal control. Formal studies of such perturbation problems were made in Dorato et al. (7) and Stratonovich (29). It is known (11) that a rigorous treatment of this problem is hampered by the fact that both $(\partial V_1(x)/\partial x)$ and $(\partial^2 V_D(x)/\partial x^2)$ are discontinuous across a switching surface for the deterministic optimal control.

Fleming (11) studied the mathematical validity of approximate solution by a singular perturbation. He showed that approximate solutions hold in regions where the solution of the Hamilton-Jacobi equation for the deterministic problem is sufficiently well-behaved.

Dorato *et al*. (7) reported a method for an approximate solution to time-optimal control problems in stochastic system with low noise levels using a singular perturbation technique. The method was applied to a second order purely inertial system with additive gaussian white noise with variance of 0.3 at the control input. Two different terminal regions, the closed interval and the circular, were considered. With the terminal region of the closed interval, identical switching curves for stochastic and deterministic cases were obtained, but with the circular the stochastic switching curve somewhat different and lay slightly clockwise from the deterministic one.

Kushner and Kleinman (19) considered direct numerical solution of the second order partial differential equation, equation (3-10), of elliptic type in a bounded region. An artificial outer boundary was imposed for the numerical analysis. The nonlinear partial differential equation was then approximated using finite difference equations. The discretized problem corresponds to an optimal control problem for Markov chains whose states are the points of the grid.

An iterative method was derived for finding the optimum, which is basically similar to the usual Jacobi and Gauss-Siedel method for linear problems. It converged at least as good as the schemes of Howard (12) and Eaton-Zadeh (9). Later (20, 21) procedures were defined for accelerating the convergence of iterative methods. It is reported that one of the methods gave a ten-fold decrease in computation time over a more usual procedure of dynamic programming. Furthermore, as the artificial outer boundary increases, convergence of the corresponding sequence of V(x) to the solution of original problem can often be proved. The same system as Dorato *et al*. (7) used was considered with the rectangular terminal region. The numerical results indicated that the optimal switching curve for the stochastic system would lie somewhat counterclockwise from the deterministic curve, which is different from the results of Dorato *et al*. (7) in that slightly clockwise location was obtained in a different terminal region configuration.

Robinson and Yurtseven (27) solved the partial differential equation using a Monte Carlo technique on an analog computer. The optimal switching curve was found using an iterative technique called approximation-in-policy-space. For time-optimal control in the second order inertial system, there was close agreement between the switching curves obtained by perturbation technique by Dorato *et al*. (7) and this procedure.

CHAPTER V

DEVELOPMENT OF
DIRECT SOLUTION FOR SWITCHING POLICY

Direct Solution for Switching Surface

A direct approach is developed in this work to compute the switching surface for time optimal control in stochastic systems. The partial differential equation (3-6) differs from the corresponding equation for a deterministic system having the same plant characteristics only by the first term of second order partial derivatives. Thus the switching surface (equation (3-9)) for the stochastic system differs from the relatively easily found solution for a deterministic but otherwise similar system only by an analytic function.

The analytic function of the difference in switching functions is expanded into a Taylor's series taking as the variable the distance from the origin in state space along the switching surface. In other words, the difference between equation (3-9) for the stochastic switching surface and the corresponding equation for the deterministic case is formulated as a power series in s, the distance from the origin along the switching surface. Thus

$$B'(\partial V_S/\partial x) = B'(\partial V_D/\partial x) + a_o + a_1 s + a_2 s^2 + \ldots.. \qquad (5\text{-}1)$$

where s is the distance from the origin, positive to the right, negative to the left;

V_S is the minimum expected time for the stochastic system;

V_D is the minimum time for the deterministic system;

a_i are constants.

The coefficients a_i for the truncated series in equation (5-1) may then be determined by experimental or Monte Carlo parameter search methods. The objective function to be minimized is the expected value of the time to reach the target set.

Characteristics of the Stochastic Switching Surface

For a deterministic linear time-invariant system, it can be shown that, at most, (n-1) sign switches are required on the time optimal path if the system is normal and if all eigenvalues of the matrix A are real (where the matrix A is a constant $n \times n$ matrix in the linear time -invariant system, $\dot{x} = Ax + Bu$). No such theorem can exist for the linear, time-invariant stochastic system. The reason may be seen by reference to Figure 1. After reaching the switching surface, the state would tend to be driven off the switching surface by the stochastic input. If another switching were not used to return the system to the prescribed trajectory, this behavior could result not only in a longer settling time but also it could cause the state trajectory to miss the target set altogether.

In this work only the linear coefficients have been determined with the higher order difference coefficients assumed to be small. The numerical results of tests on a simple example system to be discussed reflect the magnitude of deviations from that assumption.

The stochastic switching surface must more nearly approximate the deterministic surface as the target set is approached. In this work it is arbitrarily assumed that two curves intersect at the border of the

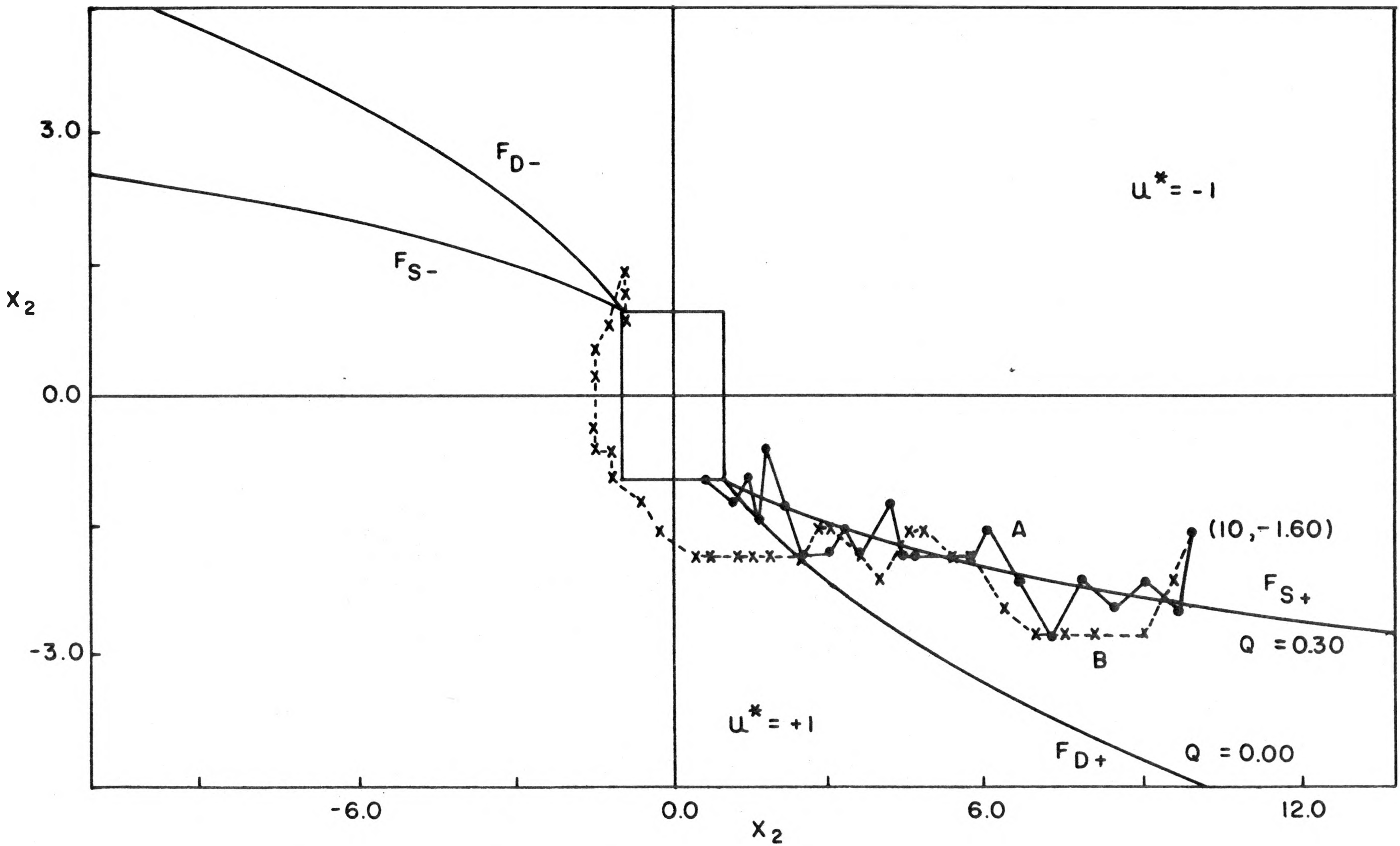

Figure 1. Stochastic Time-Optimal Trajectories

target set S. This assumption further simplifies the task of determining the coefficients. This assumption along with the linearity assumption means that only a_1 in equation (5-1) need be found.

CHAPTER VI

NUMERICAL RESULTS AND DISCUSSIONS

A Numerical Example

The plant under consideration for a numerical example is a second order linear time-invariant system with constrained control $|u| \leq 1$. Scalar gaussian white noise is added to the system with the control variable:

$$\begin{aligned} dx_1 &= x_2 dt \\ dx_2 &= u dt + d\eta \end{aligned} \tag{6-1}$$

or

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= u + w \end{aligned} \tag{6-2}$$

where $x(t_o) = x_o$ is a fixed constant as the initial state, and w is a scalar gaussian white noise with $E\{w(t)\} = 0$ and $E\{w(t)w'(\tau)\} = Q(t)\delta(t-\tau)$ = constant for $t, \tau \geq t_o$.

The target set was taken as the rectangle

$$S = \{x: |x_i| \leq 1,\ i = 1,\ 2\} \tag{6-3}$$

For this system equation (3-6) takes the following form:

$$\min_{|u| \leq 1} \{\tfrac{1}{2} Q(\partial^2 V_s/\partial x_2^2) + x_2(\partial V_s/\partial x_1) + u(\partial V_s/\partial x_2) + 1\} = 0 \tag{6-4}$$

so that the optimal control becomes

$$u^* = - \operatorname{sgn}(\partial V_S/\partial x_2) \tag{6-5}$$

The first step for solution to this stochastic problem is to solve the associated deterministic problem. These are the same as equations (6-4) and (6-5) except that the second order derivative is missing:

$$\begin{aligned} &\min_{|u|\leq 1} \{x_2(\partial V_D/\partial x_1) + u(\partial V_D/\partial x_2) + 1\} = 0 \\ &u^* = - \operatorname{sgn}(\partial V_D/\partial x_2) \end{aligned} \tag{6-6}$$

The solution to these equations is most easily found in an indirect method by use of the Minimum Principle. Direct solution for system trajectories is then possible using the bang-bang control u* in equation (6-6). The details of this solution appear in the Appendix A. The optimal switching curve is given by (Figure 2):

$$F_D = - x_1 + \tfrac{1}{2} u^* x_2{}^2 + \tfrac{1}{2} u^* = 0$$

or (6-7)

$$F_D = - x_1 \pm \tfrac{1}{2} x_2{}^2 \pm \tfrac{1}{2} = 0$$

where F_D defines the deterministic switching curve.

For the stochastic problem, a solution is assumed to be of the form of equation (5-1) but that equation is modified in two ways. First, a linear approximation is taken neglecting the higher order terms. Secondly, the coordinate x_1 is used as the variable in the Taylor's series expansion which, in this example, is a monotone function of the

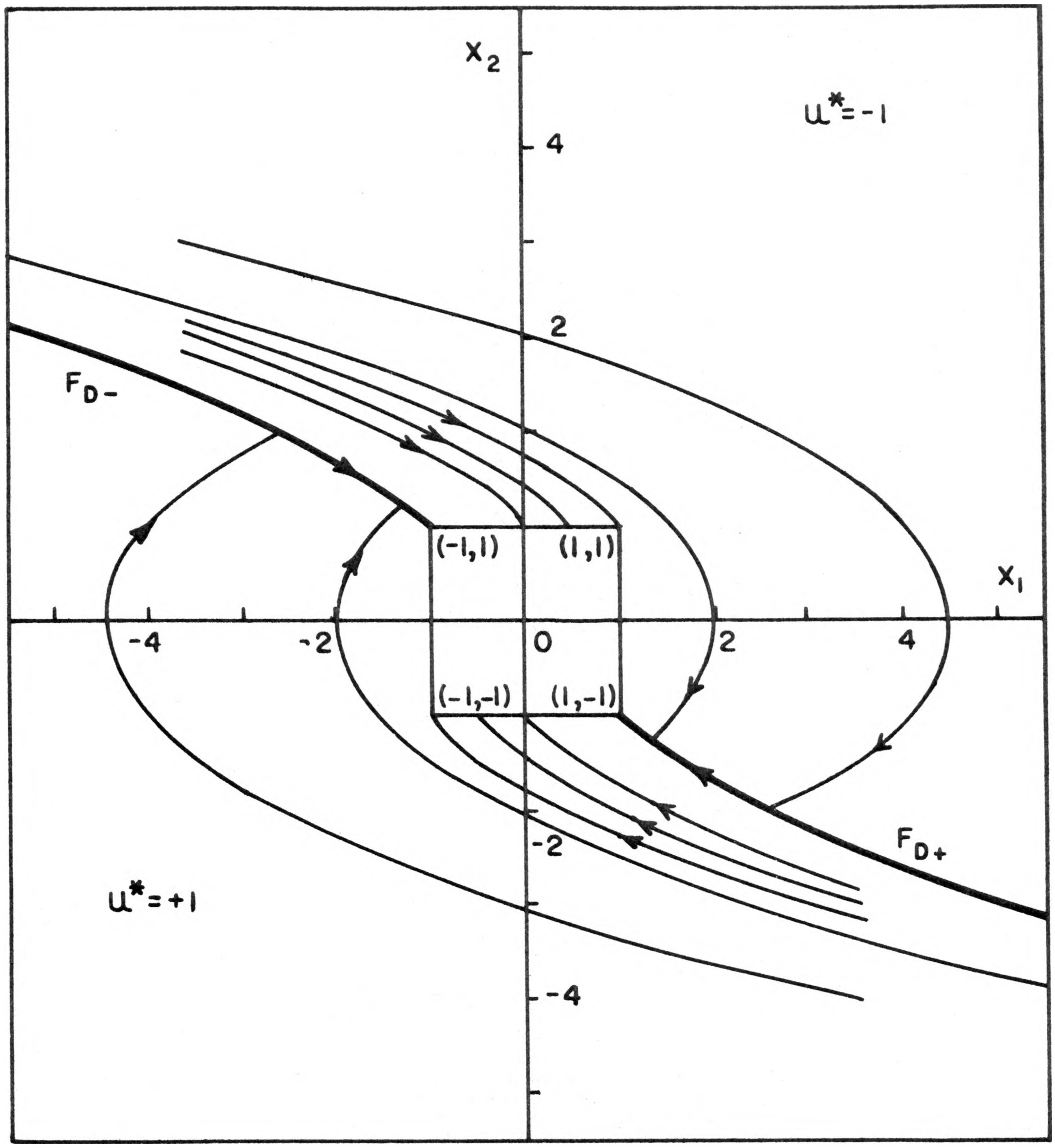

Figure 2. Deterministic Time-Optimal Trajectories

distance s along the switching curve. Thus equation (5-1) becomes

$$F_S = F_D + a_o + a_1 x_1 = 0 \tag{6-8}$$

where F_S defines the stochastic switching curve.

The two curves were assumed to intersect on the border of the target set S, at points (1, -1) and (-1, 1). Thus for the case of intersecting at (1, -1), the values 1, 1 and -1 are substituted into equations (6-7) and (6-8) for u*, x_1 and x_2 respectively, giving a_o equal $-a_1$. For the intersection at (-1, 1), substituting -1, -1 and 1 for u*, x_1 and x_2 in equations (6-7) and (6-8), $a_o = a_1$ is obtained. Hence, the switching curve can be expressed as follows:

$$F_S = F_D - u^* a_1 + a_1 x_1 = 0 \tag{6-9}$$

Substituting equation (6-7) into equation (6-9) yields

$$F_S = x_1(1 - a_1) - \tfrac{1}{2} u^* x_2^2 - u^* (\tfrac{1}{2} - a_1) = 0 \tag{6-10}$$

The value of parameter a_1 in equation (6-10) is to be varied to find an optimum value for a_1 which gives the minimum value of the expected time to reach the target set S from the given initial state. With the optimum value of a_1, the switching curve given by equation (6-10) is to be used for the stochastic bang-bang controller.

For a fixed value of a_1, the following procedure was used to estimate the expectation of the time to reach the given target set. The dynamic system equation (6-1) or (6-2) was integrated along with the additive gaussian white noise input to obtain a stochastic trajectory as shown in Appendix B. The time required to first reach the target set S from a given initial state was measured. This procedure was then repeated twenty times using different random noise sequences as the input. The average of the twenty runs was taken as an estimate of the expected value.

A special switching strategy was necessary to guarantee that the stochastic system would converge to the target set. As shown in Figure 1, the first switching was made when optimal trajectory first fell below the switching curve F_{s+}, a portion of the stochastic switching curve with $u^* = +1$. Thereafter whenever the trajectory crossed above the curve F_{s+}, the control was changed to drive it immediately below the switching curve. This switching policy allowed trajectories to remain below the switching curve F_{s+} and above F_{s-}, a portion of the stochastic switching curve with $u^* = -1$. Although there is no guarantee that the stochastic time -optimal trajectory follows exactly the switching curve after switching occurs, this policy guaranteed that the process would be transferred to the desired target set from a given initial state. In the worst case, additional clockwise cycles around the target set S could be required (Figure 1). This type of switching policy should find usefulness in practice in real applications.

For the example considered here two levels of variance of input noise were considered -- 0.05 and 0.30 for a gaussian white noise. A gaussian white noise was approximated by a piecewise constant gaussian random sequence.

For a given magnitude the variance of an approximated gaussian white noise depends strongly on the size of time increment, Δt. That is, the variance of a piecewise gaussian random sequence is inversely proportional to the time increment. In this numerical example $\Delta t = 0.25$ of time increment was used, which implies that variances used for generation of a gaussian random sequence were 0.2 and 1.2. From now on, the term "variance" of input noise will refer to the variance of a gaussian white noise.

Results and Discussions

Two typical trajectories of the stochastic system are shown in Figure 1. The switching curve shown for the noise level of zero is the optimum for the deterministic system.

The value for a_1 that minimizes the expected time to reach the target set S from a given initial state could be found by a one-dimensional search method such as stochastic approximation. In this work however, a case study type "search" was made by computing the estimated trajectory time over a wide range of values for a_1. Presented in Tables 1 through 6 are the estimate of the expected length of time required to reach the target set for various values of a_1. These data are plotted in Figures 3 through 8 where Figures 3, 4 and 5 are for the noise level of $Q = 0.30$ and Figures 6, 7 and 8 the noise variance of 0.05. These plots indicate not only the probable locations of the minima but also show the sensitivity of the minima to changes in the parameter a_1.

Three different initial states were explored -- $x_1 = 5$, $x_1 = 10$ and $x_1 = 15$. Several considerations entered into selection of the initial x_2 coordinate. Since system trajectories flow generally clockwise (Figure 2), the switching curve would be reached moving in a clockwise direction. That is, the initial point should be above the

TABLE 1

THE MEAN VALUE OF THE LENGTH OF TIME TO REACH THE TARGET SET FOR VARIOUS VALUES OF PARAMETER a_1 AT NOISE LEVEL OF Q = 0.30

Initial State $(x_1, x_2) = (5.00, -1.34)$

The Minimum Length of Time for the Deterministic Problem = 2.306

Parameter a_1	Mean Value of the Length of Time	Standard Deviation of Mean	95.0% Confidence Interval for Mean
-0.40	4.32969	0.78270	5.86378 - 2.79560
-0.20	4.72812	0.78710	6.27083 - 3.18541
0.00	3.96328	0.54853	5.03841 - 2.88815
0.10	3.66094	0.50680	4.65427 - 2.66761
0.20	3.68125	0.51986	4.70017 - 2.66232
0.25	3.69375	0.51830	4.70961 - 2.67789
0.30	3.43672	0.45558	4.32965 - 2.54379
0.35	3.47422	0.45102	4.35822 - 2.59022
0.40	3.53984	0.44346	4.40902 - 2.67066
0.50	3.58906	0.45319	4.47731 - 2.70081
0.60	3.52891	0.41445	4.34123 - 2.71659
0.70	3.55703	0.34821	4.23953 - 2.87453
0.80	3.63984	0.31084	4.24909 - 3.03059
0.90	3.91641	0.39466	4.68994 - 3.14288

TABLE 2

THE MEAN VALUE OF THE LENGTH OF TIME TO REACH THE TARGET SET

FOR VARIOUS VALUES OF PARAMETER a_1 AT NOISE LEVEL OF Q = 0.30

Initial State $(x_1, x_2) = (10.0, -1.60)$

The Minimum Length of time for the Deterministic Problem = 3.967

Parameter a_1	Mean Value of the Length of Time	Standard Deviation of Mean	95.0% Confidence Interval for Mean
-0.40	9.61797	0.99332	11.56487 - 7.67107
-0.20	8.62578	1.00071	10.58718 - 6.66438
0.00	7.28203	0.81170	8.87297 - 5.69109
0.10	7.21406	0.76920	8.72169 - 5.70643
0.20	7.40234	0.71707	8.80779 - 5.99689
0.30	6.86344	0.69437	8.23440 - 5.51248
0.40	6.70313	0.65064	7.97838 - 5.42788
0.50	6.45313	0.55572	7.54234 - 5.36392
0.60	6.11016	0.51141	7.11253 - 5.10779
0.70	5.84375	0.32265	6.47615 - 5.21135
0.75	5.62656	0.16982	5.95942 - 5.29370
0.80	5.86953	0.21894	6.29865 - 5.44040
0.85	6.03594	0.15849	6.34658 - 5.72530
0.90	7.45469	0.61102	8.65229 - 6.25709

TABLE 3

THE MEAN VALUE OF THE LENGTH OF TIME TO REACH THE TARGET SET FOR VARIOUS VALUES OF PARAMETER a_1 AT NOISE LEVEL OF Q = 0.30

Initial State $(x_1, x_2) = (15.0, -1.90)$

The Minimum Length of Time for the Deterministic Problem = 5.176

Parameter a_1	Mean Value of the Length of Time	Standard Deviation of Mean	95.0% Confidence Interval for Mean
-0.40	13.04688	1.24517	15.48742 - 10.60634
-0.20	12.10469	1.08969	14.24048 - 9.96889
0.00	10.11016	0.94364	11.95970 - 8.26062
0.10	9.35234	0.97951	11.27218 - 7.43250
0.20	8.73438	0.80353	10.30931 - 7.15945
0.30	8.43359	0.76234	9.92777 - 6.93941
0.40	7.89922	0.72782	9.32575 - 6.47269
0.50	7.25391	0.56637	8.36399 - 6.14383
0.55	7.20703	0.48514	8.15790 - 6.25616
0.60	7.19609	0.44118	8.06080 - 6.33138
0.65	7.27109	0.41480	8.08410 - 6.45807
0.70	7.74609	0.44739	8.62297 - 6.86921
0.75	7.58906	0.22481	8.02968 - 7.14844
0.80	8.11016	0.32478	8.74674 - 7.47358
0.85	8.61328	0.26887	9.14027 - 8.08629
0.90	9.57734	0.43506	10.43006 - 8.72462

TABLE 4

THE MEAN VALUE OF THE LENGTH OF TIME TO REACH THE TARGET SET FOR VARIOUS VALUES OF PARAMETER a_1 AT NOISE LEVEL OF Q = 0.05

Initial State $(x_1, x_2) = (5.00, -1.34)$

The Minimum Length of Time for the Deterministic Problem = 2.306

Parameter a_1	Mean Value of the Length of Time	Standard Deviation of Mean	95.0% Confidence Interval for Mean
-0.40	3.25234	0.29307	3.82677 - 2.67792
-0.20	3.15391	0.28702	3.71647 - 2.59134
0.00	2.75703	0.14693	3.04501 - 2.46905
0.10	2.65703	0.06847	2.79123 - 2.52283
0.20	2.65078	0.05997	2.76833 - 2.53323
0.30	2.65078	0.05997	2.76833 - 2.53323
0.40	2.65859	0.06060	2.77738 - 2.53981
0.50	2.80234	0.04998	2.90031 - 2.70438
0.60	2.85000	0.04275	2.93379 - 2.76621
0.70	3.01094	0.04513	3.09939 - 2.92248
0.80	3.14687	0.03934	3.22398 - 2.06977
0.90	3.51250	0.05817	3.62652 - 3.39848

TABLE 5

THE MEAN VALUE OF THE LENGTH OF TIME TO REACH THE TARGET SET FOR VARIOUS VALUES OF PARAMETER a_1 AT NOISE LEVEL OF Q = 0.05

Initial State $(x_1, x_2) = (10.0, -1.60)$

The Minimum Length of Time for the Deterministic Problem = 3.967

Parameter a_1	Mean Value of the Length of Time	Standard Deviation of Mean	95.0% Confidence Interval for Mean
-0.40	7.71719	0.47566	8.64948 - 6.78489
-0.20	6.22500	0.43248	7.07267 - 5.37733
0.00	5.25625	0.36304	5.96780 - 4.54470
0.10	4.90000	0.28942	5.46726 - 4.33274
0.20	4.56094	0.15434	4.86344 - 4.25843
0.30	4.47187	0.07194	4.61288 - 4.33087
0.40	4.51172	0.05865	4.62668 - 4.39676
0.50	4.61328	0.04316	4.69788 - 4.52868
0.60	4.86641	0.05754	4.97918 - 4.75363
0.70	5.17500	0.03753	5.24855 - 5.10144
0.80	5.87969	0.05279	5.98316 - 5.77621
0.90	6.74219	0.05950	6.85882 - 6.62556

TABLE 6

THE MEAN VALUE OF THE LENGTH OF TIME TO REACH THE TARGET SET FOR VARIOUS VALUES OF PARAMETER a_1 AT NOISE LEVEL OF Q = 0.05

Initial State (x_1, x_2) = (15.0, -1.90)

The Minimum Length of Time for the Deterministic Problem = 5.176

Parameter a_1	Mean Value of the Length of Time	Standard Deviation of Mean	95.0% Confidence Interval for Mean
-0.40	10.42969	0.46225	11.33571 - 9.52367
-0.20	8.76406	0.49389	9.73209 - 7.79603
0.00	7.26484	0.45589	8.15839 - 6.37129
0.10	6.53750	0.34484	7.21338 - 5.86161
0.20	5.76875	0.16631	6.09472 - 5.44278
0.30	5.58281	0.05078	5.68234 - 5.48328
0.40	5.68203	0.04339	5.76708 - 5.59698
0.50	6.08672	0.05143	6.18752 - 5.98591
0.60	6.42969	0.05219	6.53198 - 6.32740
0.70	7.03437	0.04654	7.12559 - 6.94316
0.80	7.92188	0.06211	8.04361 - 7.80014
0.90	9.41719	0.04908	9.51338 - 9.32099

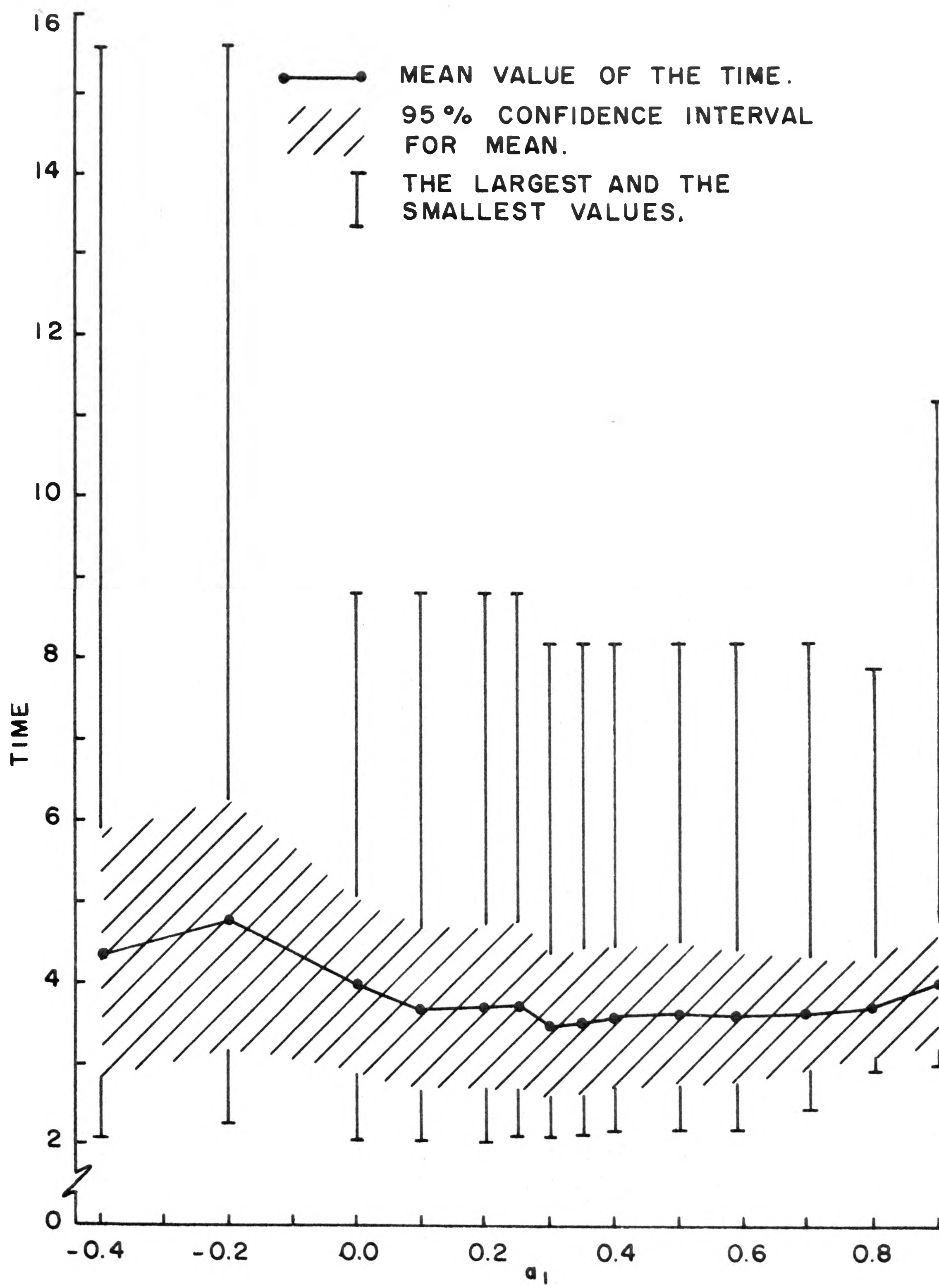

Figure 3. The Length of Time Required to Reach the Target Set for Different Rotations of the Switching Curve at Noise Level Q = 0.30. Initial State $(x_1, x_2) = (5.0, -1.34)$

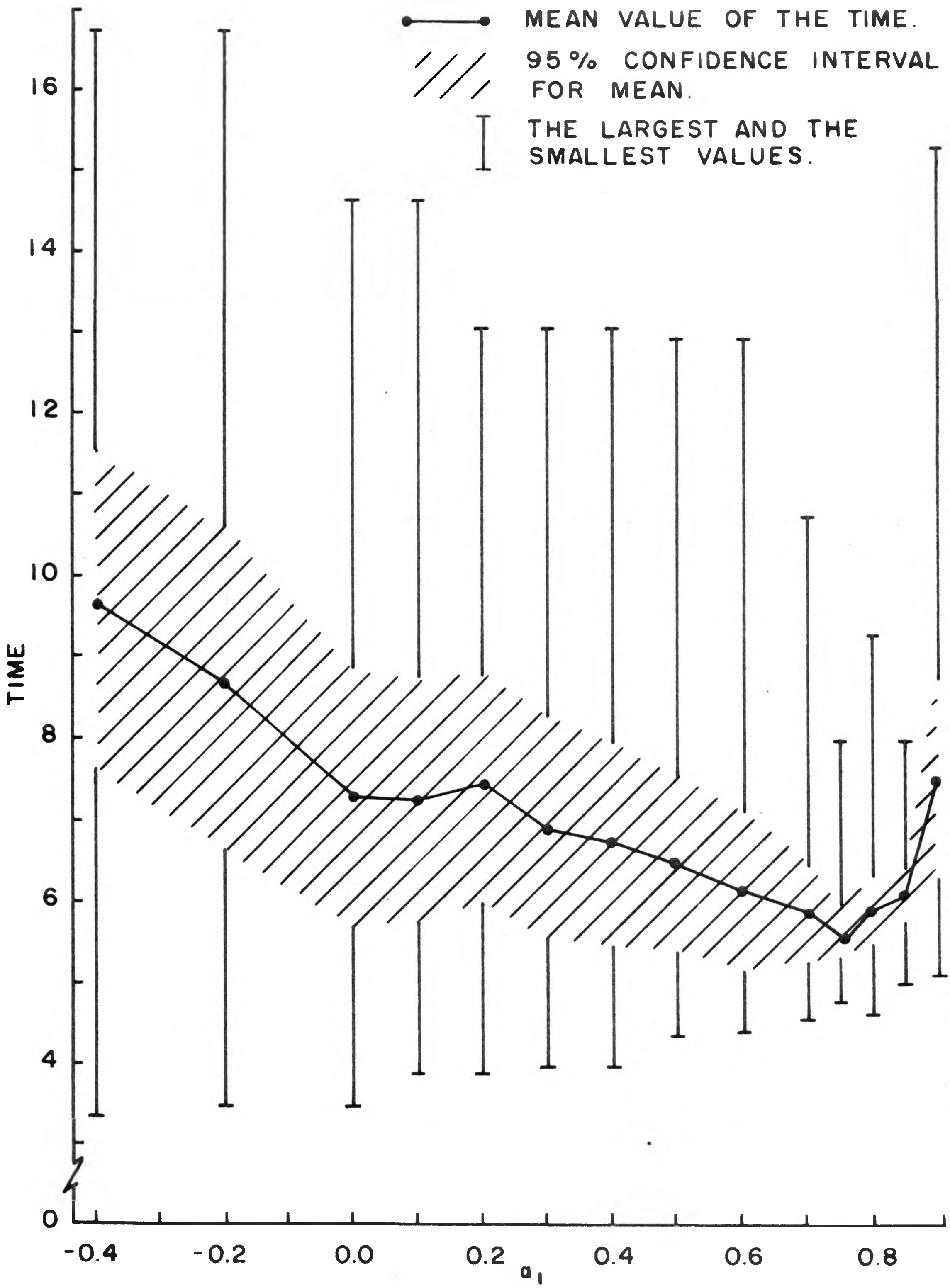

Figure 4. The Length of Time Required to Reach the Target Set for Different Rotations of the Switching Curve at Noise Level Q = 0.30. Initial State $(x_1, x_2) = (10.0, -1.60)$

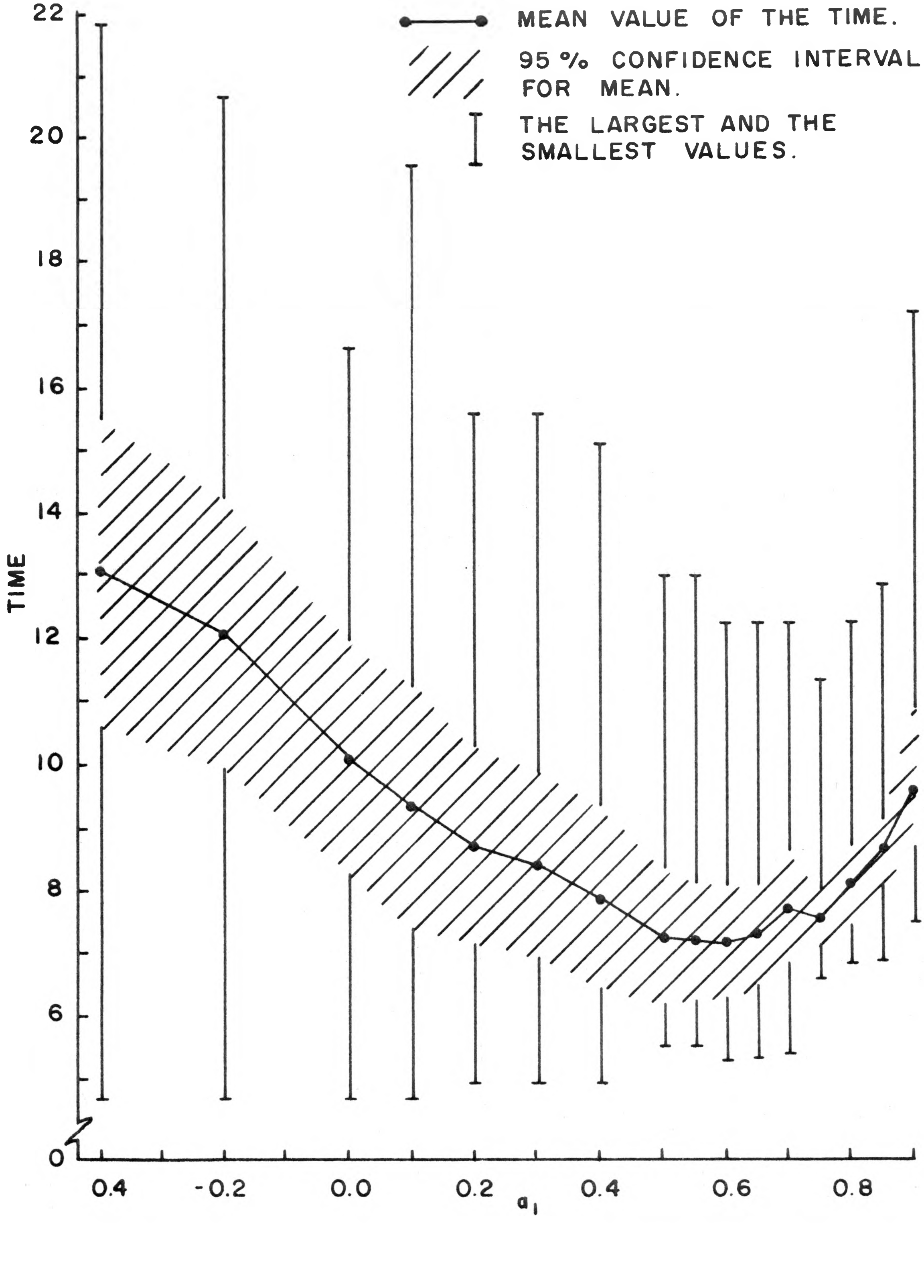

Figure 5. The Length of Time Required to Reach the Target Set for Different Rotations of the Switching Curve at Noise Level Q = 0.30. Initial State (x_1, x_2) = (15.0, -1.90)

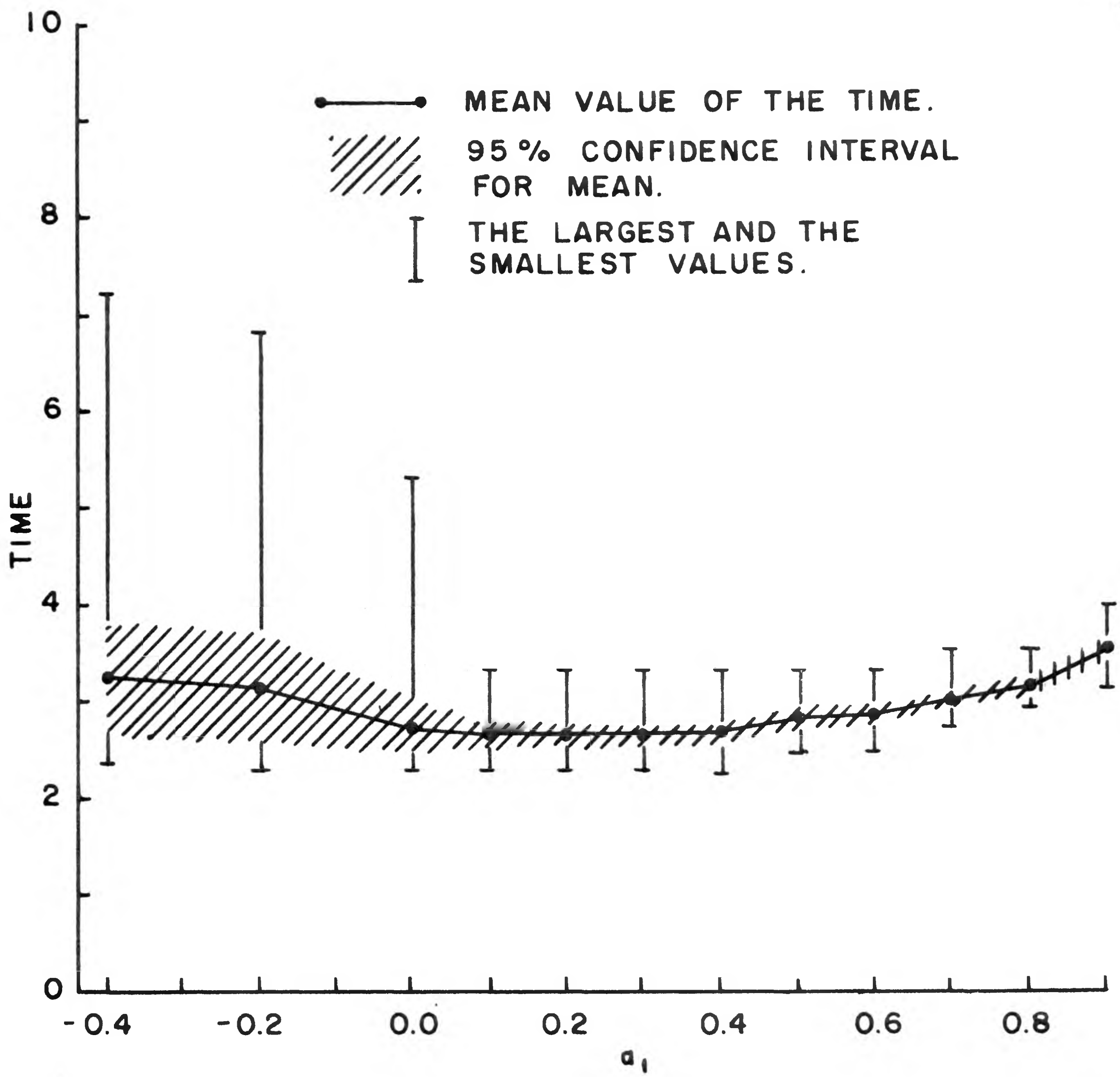

Figure 6. The Length of Time Required to Reach the Target Set for Different Rotations of the Switching Curve at Noise Level Q = 0.05. Initial State $(x_1, x_2) = (5.0, -1.34)$

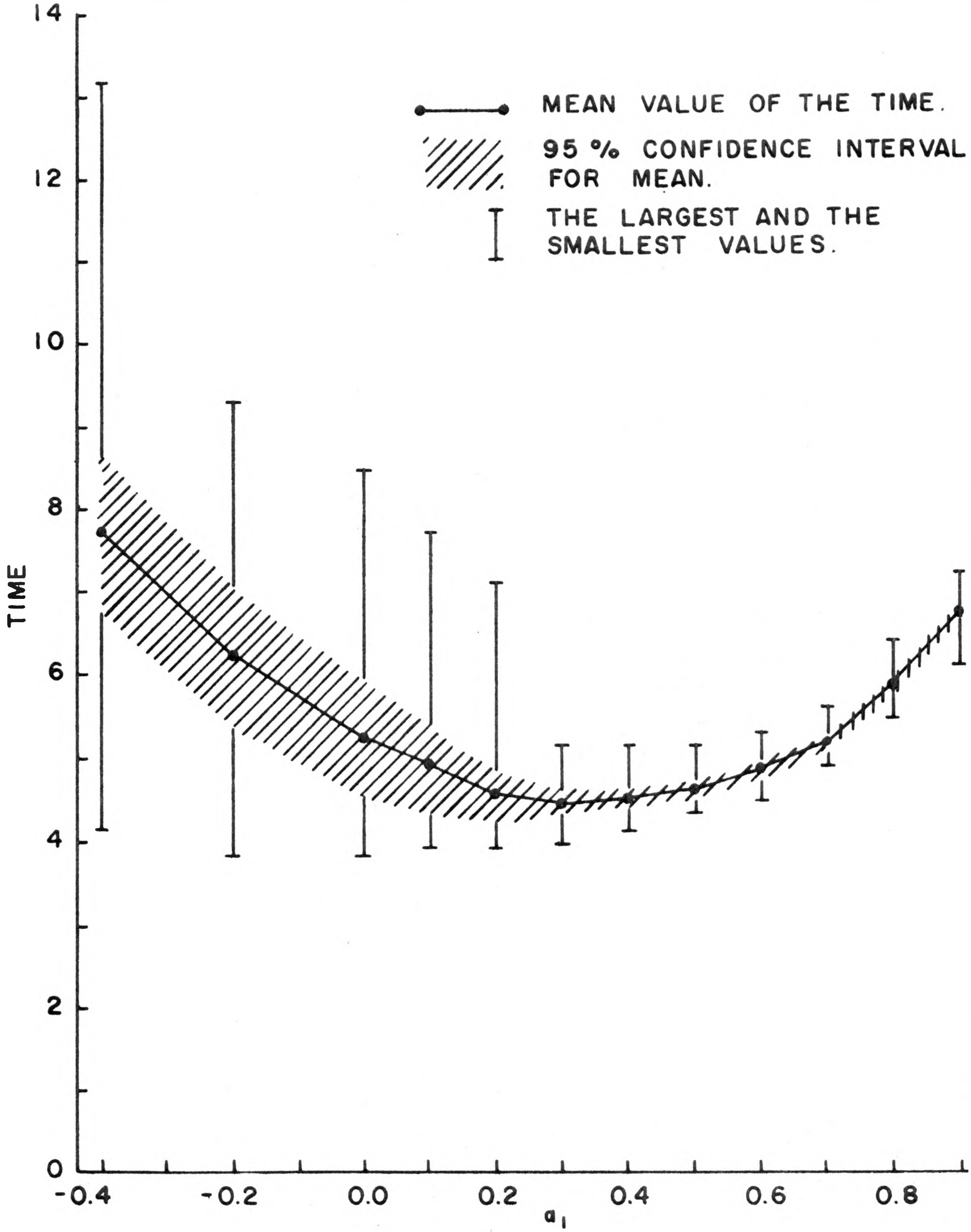

Figure 7. The Length of Time Required to Reach the Target Set for Different Rotations of the Switching Curve at Noise Level Q = 0.05. Initial State $(x_1, x_2) = (10.0, -1.60)$

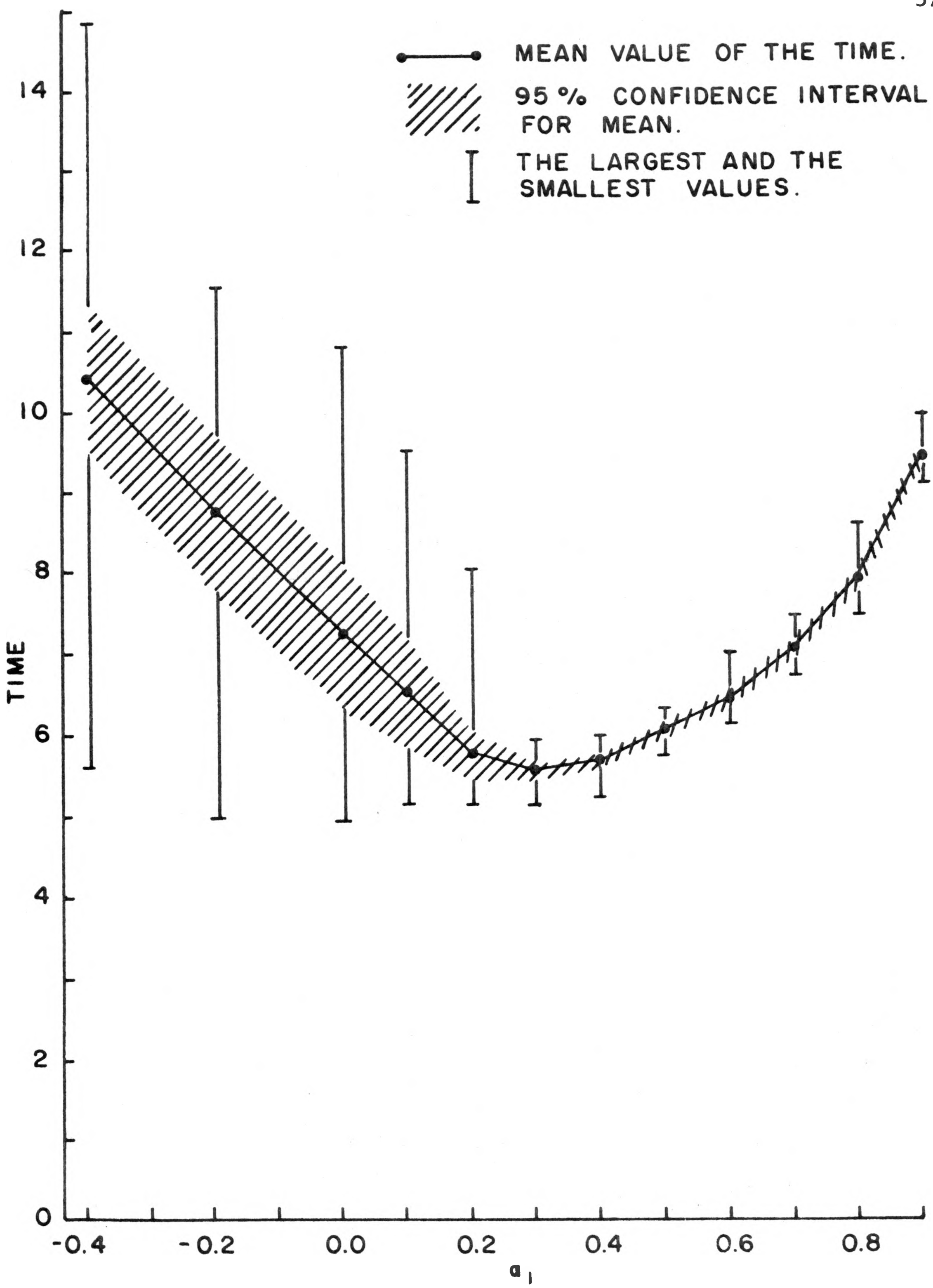

Figure 8. The Length of Time Required to Reach the Target Set for Different Rotations of the Switching Curve at Noise Level Q = 0.05. Initial State $(x_1, x_2) = (15.0, -1.90)$

right branch F_{s+} of the switching curve or below the left branch F_{s-}.

It would not be desirable to start from a point too far removed from the switching curve since the stochastic input introduces a random element into the length of time required in the repetitions for the first switching to occur. This variation would be irrelevant to the task of finding the optimal switching curve. Thus ideally the starting point should be just barely counterclockwise from the largest value of a_1 to be tested.

Some preliminary scouting computations indicated that the optimal switching curve for the stochastic system would lie somewhat counterclockwise from the deterministic curve -- i.e., the optimal value a_1* would be positive but less than unity. Thus the x_2 coordinates were computed using a value of 0.9 for a_1 in equation (6-10).

It is easy to see from Figures 3 through 8 that the optimum a_1 is positive. The phenomological reason for this behavior is that the controller cannot bring the system back to the switching curve if the stochastic inputs force it clockwise. Thus if the system wandered below F_{s+}, it frequently completely missed the target set on that trajectory. A convergence was not obtained until another switching sequence was encountered on F_{s-} (see trajectory B on Figure 1). Of course, such a path required more time than the more direct route. A more counterclockwise stochastic switching curve increased the probability of hitting the target on a direct route.

The confidence intervals for mean values that appeared in Tables 1 through 6 and Figures 3 through 8 were computed assuming a normal distribution. Even though the distribution of the repetitive trajectory times was skewed, the distribution of the means approached a normal distribution by virtue of the Central Limit theorem.

The variation of the length of trajectory time as a function of the parameter a_1 is very shallow for the case of the initial state nearest the target set (Figures 3 and 6). This implies that the probability of hitting the target set remains almost the same over a wide range of values of a_1. If this is the case, it appears that the optimal switching curve for the deterministic problem could be used as a good approximate solution to the stochastic problem. A stochastic switching strategy of multiple switchings is still required however.

The uncertainty of the required time at a given value of the parameter a_1 overshadowed the variations in required time due to changes in that parameter. Thus the exact location of the minimum is quite uncertain but the mean value of the corresponding penalty for operating at a near but a non-optimal point is small.

The minima shown in Figures 4, 5, 7 and 8 are more pronounced in relation to the uncertainty of the means. It is easy to see from Figures 7 and 8 that the minima in the low noise level case are more pronounced in relation to the high noise level.

If equation (6-10) were rigorous, the a_1* at minimal trajectory time would be the same for all initial states. The apparent differences shown on Figures 3, 4 and 5 for the high noise level case could be taken as an indication that higher order terms should be included in the Taylor's series expansion. Because of the large uncertainty in the location of the minima, however, retention of the higher order terms can not be justified.

Thus, the overall optimal switching curve can be constructed using nominal values of a_1*, that is, 0.75 for the variance of noise of 0.3 and 0.3 for the variance of noise of 0.05 as shown in Figure 9.

However, the exact locations of these minima are still subject to uncertainty. The statistical one-sided "t" test was used to establish whether minimum average time was significantly smaller than its neighbours at 0.05 level. The results are listed in Table 7.

TABLE 7

RANGE OF VALUES FOR a_1 FOR WHICH MINIMUM AVERAGE TIME VALUES ARE NOT SIGNIFICANTLY SMALLER THAN NEIGHBOURING VALUES AT 0.05 LEVEL

Initial State (x_1, x_2)	Noise Level Q = 0.30	Noise Level Q = 0.05
(5.0, -1.34)	-0.4 - 0.90	0.0 - 0.40
(10.0, -1.60)	0.4 - 0.80	0.1 - 0.40
(15.0, -1.90)	0.3 - 0.75	0.2 - 0.40

These results indicate that the following alternatives could be made for constructing an overall optimal switching curve. The optimal switching curve could be constructed in sections using nominal values of a_1* in the vicinity of each of the tested initial points. Alternatively, a simple polynomial for a_1* in terms of x_1 could be derived so that these "constants" in the Taylor's series would themselves be functions of distance. Finally, a switching curve could be constructed using an average value for a_1*. This average would be weighted more heavily for

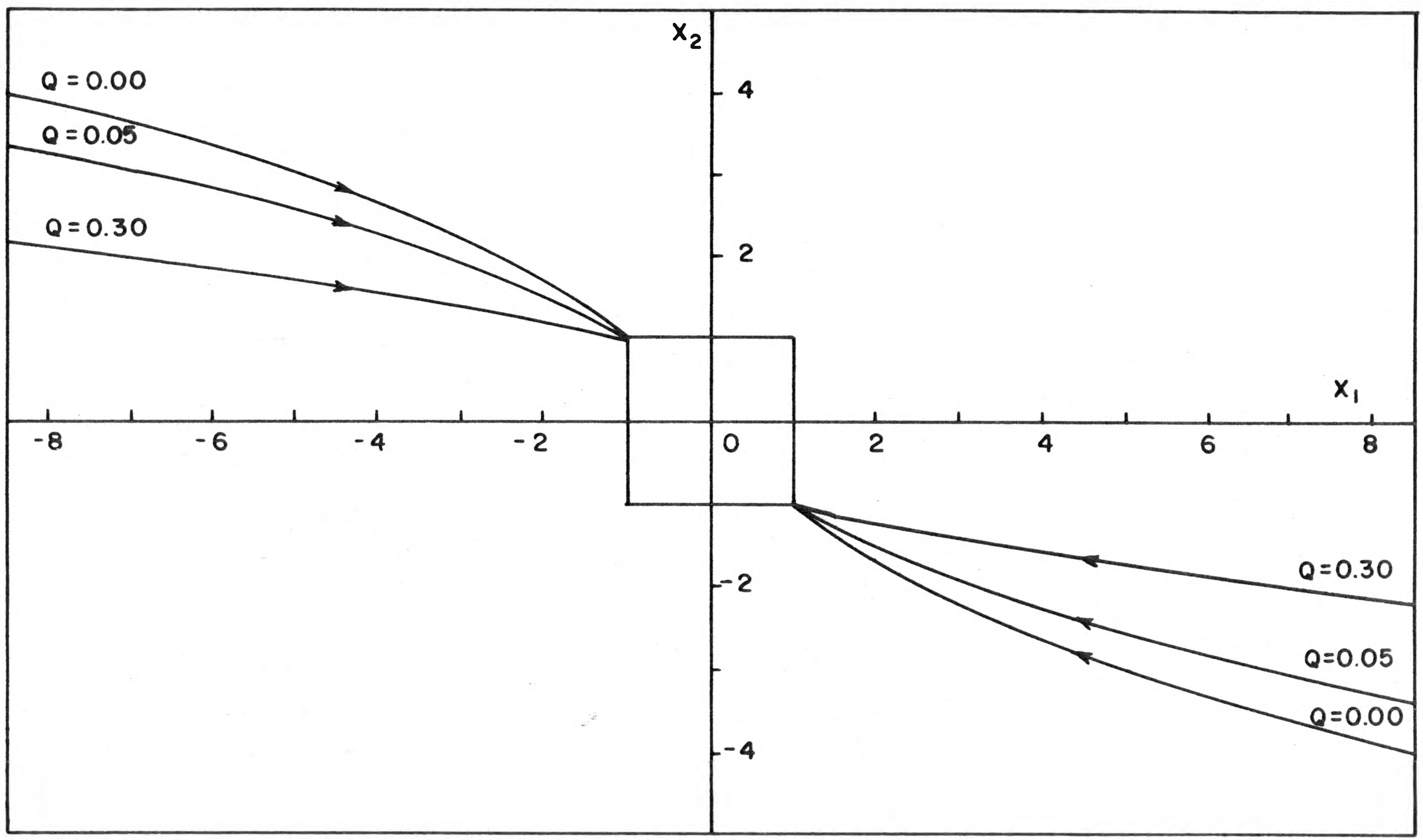

Figure 9. Location of the Optimal Switching Curve for Different Levels of Noise

the points found at the greater distance because the minima are more pronounced in those regions.

In view of the large variations seen in the length of times, only small penalties could be expected in operation using these various suboptimal switching strategies.

CHAPTER VII

CONCLUSIONS

The direct computational search using repetitive simulations for the difference between the deterministic and stochastic switching curves is a practical and useful approach to find time optimal controls for stochastic systems. This method also allows evaluation of the sensitivity of the expected value of the length of time to the switching curve location which corresponds to changes in the linear coefficient used in a truncated power series.

It is not necessary to retain more than the first order term in Taylor's series expansion for the difference because random variations in repetitive trajectory times overshadow the variations in the linear coefficient as a function of distance from the target set.

For systems with noise corrupted observations, the procedure described in this thesis can be followed except that the estimators for the state variables must be used.

CHAPTER VIII

RECOMMENDATIONS

The following recommendations are offered on the basis of the foregoing discussion:

1. In this work, a direct search using repetitive simulations was used for constructing the optimal switching curve for time optimal control in stochastic systems. Numerical results have shown that the exact locations of the minimum average time are subject to uncertainty. Further investigation for the optimal number of repetitions should be considered, and the direct search method can be replaced by another search procedure such as stochastic approximation.

2. There were the apparent differences in the optimum value of parameter a_1 for three different initial states. Furthermore the result of statistical test suggested a range of values for probable location of the optimum value of parameter a_1. In view of those uncertainties, switching range could be used for the "bang-bang" controller in stochastic systems with minimal penalty. Control switching does not occur as long as the state trajectory of the system remains within switching range. In this way less switchings would be required and the system could be moved to the desired terminal region without making additional clockwise cycle around the terminal region. The switching range method would be useful for implementation.

3. The rectangle was taken as the desired terminal region in this investigation. With that rectangular terminal region it was concluded that the stochastic switching curve would be located somewhat

counterclockwise from the deterministic curve. However, slightly clockwise location had been reported with the circular terminal region by others (7, 27). Thus further investigations should be made for the effect of terminal region configuration on the location of the optimal stochastic switching curve.

4. Since there is no guarantee that the state will remain in the desired terminal region once the region was reached, time optimal control problem in stochastic systems could be completely solved by combining the problem of maximizing the probability or the expected length of time that the state will stay in the terminal region.

NOMENCLATURE

NOMENCLATURE

Unless otherwise stated, numbers in parentheses at the end of the definitions refer to the equation number where the symbol was defined or first appeared.

A	=	Constant state coefficient 2×2 matrix (B-1)
a_i	=	Constant coefficient defined by (5-1)
$B(t)$, B	=	Control coefficient $n \times m$ matrix (3-1), (3-6)
B'	=	Transpose of matrix B (3-6)
$E\{\ \}$	=	Mathematical expectation (3-3)
F_D	=	Switching function for time optimal control in deterministic systems (6-7)
F_{D+}	=	Portion of the deterministic switching curve with the optimal control $u^* = +1$ (Figure 1)
F_{D-}	=	Portion of the deterministic switching curve with the optimal control $u^* = -1$ (Figure 1)
F_s	=	Switching function for time optimal control in stochastic systems (6-8)
F_{s+}	=	Portion of the stochastic switching curve with the optimal control $u^* = +1$ (Figure 1)
F_{s-}	=	Portion of the stochastic switching curve with the optimal control $u^* = -1$ (Figure 1)
$f(x(t,\omega),t)$, f	=	n-Vector valued function (3-1), (3-6)
$G(t)$, G	=	Noise coefficient $n \times p$ matrix (3-1), (3-6)
G'	=	Transpose of matrix G (3-6)
(g_{ij})	=	$n \times n$ Matrix defined by (3-6)
$H(x,p,u,t)$, H	=	Hamiltonian function in the Minimum Principle (A-2), (A-4)
I	=	2×2 Identity matrix (B-9)

$J(x_o,u)$, $J(x,u)$	=	Performance criterion for time optimal control in stochastic systems defined by (3-4), (3-5)
K	=	Positive constant (3-11)
k	=	Constant element of vector q defined by (A-7)
k_1	=	Constant element of vector q defined by (A-11)
n	=	Number of time intervals for approximating a continuous-time gaussian white noise (B-5)
$p_i(t)$	=	Costate variable (A-3)
$p_i(t_1)$	=	Costate variable $p_i(t)$ at the terminal time t_1 (A-9)
$\dot{p}_i(t)$	=	Time derivative of the costate variable $p_i(t)$ defined by (A-5)
Q	=	Constant variance of a scalar gaussian white noise (6-4)
$Q(t)$	=	Covariance matrix of p-vector Wiener process, $p \times p$ non-negative definite matrix, (3-3)
$Q_A(i)$	=	Covariance of a piecewise constant gaussian white sequence at time index i (B-7)
Q_+	=	Set of states (x_1, x_2) to the right of the deterministic switching curve F_D in the state space (Figure A-1)
Q_-	=	Set of states (x_1, x_2) to the left of the deterministic switching curve F_D in the state space (Figure A-1)
q	=	2-Dimension vector in tangent hyperplane to the target set S at the terminal time t_1 (A-8)
R^m	=	Space of m-tuples of real numbers (3-1)
S	=	Terminal region (target set) (6-3)
s	=	Distance from the origin along the switching surface (5-1)
sgn	=	Sign function (3-8)
t	=	Time (3-1)

t_i	=	Time index used in approximation of a gaussian white noise (B-8)
t_o	=	Initial time (3-1)
t_s	=	Time required to first reach the target set S (3-4)
t_1	=	Terminal time (A-9)
tr( )	=	Trace (3-6)
U	=	Compact convex set of admissible controls (3-5)
$u(x(t,\omega),t)$, $u(t)$, u	=	m-Vector control variable (3-1), (A-1), (3-5)
u^*, $u^*(t)$	=	Optimal control (3-8), (A-3)
$\|\|u(x,t)\|\|$	=	Norm of control vector u(x,t) (3-11)
$\|u\|$, $\|u(t)\|$	=	Absolute value of scalar control (6-4), (A-2)
$V(x)$	=	Optimal value of J(x,u) starting in the state x at time t and using optimal control (3-5)
$V_D(x)$, V_D	=	Minimum value of J(x,u) for the deterministic system (4-1), (5-1)
$V_s(x)$, V_s	=	Minimum value of J(x,u) for the stochastic system (4-1), (5-1)
$V_i(x)$	=	Coefficient defined in a singular perturbation method (4-1)
$w(t,\omega)$, $w(t)$	=	p-Vector gaussian white noise (3-1), (3-3)
$w'(t)$	=	Transpose of vector w(t) (3-3)
$w^{(n)}(t)$	=	Piecewise constant gaussian white sequence (B-5)
$w_i^{(n)}(t)$	=	Component of piecewise constant gaussian white sequence (Figure B-1)
$x(t,\omega)$, $x(t)$	=	n-Vector state (3-1), (B-1)
$x(t_o)$, x_o	=	Initial state at the initial time t_o (3-1)
$\dot{x}(t,\omega)$	=	Time derivative of $x(t,\omega)$ (3-2)
$\|\|x\|\|$	=	Norm of state vector x (3-11)

$x_i(t)$	=	ith Element of the state vector $x(t)$ (6-1)
$\dot{x}_i(t)$	=	Time derivative of $x_i(t)$ (6-2)
$\lvert x_i(t)\rvert$	=	Absolute value of $x_i(t)$ (6-3)

<u>Greek Letters</u>

$\delta(t)$	=	Diract delta function (3-3)
Δt	=	Time increment (3-11)
$\lvert\Delta t\rvert$	=	Absolute value of Δt (3-11)
Δx	=	State increment vector (3-11)
$\lVert\Delta x\rVert$	=	Norm of vector Δx (3-11)
$\eta(t,\omega)$, $\eta(t)$	=	p-Vector Wiener process (brownian motion) (3-1), (3-3)
$\eta'(t)$	=	Transpose of vector $\eta(t)$ (3-3)
τ	=	Time variable (B-2)
π_i	=	Initial value of the costate variable $p_i(t)$ at the initial time $t = 0$ (A-6)
$\Phi(t)$	=	State transition matrix (B-2)
$\dot{\Phi}$	=	Time derivative of $\Phi(t)$ (B-3)
Ω	=	Probability space (3-3)
ω	=	Element of a probability space Ω (3-1)

BIBLIOGRAPHY

BIBLIOGRAPHY

1. Ames, W. F. Nonlinear Partial Differential Equations in Engineering. New York: Academic Press, 1965.

2. Aoki, M. "Stochastic Time-Optimal Control System." Trans. AIEE (Applications and Industry), 80, 41 (1961).

3. Astrom, K. J. Introduction to Stochastic Control Theory. New York: Academic Press, 1970.

4. Benes, V. E. "Existence of Optimal Stochastic Control Law." SIAM Journal on Control, 9, 446 (1971).

5. Bryson, A. E., Jr., and Ho, Yu-Chi. Applied Optimal Control: Optimization, Estimation and Control. Waltham, Mass.: Blaisdell Publishing Company, 1969.

6. Bucy, R. S., and Joseph, P. D. Filtering for Stochastic Processes with Applications to Guidance. New York: Interscience Publishers, 1968.

7. Dorato, P., Hsieh, C. M., and Robinson, P. N. "Optimal Bang-Bang Control of Linear Stochastic Systems with a Small Noise Parameter." IEEE Trans. on Automatic Control, AC-12, 682 (1967).

8. Dreyfus, S. E. Dynamic Programming and Calculus of Variations. New York: Academic Press, 1965.

9. Eaton, J. H., and Zadeh, L. A. "Optimal Pursuit Strategies in Discrete State Probabilistic Systems." Trans. of ASME. Journal of Basic Engineering, 84, 23 (1962).

10. Fleming, W. H. "Optimal Continuous-Parameter Stochastic Control." SIAM Review, 11, 470 (1969).

11. Fleming, W. H. "Stochastic Control for Small Noise Intensities." SIAM Journal on Control, 9, 473 (1971).

12. Howard, R. A. Dynamic Programming and Markov Processes. Cambridge, Mass.: The M.I.T. Press, 1960.

13. Jazwinski, A. H. Stochastic Processes and Filtering Theory. New York: Academic Press, 1970.

14. Kailath, T., and Frost, P. "Mathematical Modeling of Stochastic Processes." Stochastic Problems in Control. A Symposium of the American Automatic Control Council, University of Michigan, Ann Arbor, Michigan, 1968. Edited by B. Friedland. New York: The American Society of Mechanical Engineers, 1968.

15. Katayama, T. "Stochastic Bang-Bang Controls that Maximize the Expectation of First Passage Time." International Journal of Control, 14, 177 (1971).

16. Kushner, H. J. Stochastic Stability and Control. New York: Academic Press, 1967.

17. Kushner, H. J. "On the Existence of Optimal Stochastic Controls." SIAM Journal on Control, 3, 463 (1966).

18. Kushner, H. J. "Sufficient Conditions for the Optimality of a Stochastic Control." SIAM Journal on Control, 3, 499 (1966).

19. Kushner, H. J., and Kleinman, A. J. "Numerical Methods for the Solutions of the Degenerate Elliptic Equations Arising in Optimal Stochastic Control Theory." IEEE Trans. on Automatic Control, AC-13, 344 (1968).

20. Kushner, H. J., and Kleinman, A. J. "Accelerated Procedures for the Solution of Discrete Markov Control Problems." IEEE Trans. on Automatic Control, AC-16, 147 (1971).

21. Kushner, H. J. Introduction to Stochastic Control. New York: Holt, Rinehart, and Winston, Inc., 1971.

22. Lim, Y. S. "Asymptotic Solution of an Optimal Servo Problem." IEEE Trans. on Automatic Control, AC-13, 45 (1968).

23. Meditch, J. S. Stochastic Optimal Linear Estimation and Control. New York: McGraw-Hill Book Company, 1969.

24. Novosel'tsev, V. N. "Time-Optimal Control Systems in the Presence of Random Noise." Automation and Remote Control, 23, 1519 (1962).

25. Rishel, R. "Necessary and Sufficient Dynamic Programming Conditions for Continuous Time Stochastic Optimal Control." SIAM Journal on Control, 8, 559 (1970).

26. Rishel, R. W. "Weak Solutions of a Partial Differential Equation of Dynamic Programming." SIAM Journal on Control, 9, 519 (1971).

27. Robinson, P. N., and Yurtseven, H. O. "Monte Carlo Methods for Stochastic Time-Optimal Control." IEEE Trans. on Automatic Control, AC-14, 574 (1969).

28. Schultz, P. R. "An Optimal Control Problem with State Vector Measurement Errors." Vol. 1 of Advances in Control Systems: Theory and Applications. Edited by C. T. Leondes. New York: Academic Press, 1964.

29. Stratonovich, R. L. "On the Theory of Optimal Control. An Asymptotic Method for Solving the Diffusive Alternative Equation." Automation and Remote Control, 23, 1352 (1962).

30. Van Mellaert, L. J. "Inclusion-Probability-Optimal Control." Research Report No. PIBMRI 1364-67, Department of Electrical Engineering, Polytechnic Institute of Brooklyn, 1967.

APPENDICES

APPENDIX A

DETAILED SOLUTION OF
TIME OPTIMAL CONTROL PROBLEM IN DETERMINISTIC SYSTEMS

The following outlines solution to the time optimal control problem in deterministic system.

Problem

The system is represented by

$$\begin{aligned} \dot{x}_1(t) &= x_2(t) \\ \dot{x}_2(t) &= u(t). \end{aligned} \tag{A-1}$$

The control u(t) is assumed to be constrained in magnitude by the relation

$$|u(t)| \leq 1 \text{ for all } t. \tag{A-2}$$

The problem is to find the admissible feedback control that transfers the system from the given initial state (x_{10}, x_{20}) at time t=0 to the target set $S = \{(x_1, x_2): |x_1| \leq 1 \text{ and } |x_2| \leq 1\}$ in the shortest possible time.

Detailed Solution

The Hamiltonian function is given by

$$H(x,p,u,t) = 1 + x_2(t)\ p_1(t) + u(t)\ p_2(t). \tag{A-3}$$

The control which minimizes the Hamiltonian is obtained:

$$u^*(t) = -\mathrm{sgn}\{p_2(t)\} = \pm 1. \tag{A-4}$$

The costate variables $p_1(t)$ and $p_2(t)$ satisfy the following equations:

$$\begin{aligned} \dot{p}_1(t) &= -\{\partial H/\partial x_1(t)\} = 0 \\ \dot{p}_2(t) &= -\{\partial H/\partial x_2(t)\} = -p_1(t). \end{aligned} \tag{A-5}$$

Let π_1 and π_2 be the initial values of the costate variable as follows:

$$\begin{aligned} \pi_1 &= p_1(0) \\ \pi_2 &= p_2(0). \end{aligned} \tag{A-6}$$

Then from equation (A-5) it follows that

$$\begin{aligned} p_1(t) &= \pi_1 = \text{constant} \\ p_2(t) &= \pi_2 - \pi_1\, t. \end{aligned} \tag{A-7}$$

In the problem of hitting the target set, the transversality conditions require that, at the terminal time t_1, the costate vector $p(t_1)$ be normal to a vector q which belongs to a hyperplane that is tangent to the target set S.

Suppose that the terminal state of the optimal trajectory belongs to the interior of $\overline{ab}$ in Figure A-1. Then the transversality conditions will hold. If, on the other hand, the terminal state is either at (1, 1) or at (-1, 1), then a unique tangent line to $\overline{ab}$ cannot be defined, and so any transversality conditions cannot be found. For this reason all four control sequences {+1}, {-1}, {+1, -1} and {-1, +1} can be candidates for optimal control for the terminal state at either (1, 1) or at (-1, 1).

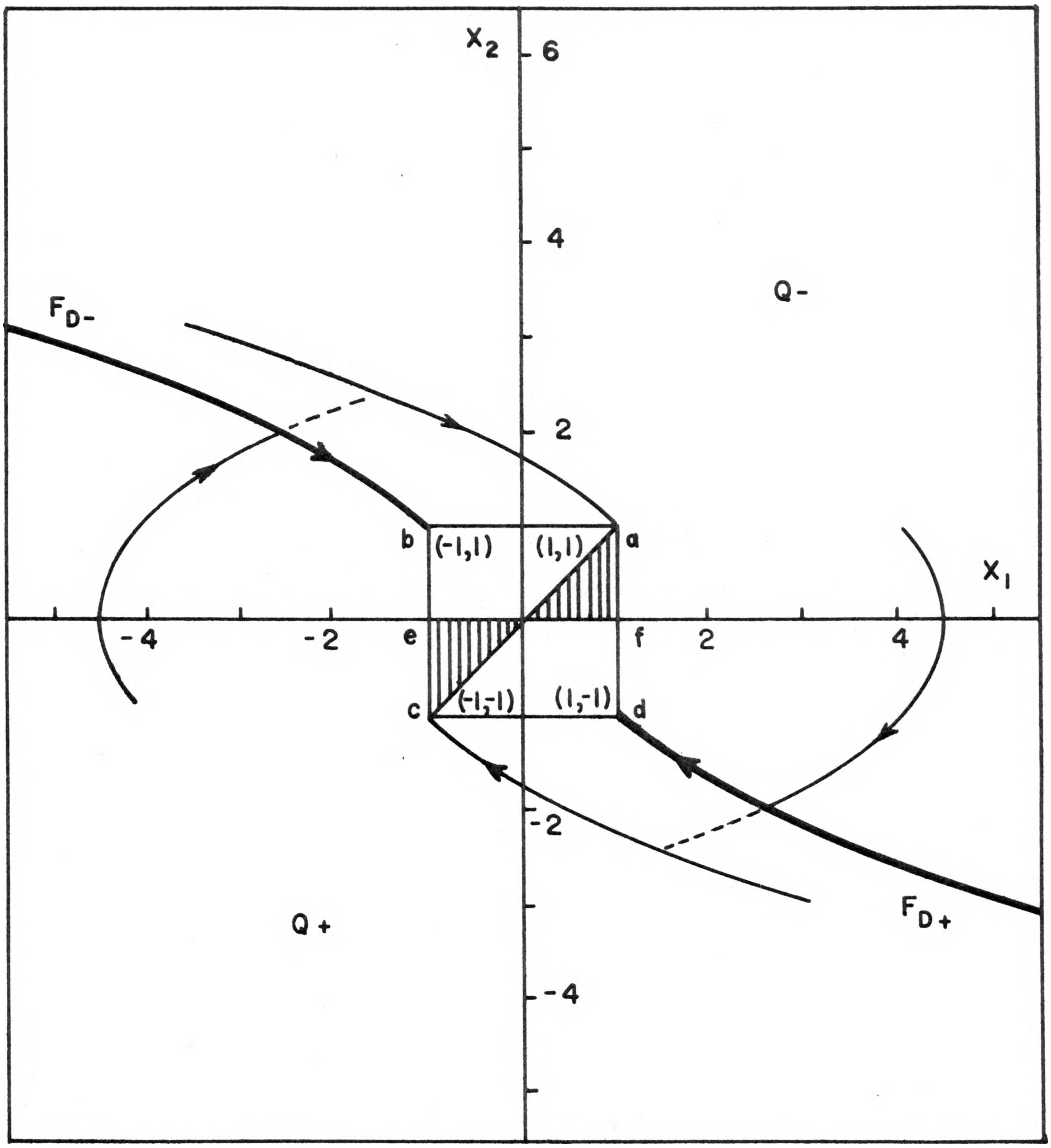

Figure A-1. Construction of the Deterministic Switching Curve for Time Optimal Control

Since $\overline{ab}$ is a straight line, a vector q in the tangent hyperplane which can be used in the transversality conditions has the form:

$$q = (k, 0)', \; k \neq 0. \tag{A-8}$$

Since p and q are normal at the terminal time t_1, it follows that

$$(p_1(t_1), p_2(t_1))(k, 0)' = 0, \tag{A-9}$$

from which it is found that

$$p_1(t_1) = 0. \tag{A-10}$$

But $p_1(t) = \pi_1$ is constant for all time; hence, $\pi_1 = 0$, and it follows that

$$p_2(t) = \pi_2 \neq 0. \tag{A-11}$$

Thus the time optimal control is u* = +1 or -1 and no switching can occur. Similarly, if the terminal state of the optimal trajectory belongs to the interior of $\overline{cd}$, the time optimal control is u* = +1, or -1 and no switching can occur. From the Minimum Principle, p(t) must be an outward directed normal to q. Hence it is seen that $p_2(t)$ is positive for the terminal state at the interior of $\overline{ab}$, i.e., u* = -1, and vice versa for $\overline{cd}$. For the terminal state either at (-1, -1) or at (1, -1), all four control sequences {+1}, {-1}, {+1, -1} and {-1, +1} can be candidates for optimal control.

Consider the terminal state belonging to the interior of $\overline{ad}$. A vector q on this boundary can be given by

$$q = (0, k_1)', \; k_1 \neq 0. \tag{A-12}$$

Hence

$$(p_1(t_1), p_2(t_1)) (0, k_1)' = 0. \tag{A-13}$$

implies $0\ p_1(t_1) + k_1\ p_2(t_1) = 0$ from which $p_2(t_1) = 0$. It follows immediately that $p_2(t) = \pi_1\ (t_1 - t)$, and $u^* = -1$ since $p_1(t_1) = \pi_1$ is positive in this case and furthermore no switching can occur.

By simple geometry it is obvious that no state on $\overline{af}$ can qualify as the terminal state of the time-optimal trajectory. Similarly, no switching can occur for the state on $\overline{bc}$ and the optimal control $u^* = +1$. No state on $\overline{ec}$ can qualify as the terminal state of the time-optimal trajectory.

Since switching can occur for the initial state to reach the terminal state at one of four states (1, 1), (-1, 1), (1, -1) and (-1, -1), it leads us to say that the switching curve is a portion of the time-optimal trajectory. It is easy to see from Figure A-1 that from any state in Q_- it takes less time to reach the state (1, -1) than the state (-1, -1). Similar reasoning can be used to establish that if state (x_1, x_2) is in Q_+, then only the control sequence $\{+1, -1\}$ can transfer (x_1, x_2) to either (-1, 1) or (1, 1) and that it takes less time to reach the state (-1, 1) than the state (1, 1).

By integration of equation (A-1)

$$\begin{aligned} x_1(t) &= x_1(0) + x_2(0)t + \tfrac{1}{2}\, u^* t^2 \\ x_2(t) &= x_2(0) + u^* t. \end{aligned} \tag{A-14}$$

By eliminating the time $t = u^*(x_2(t) - x_2(0))$, the optimal trajectory is obtained:

$$x_1(t) = x_1(0) + \tfrac{1}{2}\, u^*\, x_2(t)^2 - \tfrac{1}{2}\, u^*\, x_2(0)^2. \tag{A-15}$$

Since the terminal state belongs to either (-1, 1) or (1, -1) after switching, the switching function is obtained:

$$x_1 = u^* \, (x_2{}^2 + 1)/2. \tag{A-16}$$

APPENDIX B

INTEGRATION OF STOCHASTIC DIFFERENTIAL EQUATION

An outline of the mathematical formulation for a solution to stochastic differential equation (6-1) is presented and discussed in this Appendix. When written in vector-matrix notation, equation (6-1) becomes

$$\begin{aligned} dx(t) &= \{Ax(t) + Bu(t)\}dt + Gd\eta(t) \\ x(t_o) &= x_o \end{aligned} \tag{B-1}$$

where $\eta(t)$ is a scalar Wiener process (Brownian motion) satisfying $E\{d\eta(t)\} = 0$ and $E\{d\eta(t)d\eta'(t)\} = Q(t)dt$;

x_o is a given initial state as the fixed constant;

A, B and G are given by

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \; B = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \text{ and } G = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Since equation (B-1) is a linear time-invariant vector stochastic differential equation, it has a solution given by

$$x(t) = \Phi(t-t_o)x(t_o) + \int_{t_o}^{t} \Phi(t-\tau)Bu(\tau)d\tau + \int_{t_o}^{t} \Phi(t-\tau)Gd\eta(\tau), \tag{B-2}$$

where $\Phi(t)$ is the state transition matrix satisfying,

$$\dot{\Phi}(t) = A\Phi(t). \tag{B-3}$$

Equation (B-2) is the unique solution of equation (B-1), since the linear stochastic differential equation (B-1) satisfies the conditions of the existence and uniqueness theorem for stochastic differential equations (6).

As a good approximation to $\int\Phi(t)Gd\eta(t)$ in equation (B-2), the physically realizable process $\int\Phi(t)Gw(t)dt$ has been used (14). Thus equation (B-2) becomes

$$x(t) = \Phi(t-t_o)x(t_o) + \int_{t_o}^{t} \Phi(t-\tau)Bu(\tau)d\tau + \int_{t_o}^{t} \Phi(t-\tau)Gw(\tau)d\tau, \quad \text{(B-4)}$$

where w(t) is a scalar gaussian white noise satisfying $E\{w(t)\} = 0$ and $E\{w(t)w'(\tau)\} = Q(t)\delta(t-\tau)$ for $t, \tau \geq t_o$. Further discussions of stochastic integrals and additional references are given (10, 13, 14, 28).

In applications where a digital computer is used for a solution, it is usually necessary to approximate a continuous-time gaussian white noise by a multistage process. A continuous-time gaussian white noise $\{w(\tau),\ t_o \leq \tau \leq t\}$ can be defined to be the limit of the gaussian white sequence (5, 23):

$$\{w(\tau),\ t_o \leq \tau \leq t\} = \lim_{n\to\infty} \{w^{(n)}(\tau),\ t_o \leq \tau \leq t,\ n\Delta t = t-t_o\}. \quad \text{(B-5)}$$

For some given value of n, $\{w^{(n)}(\tau),\ t_o \leq \tau \leq t\}$ denotes the piecewise constant gaussian white sequence as depicted in Figure B-1. The value over each interval is defined from the left.

The covariance $Q_A(t_o + i\Delta t) = Q_A(i)$ of the piecewise constant gaussian white sequence is to be replaced by $Q(t_o + i\Delta t)/\Delta t$ in order to approximate the gaussian white noise w(t) with the given Q(t) as defined in equation (B-5). The covariance $Q_A(i)$ of the approximated gaussian white noise obviously depends strongly on the size of time increment Δt. The time increment Δt in approximation of the gaussian white noise should be smaller than time constants of the system for a gaussian random

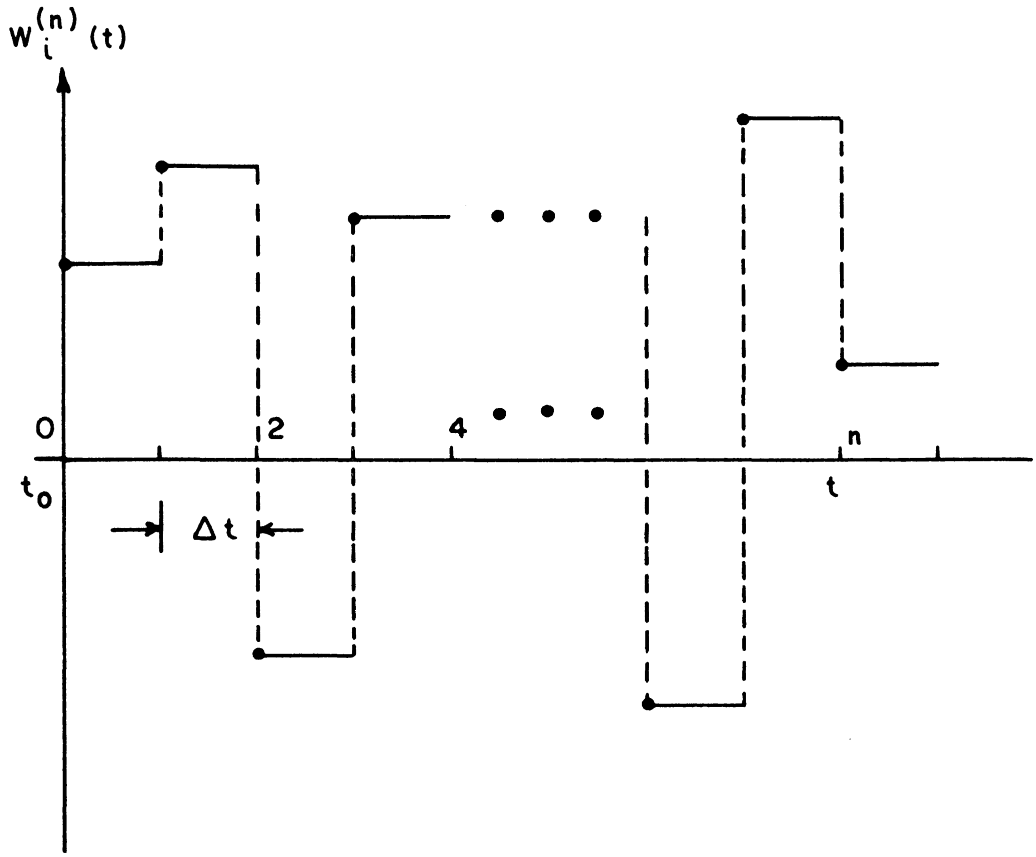

Figure B-1. Components of Piecewise Constant Gaussian White Sequence Sample Function sequence being regarded as white noise relative to the system.

Thus scalar gaussian white noise is approximated by a piecewise constant gaussian random sequence,

$$w^{(n)}(\tau) = w^{(n)}(t) = \text{constant, for } t \le \tau \le t + \Delta t \tag{B-6}$$

with the covariance $Q_A(i) = Q(t)/\Delta t$, (B-7)

where t corresponds to the time point i.

Equation (B-4) is then approximated as follows:

$$x(t) = \Phi(t-t_o)x(t_o) + \int_{t_o}^{t} \Phi(t-\tau)Bu(\tau)d\tau + \sum_{i=0}^{n-1} \{\int_{t_i}^{t_{i+1}} \Phi(t_{i+1} - \tau)Gw^{(n)}(t_i)d\tau\}, \tag{B-8}$$

where $t_i = t_o + i\Delta t$;

$\Delta t = (t - t_o)/n.$

It is straightforward from equation (B-3) that the state transition matrix $\Phi(t-\tau)$ is $\exp\{A\ (t - \tau)\}$ and is most easily computed directly in this problem as given below:

$$\exp\{A\ (t - \tau)\} = I + A\ (t - \tau) + A^2\ (t - \tau)^2/2! + \ldots\ldots, \tag{B-9}$$

where I is the identity matrix. Since $A^2 = A^3 = \ldots\ldots = 0$, the state transition matrix is given by

$$\Phi(t-\tau) = \exp\{A\ (t - \tau)\} = I + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} (t - \tau) = \begin{pmatrix} 1 & (t-\tau) \\ 0 & 1 \end{pmatrix}. \tag{B-10}$$

By substituting equation (B-10) into equation (B-8) and using the optimal admissible feedback control u*, a solution of the numerical example given in Chapter VI is represented by

$$
\begin{aligned}
x_1(t) &= x_1(t_o) + x_2(t_o)(t - t_o) + u^* (t - t_o)^2/2 + \\
&\qquad \sum_{i=0}^{n-1} \{(\Delta t^2/2)\ w^{(n)}(t_o + i\Delta t)\} \\
x_2(t) &= x_2(t_o) + u^* (t - t_o) + \sum_{i=0}^{n-1} \{(\Delta t)\ w^{(n)}(t_o + i\Delta t)\}.
\end{aligned}
\tag{B-11}
$$

Note that if numerical method is used for integration, the time increment for numerical integration should be, at most, same as the time increment used in the approximation of a gaussian white noise.

The undersigned, appointed by the Dean of the Graduate Faculty, have examined a thesis entitled

TIME OPTIMAL CONTROL IN STOCHASTIC SYSTEMS

presented by Won-Kyoo Lee

a candidate for the degree of Doctor of Philosophy

and hereby certify that in their opinion it is worthy of acceptance.

Richard H. Luecke

Richard M. Angus

Robert W. McLaren