CONTROL AND STABILITY THEORY IN THE SPACE OF MEASURES

Electrical Abraham Boyarsky, B.Eng. (Hons.), M.Eng. Ph.D.

ABSTRACT

This thesis treats some problems in stochastic control and stability theory from the point of view of flows (induced by the stochastic systems) in the space of probability measures. In the first part of the dissertation, the concept of attainable set of probability measures for a stochastic process is introduced, and the following results are obtained for a control system modelled by a stochastic differential equation, where the control is additive in the drift coefficient :

(i) Employing the martingale approach initiated by Stroock and Varadhan, a "stochastic bang-bang principle" is proved. It follows from the proof of this principle that, for a certain class of controls, the attainable set of probability measures is weak compact.

(ii) Assuming the "target" set of probability measures is a continuous function of time with respect to a certain topology, a time-optimal stochastic control theorem is demonstrated.

(iii) The existence of unique quasi-diffusions for the class of drift coefficients, which are bounded and integrable, is verified, and a necessary and sufficient condition for the average of a cost functional, to be minimized by a feedback control, is derived.

In the second part of the dissertation, stability properties of general stochastic systems are investigated. The following work is carried out :

(iv) A theory of dynamical systems on the space of probability measures is formulated where the relevant topology is that of weak convergence.

(v) In this dynamical system framework, a new definition for stochastic stability is proposed, which is weaker than any other previously studied. Employing the concept of D-functions, some conditions for stability are obtained for trajectories in the space of probability measures.

CONTROL AND STABILITY THEORY IN THE SPACE OF MEASURES

by

Abraham Boyarsky, B.Eng. (Hons.), M.Eng.

Department of Electrical Engineering

McGill University,

Montreal, Canada.

CONTROL AND STABILITY THEORY IN THE SPACE OF MEASURES

by

Abraham Boyarsky, B.Eng. (Hons.), M.Eng.

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy.

Department of Electrical Engineering,
McGill University,
Montreal, Canada.

August, 1970.

ABSTRACT

This thesis treats some problems in stochastic control and stability theory from the point of view of flows (induced by the stochastic systems) in the space of probability measures. In the first part of the dissertation, the concept of attainable set of probability measures for a stochastic process is introduced, and the following results are obtained for a control system modelled by a stochastic differential equation, where the control is additive in the drift coefficient :

(i) Employing the martingale approach initiated by Stroock and Varadhan, a "stochastic bang-bang principle" is proved. It follows from the proof of this principle that, for a certain class of controls, the attainable set of probability measures is weak compact.

(ii) Assuming the "target" sets of probability measures is a continuous function of time with respect to a certain topology, a time-optimal stochastic control theorem is demonstrated.

(iii) The existence of unique quasi-diffusions for the class of drift coefficients, which are bounded and integrable, is verified, and a necessary and sufficient condition for the average of a cost functional, to be minimized by a feedback control, is derived.

In the second part of the dissertation, stability properties of general stochastic systems are investigated. The following work is carried out :

(iv) A theory of dynamical systems on the space of probability measures is formulated where the relevant topology is that of weak convergence.

(v) In this dynamical system framework, a new definition for stochastic stability is proposed, which is weaker than any other previously studied. Employing the concept of D-functions, some conditions for stability are obtained for trajectories in the space of probability measures.

ACKNOWLEDGEMENTS

The author wishes to express his sincere thanks to Professor D. Dawson of the Department of Mathematics for his guidance and patience during the preparation of this dissertation.

The author also wishes to acknowledge his co-director, Professor M. Fukada, Professor W. Fleming of Brown University for some helpful comments, and Dr. J. Komlos for his assistance with Lemma 3.1 . Thanks are also extended to Mrs. Peggy Hyland, Miss Louise Smith and Mrs. Sue Brunton for their fine job of typing, and to the National Research Council of Canada for their financial assistance. Special gratitude is due 'frieda' for her unfailing support and for her optimism when, so very often, all seemed lost.

TABLE OF CONTENTS

ABRIDGED LIST OF SYMBOLS

$I \; = \; [0, t_f] \subset R, \; t_f < +\infty, \; R$ - the real line.

θ A real number.

$E = R^n$ n-dimensional euclidean space.

l Lebesque measure on R.

$B(R)$ The space of real valued bounded measurable functions on R.

$\mathcal{B}(E)$ The σ-algebra on E generated by the open sets of E.

$\tilde{\mathcal{B}}(I) \equiv \left\{ \chi_{[0, t]} : t \in I \right\}$.

$B(I, \mathfrak{X})$ The space of Bochner integrable functions from I, with the Lebesque measure, into the Banach space $\mathfrak{X}$.

$C(E)$ The space of real valued bounded continuous functions on E.

$C_o(E)$ The space of real valued bounded continuous functions on E vanishing at ∞.

$\Omega \equiv C(I, E)$ The space of continuous functions from I into E.

$C_o^\infty(E)$ The space of real valued C^∞ functions on E having compact support.

$\mathcal{G}$ The space of bounded measurable functions from $I \times R$ into R.

$\Sigma^o \equiv \left\{ \chi_F : F \text{ measurable subset of } I \right\}$.

$\Sigma \equiv \left\{ u \in \mathcal{L}_\infty(I) : 0 \le u(t) \le 1, \; t \in I \right\}$.

$\tilde{\Sigma} \equiv \left\{ u \in \mathcal{L}_\infty(I \times R) : 0 \le u(t, x) \le 1, \; t \in I, x \in R \right\}$

h The Prohorov-Hausdorff metric on $\mathcal{K}$.

$\mathcal{K}$ The space of non-empty weak compact subsets of $\mathcal{M}_1(R)$.

$\mathcal{L}(\mathfrak{X}, \mathfrak{X})$	The space of bounded linear operators from $\mathfrak{X}$ into $\mathfrak{X}$, where $\mathfrak{X}$ is a Banach space.
$\mathcal{M}(E)$	The space of signed measures on E.
$\mathcal{M}_1(E)$	The space of probability measures on E.
ρ	Prohorov metric on $\mathcal{M}_1(E)$.
P_x^n, Q	Probability measures on Ω.
E^P	The integral over Ω with respect to the probability measure P.

$$\eta^u(t)\ \varphi(\cdot) = \int_R P^u(0, x, t, \cdot)\ \varphi(dx) .$$

$$\mathring{\mathcal{R}}_\varphi^t \equiv \left\{ \eta^u(t)\,\varphi : u \in \Sigma^o \right\} , \quad \mathring{\mathcal{R}}_\varphi^t(f) \equiv \left\{ \eta^u(t)\,\varphi(f) : u \in \Sigma^o \right\}$$

$$\mathcal{R}_\varphi^t \equiv \left\{ \eta^u(t)\,\varphi : u \in \Sigma \right\} , \quad \mathcal{R}_\varphi^t(f) \equiv \left\{ \eta^u(t)\,\varphi(f) : u \in \Sigma \right\}$$

$$\tilde{\mathcal{R}}_\varphi^t \equiv \left\{ \eta^u(t)\,\varphi : u \in \tilde{\Sigma} \right\} , \quad \tilde{\mathcal{R}}_\varphi^t(f) \equiv \left\{ \eta^u(t)\,\varphi(f) : u \in \tilde{\Sigma} \right\}$$

$S(\varphi, \delta)$	A closed δ ball of $\varphi \in \mathcal{M}_1(R)$ in the weak topology on $\mathcal{M}_1(R)$.
$\tau(t)$	The target set in $\mathcal{M}_1(R)$.
$\left\{T(s, t) : s, t \in I\right\}$	The family of bounded linear operators from $C_o(R)$ into $C_o(R)$ generated by (2.4) with $u \equiv 0$.
$\left\{U^u(s, t) : s, t \in I\right\}$	The family of bounded linear operators from $C_o(R)$ into $C_o(R)$ generated by (2.4) with control u.

W	Wiener measure on Ω .
z_t	Brownian motion.
ϵ_x	Dirac measure at $x \in E$
φ, ψ, μ	Elements in $\mathfrak{M}_1(R)$.
f, g	Elements in $C_o(R)$.
$\sigma(\mathfrak{X}^*, \mathfrak{X})$	Weak*topology on $\mathfrak{X}^*$; $\mathfrak{X}$ is a Banach space.
π_t	The coordinate function from $C(I, E)$ into E .
$\psi(t) \equiv M\sqrt{\frac{2\pi}{a t}}$	Where M and a are positive constants.
$\rho \equiv \int_0^{t_f} \psi(t)\, dt$	
$\lambda(t) \equiv \psi(t) + \gamma\rho e^{\gamma\rho t}$	Where γ is a positive constant.
$\theta \equiv \int_0^{t_f} \lambda(s)\, ds$	
$\Vert \cdot \Vert$	Is used to represent both the norms on $C_o(R)$ and $\mathfrak{M}(R)$; which is meant will be obvious from the element it operates on.
$\eta^u(t) \equiv U^u(t)^*$	The adjoint of $U^u(t)$.
$m(t) \equiv T^*(t)$	

CHAPTER I

INTRODUCTION

1.1 Historical Background

The classical theory of deterministic linear control systems was developed during the 1930's and 1940's. Methods such as the Nyquist criterion and root locus technique are some of the classical tools that evolved. The systems were all modelled by autonomous differential equations since the control analysis rested heavily, in one way or another, on controlling the location of the poles and zeros of the system transfer function.

During the 1950's, aerospace applications revealed the inadequacies of the stationarity assumption in the control system models, and led to further investigation of time domain methods. The calculus of variations, dynamic programming and Pontryagin's maximum principle are the most important state space methods that developed, providing a satisfactory deterministic control theory.

The first results in stochastic stability theory appear to be in the work by Andronov, Pontryagin and Witt [28]. Motivated by Lyapunov's work on the stability of ordinary differential equations and the fundamental studies of Kolmogorov on Markov processes, Andronov et al. investigated the probabilistic behaviour of the sample paths of some Markov processes. However, not until the modern theory of Markov processes was developed did much progress take place in stochastic stability theory. With the advent of the Ito stochastic integral and its calculus [49], many stability results were obtained which were in spirit very similar to their deterministic analogues. Much of the early work on the stability of diffusions is due to Khasminskii [29] and Kushner [34].

The connections between stability and control theory were first investigated in the West in the early 1960's by Kalman and Bertram [30]. The necessity for stability analysis of control systems is due to the fact that optimized control systems may not be stable. Using the second method of Lyapunov for the design of optimal controllers assures that the optimal solution is asymptotically stable. Thus, for optimal system design, the problems of control and stability are closely related.

Since there existed substantial results on the control and stability theories for deterministic systems by the early 60's, it seemed plausible that at least some of the deterministic results could be extended to stochastic systems. Although, conceptually, the extension is straight-forward, the mathematics involved in the stochastic analysis is fairly complex, depending heavily on the theory of Markov processes and involving such concepts as infinitesimal generators, supermartingales and stopping times. Most of this basic work can be found in the book by Kushner [34] which includes a chapter on the design of stochastic controllers using stochastic Lyapunov functions. Some other extensions of the deterministic control theory, not dealt with in [34], are as follows. In [22] stochastic Lagrange multipliers are studied. The problem of partial observability of diffusions is investigated in [23]. In [24] Fleming and Nisio consider a more general model for the control system than the usual Ito stochastic differential equation. In [26] and [27], Kushner derives two stochastic maximum principles. For a review of some recent developments in optimal stochastic control theory, see the survey paper [25].

1.2 Stochastic Control Theory

Most of the stochastic control problems studied in the literature take one of the following forms.

(i) Given an initial point x_0 in the range space, determine the control, constrained to be in a certain admissible class, which transfers x_0 to a target set in the range space with probability one.

(ii) Given x_0, determine the control which maximizes the probability of hitting a target set.

(iii) Given x_0, determine the control which transfers x_0 to a target set with probability one and minimizes the average of some preassigned cost functional.

(iv) Given x_0, determine the control which approximates as closely as possible a specified path in some suitable (probabilistic) manner.

The optimal stochastic control problem which consists of simply minimizing the average of a cost functional is readily formulated using dynamic programming [34, Chapter IV] . Unfortunately, the optimal control in general depends upon the solution of a complex non-linear partial differential equation which is rarely amenable to analytic methods.

As we can see from the above formulations of stochastic control problems, the mathematical analysis involves the process sample paths. The sample paths of a process are analogous to the trajectories of a deterministic system, and therefore studying the sample paths appears to be the natural framework for the investigation of stochastic control problems. We shall, however, take a different approach based on viewing the control process in a certain space of measures.

In general, for Markov processes, the equation for the dynamics of the probability measures, on the range space of the process, induced by the random variables as a function of time is described by the transition function of the process. For many purposes the resulting representation for the flow of probability measures is not adequate since the transition function conceals its dependence on the parameters of the process, for example, the coefficients of a diffusion. We shall employ a more revealing and descriptive characterization for the flow of probability measures associated with a non-stationary stochastic differential equation where the control is additive in the drift coefficient. The resulting representation is important in establishing a necessary and sufficient condition for the existence of optimal stochastic controls (Section 5.3).

The key point in our functional analytic approach is the correspondence of a trajectory in the space of measures to each control function. Whereas it makes no sense to consider the attainable sets in the range space of a stochastic process, it is natural to consider the attainable sets of probability measures for the process. This point is discussed further in the next section.

1.3 Stochastic Stability Theory

As in deterministic stability theory, the main tool in the study of the stability of stochastic processes is the Lyapunov function. For each Lyapunov concept of stability in the deterministic case, there exist at least three stochastic analogues, corresponding to the three usual modes of convergence of a family of random variables.

To explain this in more detail, let $x = 0$ be the equilibrium solution whose stability properties are being investigated. Let $x(t; x_0, t_0)$ denote the solution of an ordinary n-dimensional differential equation with initial state x_0 at time t_0.

Definition 1.1 Deterministic Lyapunov Stability

The equilibrium solution is said to be stable if given any $\epsilon > 0$, there exists $\delta(\epsilon, t_0) > 0$ such that for $\| x_0 \| < \delta$

$$\sup_{t \geq t_0} \| x(t; x_0, t_0) \| < \epsilon \quad ,$$

where $\| y \| = \left(\sum_{i=1}^{n} | y_i |^2 \right)^{1/2}$.

To transform Definition 1.1 to stochastic forms, we write the convergence of the random variable $\sup_{t \geq t_0} \| x(t; x_0, t_0, \omega) \|$ in the three common modes of probabilistic convergence, where $(\Omega, \mathfrak{F}, P)$ is some underlying probability space. From now on, the generic element $\omega \in \Omega$ will be suppressed.

Definition 1.2 Lyapunov Stability in Probability

The equilibrium solution is stable in probability if given ϵ, $\epsilon' > 0$, there exists $\delta(\epsilon, \epsilon', t_0)$ such that $\| x \| < \delta$ implies

$$P \left\{ \sup_{t \geq t_0} \| x(t; x_0, t_0) \| > \epsilon' \right\} < \epsilon \quad .$$

Definition 1.3 Lyapunov Stability in the rth Mean

The equilibrium solution is stable in the rth mean if the rth moments of the solution process exist, and given $\epsilon > 0$, there exists $\delta(\epsilon, t_0)$ such that $\| x_0 \|_r < \delta$ implies

$$E^P\left\{ \sup_{t \geq t_0} \| x(t\ ; x_0, t_0) \|_r^r \right\} < \epsilon \quad ,$$

where $\| y \|_r = \sum_{i=1}^{n} | y_i |^r$.

Definition 1.4 Almost Sure Lyapunov Stability

The equilibrium solution is said to be almost surely P stable if

$$P\left\{ \lim_{\| x_0 \| \to 0} \sup_{t \geq t_0} \| x(t\ ; x_0, t_0) \| = 0 \right\} = 1 \quad .$$

Definition 1.4 is equivalent to saying that Definition 1.1 holds for almost every $\omega \epsilon \Omega$ with respect to P . For each of the above definitions, there exists a related definition for asymptotic stability. These concepts and other definitions of stochastic stability can be found in the survey paper by Kozin [31] .

Lyapunov stability in probability is weaker than the stability concepts defined by Definitions 1.3 and 1.4 . For some applications, this type of convergence may be of little interest since it does not imply convergence of the sample paths. However, where the expectation of a continuous function of the process is required to converge, this mode of convergence is sufficient. In fact, one of the motivations for the stability work in this dissertation is that Lyapunov stability in the mean and a.s. Lyapunov stability demand too much of the process. We shall work with a type of stability which is, in general, even weaker than that of Definition 1.2 , but may still be quite adequate for many applications. Rather than study stability properties with respect to a point, for instance, the null solution $x = 0$, we shall be concerned with the stability properties of trajectories of probability measures.

For a stochastic process, it is more reasonable to have the distribution functions, associated with the process, approach a probability measure (in the weak topology) than to have it approach a set in the range space, in the sense of convergence in probability. Another inadequacy of Definitions 1.2 - 1.4 is that the stochastic stability is with respect to initial points in the range space; often one does not know exactly where a process starts, and has at most only an estimate of the initial distribution function. Therefore, it seems more realistic to study stochastic stability with respect to initial probability measures rather than points in the range space. The approach to stochastic stability that we shall present takes all the above points into consideration. This is accomplished through the use of dynamical system theory [32].

It is impossible to define a dynamical system for a stochastic process, when it is regarded as a measurable function on a probability space, since the theory of dynamical systems can only be used in situations where the present state completely specifies the future states. This is, of course, not true for stochastic processes. Even Roxin's theory of attainability functions [37], which gives rise to a more generalized dynamical system is of little value since the range of many stochastic processes is the entire space, thereby yielding no information. Instead of trying to define a dynamical system on the range space of a process, we shall find it rewarding to define a dynamical system on the space of probability measures.

To reiterate, the object of the stability work in this dissertation is to present a different approach to some stochastic stability problems by examining these problems from the point of view of dynamical system theory, where the basic space is the space of probability measures. The advantages of this approach for these problems (Section 7.1) are:

(i) Studying the stochastic systems in the space of measures obviates detailed knowledge of the sample path behaviour.

(ii) The stochastic theory benefits from the completeness and structure of the (deterministic) dynamical system theory.

(iii) The deterministic and stochastic theories are unified under the concept of non-deterministic dynamical system.

Using the Prohorov metric in the space of probability measures, a theory of stochastic stability is formulated based on a slightly modified form of the usual theory of dynamical systems. Then a certain continuous function, defined on the range space of the process, is introduced whose existence assures the stability of flows in the space of probability measures.

Although the control and stability sections are essentially independent of each other, we justify their presence in the same dissertation by recalling that stability concepts are important in the design of optimal control systems. Also, the fundamental concept of controllability in the theory of control is closely related to stability theory, as can be seen by Theorems 1 and 2 of [44]. A final reason for treating stochastic control and stability theory together is that the approach proposed in this work, that of studying the process as trajectories in the space of measures, can be applied effectively to both these theories. In fact, an alternate title for this dissertation could be: "Flows of Probability Measures Induced by Stochastic Differential Equations with Application to Stochastic Control and Stability Theory".

CHAPTER II

PRELIMINARIES AND SUMMARY

2.1 Notation

Let $C([0,\infty), R^n)$ be the space of continuous functions on $[0,\infty)$ into R^n. By a continuous Markov process on R^n, we mean a family of probability measures $P_{s,x}$, $s \geq 0$, $x \in R^n$ on $\Omega \equiv C([0,\infty), R^n)$ such that $P_{s,x}(\Gamma | \mathfrak{F}_t^s) = P_{t,\pi_t}(\Gamma)$ a.s. $P_{s,x}$, $\Gamma \in \mathfrak{F}_\infty^t \equiv \mathfrak{F}^t$, where $\mathfrak{F}_t^s$ is the smallest σ-algebra on Ω with respect to which all the coordinate functions $\pi_\tau = \pi(\tau, \cdot): \Omega \longrightarrow R^n$, $s \leq \tau \leq t$ are measurable.*

Let $z_t = z(t, \cdot)$ be n-dimensional Brownian motion and let $a = a_i(t,x)$, $1 \leq i \leq n$, and $b = b_{ij}(t,x)$, $1 \leq i,j \leq n$, be a set of coefficients which satisfy certain smoothness conditions. The mathematical model which will be used in this dissertation is the Ito stochastic differential equation. In Ito's theory a measure $P_{s,x}$ is defined on Ω by making a nonlinear transformation of Brownian motion: for $\omega \in \Omega$ the Brownian path $\{ z(t,\omega) : t \geq 0 \}$ is transformed into the path $\{ x(t,\omega): t \geq s \}$ such that $x(s,\omega) = x_0$, and

$$dx(t,\omega) = a(t,x(t,\omega))dt + b(t,x(t,\omega))\, dz(t,\omega), \quad t \geq s, \tag{2.1}$$

where $dx(t,\omega)$ denotes the infinitesimal increment in x during the time interval $[t, t+dt]$, and $dz(t,\omega)$ denotes the corresponding increment in z. When a and b are Lipschitz continuous, Ito shows that for a set in Ω having Wiener measure one, there exists a unique continuous solution $x(t,\omega)$ satisfying these requirements. Hence, a measure P_{s,x_0} can be defined on Ω by setting

---

* For the reader unfamiliar with basic measure theory, see [8, Vol. II, Appendix] or [42].

$$P_{s,x}\left\{ \pi_{t_1} \in \Gamma_1 , \ldots\ldots , \pi_{t_m} \in \Gamma_m \right\} = W \left\{ x_{t_1} \in \Gamma_1 , \ldots , x_{t_m} \in \Gamma_m \right\},$$

where $s \le t_1 < \ldots < t_m$, $m \ge 1$, and $\Gamma_1 , \ldots , \Gamma_m \in \mathcal{B}(R^n)$*, where W is the Wiener measure on Ω .

To explain the meaning of (2.1) , we write it as the stochastic integral equation:

$$x_t = x_0 + \int_s^t a(s,x_s)ds + \int_s^t b(s,x_s)\, dz_s \quad ,$$

where $\int bdz$ is interpreted in Ito's sense, i.e.,

$$\int_s^t b(s,x_s)\, dz_s = \underset{h \to 0}{\text{l.i.m}} \sum_{i=0}^{N-1} b(\tau_i , x_{\tau_i}) (z_{\tau_{i+1}} - z_{\tau_i}) \quad ,$$

where $s = \tau_1 < \tau_2 < \ldots < \tau_N = t$, $h = \max (\tau_{i+1} - \tau_i)$, and l.i.m means in the mean square sense on Ω .

We shall find it useful to reformulate the meaning of a stochastic differential equation. In the new formulation a solution to (2.1) is a probability measure P on Ω such that

$$d\pi_t = a(t, \pi_t)\, dt + b(t, \pi_t)\, dz_t \quad , t \ge s , \text{ a.s. } P ,$$

where $\{ z_t ; t \ge s \}$ is a Brownian motion with respect to P .

Let a be bounded and measurable, b bounded and continuous, and the matrix c strictly elliptic.† Given $x \in R^n$, the following question is asked in [5]:

---

* $\mathcal{B}(R^n)$ is the σ-algebra on R^n generated by the open sets of R^n .

† $c = b^T b$, where T represents the transpose operation. Strict ellipticity means that $\sum_{i,j=1}^n c_{ij}\, \xi_i\, \xi_j \ge \nu \sum_{i=1}^n \xi_i^2$ for all vectors ξ , where $\nu > 0$.

Does there exist a probability measure $P_{s,x}$ on $(\Omega, \mathfrak{F}^s)$ such that $P_{s,x}\{\pi_s = x\} = 1$, and for any $\theta \in R^n$

$$y_\theta^s(t) = \exp\left\{(\theta, \pi_t - \pi_s) - \int_s^t (\theta, a(\tau, \pi_\tau))\, d\tau - \frac{1}{2}\int_s^t (\theta, c(\tau,\pi_\tau)\theta)\, d\tau\right\}$$

is a martingale on $(\Omega, \mathfrak{F}_t^s, t \geq s \geq 0, P_{s,x})$, where $(\cdot,\cdot)$ is the usual scalar product in R^n. We shall refer to this as the "martingale problem" and $P_{s,x}$, if it exists, is called a solution of the martingale problem starting at time s and state x.

The following important result is proved in [5]:

Theorem 2.1

Let a and b be as in the above paragraph. Then

(i) for each pair $s \in [0,\infty)$, and $x \in R^n$, there is one and only one probability measure $P_{s,x}$ on $(\Omega, \mathfrak{F}^s)$ which solves the martingale problem starting at s and x.

(ii) The system $(\Omega, \mathfrak{F}_t^s, t \geq s \geq 0, P_{s,x}, x \in R^n)$ is a continuous strong Markov process.

(iii) For each $x \in R^n$, there is an n-dimensional Brownian motion $\{z_t : t \geq 0\}$ such that

$$\pi_t = x + \int_s^t a(\tau, \pi_\tau)d\tau + \int_s^t b(\tau,\pi_\tau)dz_\tau, \quad t \geq s$$

a.s. with respect to $P_{s,x}$.

(iv) For $f \in C_0^\infty(R^n)$, the set of C^∞ functions from R^n into R having compact support, the measure $P_{s,x}$ satisfies

$$E^{P_{s,x}}\left\{f(\pi_t)\right\} - f(x) = E^{P_{s,x}}\left\{\int_s^t A(\tau) f(\pi_\tau)d\tau\right\}, \quad t \geq s, \tag{2.2}$$

where

$$A(\tau) = \sum_{i=1}^{n} a_i(\tau,\cdot)\frac{\partial}{\partial y_i} + \frac{1}{2}\sum_{i,j=1}^{n} c_{ij}(\tau,\cdot)\frac{\partial^2}{\partial y_i \partial y_j} .$$

A process satisfying (iv) is called a <u>quasi-diffusion</u>. For quasi-diffusions, $E^{P_{s,x}} f(\pi_t)$ need not satisfy the partial differential equation $\frac{\partial v}{\partial t} = A(t)v$, which is the case for diffusions. In Chapter V, we shall prove the existence of unique quasi-diffusions for a certain class of coefficients.

The concept of weak convergence of a sequence $\{P_n\}$ of measures on $(\Omega, \mathfrak{F}^o)$ is defined as follows: P_n converges to P weakly if for every bounded and continuous function $f(\omega)$ on Ω

$$\lim_{n\to\infty} \int_\Omega f(\omega)\, P_n(d\omega) = \int_\Omega f(\omega)\, P(d\omega) .$$

We shall denote this convergence by: $P_n \Rightarrow P$. We shall also use weak convergence on the real line, i.e., the sequence of measures $\{\mu_n\}$ on R converges weakly to the measure μ on R if and only if

$$\int_R f(x)\,\mu_n(dx) \to \int_R f(x)\,\mu(dx) \qquad \text{as } n \to \infty$$

for all f in $C(R)$, the space of bounded continuous functions on R .

2.2 Adjoint Semi-Groups

Adjoint semi-groups were first studied by Feller [40]. The general theory of adjoint semi-groups, using the adjoint infinitesimal generator, was studied by Phillips [41]. Below, we state some of the main results of the theory following [32, Chapter I].

Let $\mathfrak{X}$ be a real Banach space having norm $\| \ \|$, and let $\mathcal{L}(\mathfrak{X},\mathfrak{X})$ be the Banach algebra of bounded linear operators on $\mathfrak{X}$ to $\mathfrak{X}$. If $T \in \mathcal{L}(\mathfrak{X},\mathfrak{X})$, $\|T\|$ denotes the norm of T.

Definition 2.1 If $T(t)$ is an operator function on the non-negative real axis $0 \le t < \infty$ to $\mathcal{L}(\mathfrak{X},\mathfrak{X})$ satisfying the following conditions:

(i) $T(t_1 + t_2) = T(t_1)T(t_2)$ for $t_1, t_2 \ge 0$,

(ii) $T(0) = I$,

where I is the identity operator, then $\{T(t): 0 \le t < \infty\}$ is called a one-parameter semi-group of operators in $\mathcal{L}(\mathfrak{X},\mathfrak{X})$. The semi-group $\{T(t): 0 \le t < \infty\}$ is said to be of class $(\mathcal{C}_0)$ if it also satisfies the property

(iii) $\operatorname*{s-lim}_{t \downarrow 0} T(t)x = x$ for all $x \in \mathfrak{X}$,

called the strong continuity property of $T(t)$ at the origin.

Definition 2.2 The s-infinitesimal generator A of the semi-group $\{T(t): 0 \le t < \infty\}$ is defined by

$$Ax = \operatorname*{s-lim}_{h \downarrow 0} A_h x \quad , \tag{2.3}$$

where

$$A_h = \frac{1}{h}(T(h) - I) \quad ,$$

whenever the limit exists. The domain of A, $\mathfrak{D}(A)$, is the set of elements $x \in \mathfrak{X}$ for which the limit in (2.3) exists.

Proposition 2.2

(a) $\mathfrak{D}(A)$ is a linear manifold in $\mathfrak{X}$ and A is a linear operator.

(b) If $x \in \mathfrak{D}(A)$, then $T(t)x \in \mathfrak{D}(A)$ for all $t \ge 0$ and

$$\frac{dT(t)x}{dt} = AT(t)x = T(t)Ax \qquad , \quad t \geq 0$$

$$T(t)x - x = \int_0^t T(s)Ax\,ds \qquad , \quad t \geq 0$$

(c) $\mathfrak{D}(A)$ is dense in $\mathfrak{X}$, and A is a closed operator.

Let $\mathfrak{X}^*$ be the dual (or adjoint) space of all bounded linear functionals x^* on $\mathfrak{X}$. $\mathfrak{X}^*$ is a Banach space with the norm

$$\|x^*\| = \sup_{\|x\| \leq 1,\, x \in \mathfrak{X}} |x^*(x)| .$$

<u>Proposition 2.3</u>

Let U be a linear operator with domain $\mathfrak{D}(U)$ dense in $\mathfrak{X}$ to $\mathfrak{X}$.

(a) The dual operator U^* is a weak* closed linear operator. If in addition U is bounded, then $U^* \in \mathfrak{L}(\mathfrak{X}^*, \mathfrak{X}^*)$ and $\|U^*\| = \|U\|$.*

(b) If U is closed, then $\mathfrak{D}(U^*)$ is weak* dense in $\mathfrak{X}^*$ and, if $\mathfrak{X}$ is reflexive, $\mathfrak{D}(U^*)$ is strongly dense in $\mathfrak{X}^*$.

We now state the fundamental properties of the adjoint semi-group of bounded linear operators $\{T^*(t) : t \geq 0\}$.

<u>Proposition 2.4</u>

Let $\{T(t) : t \geq 0\}$ be a semi-group of operators of class $(\mathscr{C}_0)$ in $\mathfrak{L}(\mathfrak{X}, \mathfrak{X})$. Then $T^*(t)$ is an operator function on the interval $t \geq 0$ into $\mathfrak{L}(\mathfrak{X}^*, \mathfrak{X}^*)$ with $\|T^*(t)\| = \|T(t)\|$ for $t \geq 0$. Moreover,

(i) $T^*(t_1)\,T^*(t_2) = T^*(t_1 + t_2)$ for $t_1, t_2 \geq 0$

---

* See [1] or [3] for some basic results concerning the weak* topology.

(ii) $T^*(0) = I$ (identity operator on $\mathfrak{X}^*$)

(iii) $\text{weak*} \lim_{t \downarrow 0} T^*(t)\, x^* = x^*$ for $x^* \in \mathfrak{X}^*$ (weak* continuity of $T^*(t)$ at the origin) .

Under the hypothesis of the previous proposition, we have

Proposition 2.5

(a) The dual A^* of the infinitesimal generator A of the given semi-group $\{ T(t) : t \geq 0 \}$ is a weak* closed linear operator and its domain $\mathfrak{D}(A^*)$ is weak* dense in $\mathfrak{X}^*$.

(b) If $x^* \in \mathfrak{D}(A^*)$, $T^*(t)\, x^* \in \mathfrak{D}(A^*)$ for $t \geq 0$, and $A^*T^*(t)\, x = T^*(t)A^*x$.

Furthermore,

$$T^*(t)\, x^*(x) - x^*(x) = \int_0^t T^*(s)\, A^*x^*(x)\, ds \quad \text{for all } x \in \mathfrak{X},\ t > 0 .$$

(c) An element $x^* \in \mathfrak{X}^*$ belongs to the domain of A^* if

$$A^*_h x^* \equiv \frac{(T^*(h) - I)}{h}\, x^* \qquad h > 0$$

converges in the weak* topology of $\mathfrak{X}^*$ as $h \downarrow 0$, and the weak* limit is equal to A^*x^* .

Corollary 2.6

The dual operator A^* is equal to the weak* infinitesimal generator of the dual semi-group.

2.3 Summary of Results*

Consider the one-dimensional stochastic differential equation

$$dx_t^u = \left[a(t\ ,\ x_t^u) + u(t,x_t^u) \right] dt + b(t,x_t^u)\ dz_t \tag{2.4}$$

on the finite time interval $I = [0, t_f]$, where x_t^u takes values in the entire real line R . z_t is one-dimensional Brownian motion; a , u are bounded measurable functions from the cartesian product $I \times R$ into R , and b is a bounded continuous function from $I \times R$ into R such that $b^2(t,x) \geq \nu > 0$ for all $t \in I$ and $x \in R$ i.e., b is strictly elliptic. The function u is referred to as the control and acts only on current states. Often in this dissertation, u is simply a function of t, i.e. u is an open-loop control; we shall refer to (2.4) for this case as well. The existence of a unique continuous Markov process associated with (2.4) is assured by the theory in [5] .

The transition function of the Markov process (2.4) induces a two-parameter flow $\{ U^u(s,t): s,t \in I \}$ of bounded linear operators on the Banach space of bounded measurable functions from R into R (with the supremum norm), $B(R)$. $\{ U^u(s,t):s,t \in I \}$ is defined by

$$U^u(s,t)\, f(x) \equiv \int_R f(y)\ P^u(s,x,t\ ,\ dy)\ ,$$

where $P^u(s,x,t,\Gamma)$ is the transition function of (2.4) corresponding to the control u.

By Theorem 2.1 , we know that for $f \in C_o^\infty(R)$,

$$U^u(s,t)f \ = \ f + \int_s^t U^u\,(s,\tau)(A(\tau) + u(\tau,\ \cdot)D)\ f\ d\tau\ , \tag{2.5}$$

where

$$A(\tau) \ = \ a(\tau,\cdot)\,\frac{\partial}{\partial x} \ + \ \frac{b^2(\tau,\cdot)}{2}\ \frac{\partial^2}{\partial x^2} \tag{2.6}$$

and $D \equiv \dfrac{\partial}{\partial x}$.

---

* See also Chapter VIII , Remark (vii).

Let $C_o(R)$ be the Banach space (with the supremum norm) of continuous functions that vanish at $\pm\infty$. The dual space of $C_o(R)$ is $\mathcal{M}(R)$, the space of signed measures on R. $\mathcal{M}(R)$ is a Banach space with the variation norm $\|\cdot\|$ [6,p.35]. Let $\mathcal{M}_1(R) = \{\mu \in \mathcal{M}(R): \mu \geq 0, \|\mu\| = 1\}$ represent the space of probability measures on R.

Integrating both sides of (2.5) with respect to $\varphi \in \mathcal{M}_1(R)$, we obtain for $f \in C_o^\infty(R)$

$$\varphi(U^u(t)f) = \varphi(f) + \int_0^t \varphi(U^u(s)(A(s) + u(s,\cdot)D)f)ds, \tag{2.7}$$

where $U^u(t) \equiv U^u(0,t)$, $t \in I$. Let $\eta^u(t)\varphi(\Gamma) \equiv \int_R P^u(0,x,t,\Gamma)\varphi(dx)$ for $\Gamma \in \mathcal{B}(R)$. Then, for any $f \in B(R)$, $\varphi(U^u(t)f) = \eta^u(t)\varphi(f)$. Since (2.7) holds for all $f \in C_o^\infty(R)$, it uniquely defines the flow of probability measures induced by the stochastic process $\{x_t^u : t \in I\}$. For $t \in I$, we write (2.7) as

$$\eta^u(t)\varphi(f) = \varphi(f) + \int_0^t \varphi(U^u(s)(A(s) + u(s,\cdot)D)f\,ds$$

or, abstractly, as

$$\eta^u(t)\varphi = \varphi + \int_0^t \varphi(U^u(s)(A(s) + u(s,\cdot)D)(\ \ ))\,ds, \tag{2.8}$$

and call it as the 'dynamic equation in the space of measures' associated with (2.4).

Chapter III is concerned with the attainable sets of probability measures for the stochastic control system (2.4). In Section 3.1, some simple time-varying control systems are briefly investigated to motivate the subsequent work. In Section 3.3, the following result is established: Given a solution process defined by (2.4) for a bounded measurable control, then this process can be approximated, as closely as desired, by the solution process of (2.4) for a bang-bang control. Then it will be shown that the attainable set of probability measures of (2.4), where the control functions are in a bounded set of bounded measurable functions, is compact in the weak topology on $\mathcal{M}_1(R)$ for any $t \geq 0$. A convexity result is also obtained.

Chapter IV considers some optimal stochastic control problems for the system (2.4). In Section 4.1, a definition of the stochastic control problem is presented, and the concept of controllability in the space of probability measures is discussed. In Section 4.2 it is shown that the attainable sets of probability measures of (2.4) are continuous in time with respect to a topology derived from the weak topology on $\mathcal{M}_1(R)$. With the help of this result, an existence theorem for time optimal stochastic controls is proved. In Section 4.3 it is shown that, for a very general class of problems, there exists a control which minimizes the expected value of a cost functional at any time $t_1 > 0$.

In Chapter V, we establish the existence of unique quasi-diffusions for the class of systems where the diffusion coefficients are bounded and integrable. With the aid of this result, a necessary and sufficient condition for the average of a cost functional to be minimized by a feedback control is derived.

The idea of studying stochastic control systems from the point of view of the state distribution functions is not new to the control literature. Mortensen [50] used this approach to obtain a 'Hamilton-Jacobi' type equation in function space.* In this dissertation, we shall not be concerned with the optimal stochastic control problem as formulated in [50]. Our <u>main goal</u> is to study the <u>attainable set of probability</u>

---

* The drift coefficient in [50] is restricted to be continuous.

measures of a stochastic control system : specifically, as stated above, we shall be interested in the continuity, convexity, and compactness properties of the attainable sets. Although this information does not appear to facilitate the solution of any practical problems, it does offer a relatively simple and elegant way of regarding stochastic control systems. In conclusion, we aver that the understanding and characterization of the attainable sets of probability measures is fundamental in the study of stochastic control processes.

Chapter VI deals with the straight forward application of the theory of (deterministic) dynamical systems to systems of a stochastic nature. In Section 6.1, two examples are presented which are used to motivate the definition of a non-deterministic dynamical system in Section 6.2. In Section 6.3, the concept of limit sets is employed to obtain some topological results concerning the trajectories of probability measures. In Section 6.4, the limit sets are characterized further, and it is shown how the averages of certain functions, as $t \to \infty$, are related to the limit sets.

Chapter VII is concerned with the stability of non-deterministic dynamical systems. In Section 7.1, a certain continuous function (not a Lyapunov function) is introduced whose existence ensures the stability of flows in the space of probability measures. In Section 7.2, the results of the preceding section are applied to the stochastic stability theory.

The work in this part of the dissertation is motivated by [35]. The main idea here, which was not recognized in [35], is the applicability of dynamical system theory to stochastic problems. To give some indication of the effectiveness of these methods, we mention that the major theorem of [35], Theorem 3, is an immediate consequence of a standard result in dynamical system theory (see Remark (ii) at the end of Section 6.3).

We have not yet attempted to apply the theory of Chapters VI and VII to practical problems, where the usual formulation may be inapplicable. However, the theory has provided some new results related to stochastic stability. For example, the results of Section 6.4 , as well as the definition of a generalized stochastic Lyapunov (D-function) and the work of Section 7.1 , are completely new to the stochastic theory.

As a final remark, we state, that although the techniques of dynamical system theory are incapable of establishing a.s. stability results, they do provide a unified and functional analytic approach to a large class of stochastic stability problems (see Section 7.1).

CHAPTER III

ATTAINABLE SETS OF PROBABILITY MEASURES

3.1 Simple Stochastic Control Systems

To motivate the work that follows, we consider simple control systems of the form

$$d\,x_t = (a(t) + u_1(t))\,dt + (b(t) + u_2(t))\,d z_t\ , \tag{3.1}$$

where a, b, u_1, u_2 are one-dimensional bounded measurable functions of t, and z_t is one-dimensional Brownian motion. a and b are system parameters and are assumed to be fixed. For each t, the random variable $x(t,\omega)$ induced by the stochastic differential equation (3.1) has a normal distribution with the mean

$$e_t = \int_0^t (a(s) + u_1(s))\,ds\ ,$$

and the variance

$$\sigma_t^2 = \int_0^t (b(s) + u_2(s))^2\,ds\ .\ ^*$$

Suppose that the Markov process defined by (3.1) starts at time $t=0$, and $x=0$ a.s. W, and that at some t_1 afterwards, we wish $x(t_1,\omega)$ to be equal to $\alpha \in R$ a.s. W. As stated, this problem demands a great deal; a more reasonable problem is to have the control functions u_1 and u_2 'direct' the distribution functions of the family of random variables $\{x(t,\omega) : t \in [0, t_1]\}$ in such a manner that, at time t_1, it has a preassigned fixed normal distribution, $N(\alpha, \sigma^2)$, with mean α and variance σ^2. To determine the controls necessary to accomplish this matching of distribution functions

---

* $x_o = 0$ a.s. W .

(recalling that a normal distribution is completely specified by its mean and variance), the following conditions are necessary and sufficient :

$$\int_0^{t_1} (a(t) + u_1(t))\,dt = \alpha \quad ,$$

and

$$\int_0^{t_1} (b(t) + u_2(t))^2\,dt = \sigma^2 \quad . \tag{3.2}$$

The controls

$$u_1(t) = \frac{\alpha}{t_1} - a(t) \quad \text{and} \quad u_2(t) = \frac{\sigma}{\sqrt{t_1}} - b(t) \tag{3.3}$$

are bounded measurable functions which satisfy (3.2), although they are by no means unique. Therefore, u_1 and u_2, given by (3.3), 'direct' the distribution functions of $x(t, \omega)$ to $N(\alpha, \sigma^2)$.

Note that the system (3.1) can only be controlled to a normal distribution and that controls are necessary in each of the coefficients, one controlling the mean and the other the variance.

The normal distribution of x_t plays a crucial role in the above analysis. If state variables are introduced in the coefficients of (3.1), the distribution function of x_t is no longer necessarily normal and what was a two-dimensional problem now becomes an infinite dimensional problem, i.e., above we had to match only the mean and variance of the random variable x_{t_1}, whereas now, if we wish to match distribution functions, it is necessary to match all the moments of the distribution function induced by x_{t_1} with those of the 'target' distribution function.

Another important problem in control theory is the characterization of the control function necessary to attain a certain objective. In time-optimal deterministic control theory for linear systems, it is a well-known fact that the class of 'bang-bang' controls is as effective as the larger class of bounded measurable controls. To motivate the study of analogous problems for stochastic systems, we consider

$$d\, x_t^u \;=\; (a(t) + u(t))\, dt \;+\; b(t)\, dz_t \quad , \tag{3.4}$$

where we assume the process starts at $t = 0$, $x = 0$ a.s. W , a and b are as above, u is bounded measurable, and $0 \le u(t) \le 1$, $t \in I = [0, t_f]$. At any time $\tau \in I$, the mean of the distribution function induced by x_τ^u is

$$\int_0^\tau (a(s) + u(s))\, ds \quad ,$$

and the variance is

$$\int_0^\tau b^2(\tau)\, ds \quad ,$$

which is independent of the control. We can readily find a control $\bar{u}$, $\bar{u}(s) = +1$ or 0 for all $s \in I$, and such that

$$\int_0^\tau u(s)\, ds = \int_0^\tau \bar{u}(s)\, ds \quad .$$

Therefore, at any time $\tau > 0$, we can find a 'bang-bang' control such that the distribution function induced by x_τ^u is the same as that induced by $x_\tau^{\bar{u}}$. As a consequence of this, it follows that the expected value of $f(x_\tau^u(\omega))$, conditioned to start at $x = 0$ and

$t = 0$, is the same as that of $f(x^{\bar{u}}_{\tau}(\omega))$ starting at $x = 0$, $t = 0$, for any real bounded continuous functions f.

3.2 Classes of Control Functions

The notion of attainable set is fundamental in the theory of control. We are thus motivated to investigate the attainable set of probability measures associated with the stochastic control system (2.4).

Let $I = [0, t_f]$, $t_f < +\infty$, and let $\mathcal{L}_{\infty}(I)$ be the Banach space of bounded measurable functions on I with the supremum norm. Define

$$\Sigma = \left\{ u \in \mathcal{L}_{\infty}(I) \; : \; 0 \le u(t) \le 1 \quad \text{for all } t \in I \right\}$$

to be the set of admissible control functions on I. For each $u \in \Sigma$, (2.4) defines a Markov process.

In Chapter II we defined a continuous Markov process on $[0, \infty]$ into R to be a family of probability measures $P_{s,x}$, $s \ge 0$ and $x \in R$, on $\Omega = C\big([0, \infty), R\big)$ such that $P_{s,x}(\Gamma \mid \mathfrak{F}^s_t) = P_{t, \pi_t}(\Gamma)$ a.s. $P_{s,x}$ for $\Gamma \in \mathfrak{F}^t$. The transition distribution function is defined by $P(s, x, t, B) = P_{s,x}(\pi_t \in B)$, $B \in \mathcal{B}(R)$. If $P^u(s, x, t, B)$ is the transition function of the Markov process (2.4) corresponding to $u \in \Sigma$, then $\eta^u(t)\,\varphi(\cdot) = \int_R P^u(0, x, t, \cdot)\,\varphi(dx)$ describes the flow of probability measures on I associated with (2.4).

Definition 3.1. We define the attainable set of probability measures for the control system (2.4), for the admissible class Σ, by

$$\mathcal{R}^t_\varphi = \left\{ \eta^u(t)\,\varphi : u \in \Sigma \right\}, \qquad \varphi \in \mathcal{M}_1(R) .$$

For each $f \in C(R)$, the space of bounded continuous functions on R with the supremum norm, we define

$$\mathcal{R}^t_\varphi(f) = \left\{ \eta^u(t)\,\varphi(f) : u \in \Sigma \right\} .$$

Let χ_H be the characteristic function of the set H.

Define

$$\Sigma^o = \left\{ \chi_H : H \text{ is a measurable subset of } I \right\} .$$

As in Definition 3.1, we define

$$\overset{o}{\mathcal{R}}{}^t_\varphi = \left\{ \eta^u(t)\,\varphi : u \in \Sigma^o \right\}, \qquad \varphi \in \mathcal{M}_1(R),$$

and

$$\overset{o}{\mathcal{R}}{}^t_\varphi(f) = \left\{ \eta^u(t)\,\varphi(f) : u \in \Sigma^o \right\} .$$

for $f \in C(R)$.

We now consider $\mathcal{L}_\infty(I)$ with its $\sigma(\mathcal{L}_\infty(I), \mathcal{L}_1(I))$ topology, where $\mathcal{L}_1(I)$ is the space of functions whose absolute values are integrable over I with respect to the Lebesgue measure.

A sequence $\{u_n\} \subset \mathcal{L}_\infty(I)$ converges to $u \in \mathcal{L}_\infty(I)$ in the $\sigma(\mathcal{L}_\infty(I), \mathcal{L}_1(I))$ topology if and only if

$$\int_I u_n(s)\,g(s)\,ds \rightarrow \int_I u(s)\,g(s)\,ds \qquad \text{as } n \rightarrow \infty$$

for all $g \in \mathcal{L}_1(I)$. We now present a result which shall be useful in the sequel.

Lemma 3.1

$$\Sigma^{o} \text{ is } \sigma(\mathcal{L}_{\infty}(I), \mathcal{L}_{1}(I)) \text{ dense in } \Sigma .$$

Proof : Let l be the Lebesgue measure on I. We wish to show that, given any $u \in \Sigma$, there exists a sequence $\{u_n\} \subset \Sigma^{o}$ such that

$$\int_I u_n(s)\, g(s)\, l(ds) \rightarrow \int_I u(s)\, g(s)\, l(ds) \qquad \text{as } n \rightarrow \infty$$

for all $g \in \mathcal{L}_1(I)$. This is equivalent to showing that for all $B \in \tilde{\mathcal{B}}(I) \equiv \{\chi_{[0,t]} : t \in I\}$,

$$\int_B u_n(s)\, l(ds) \rightarrow \int_B u(s)\, l(ds) \qquad \text{as } n \rightarrow \infty .$$

This result can be found in [1, p. 342, Example 27]. Since u_n is the characteristic function of a measurable set $A_n \subset I$, we would like to prove the existence of a sequence of measurable sets $\{A_n\}$ such that for all sets $B \in \tilde{\mathcal{B}}(I)$

$$l(A_n \cap B) \rightarrow \int_B u(s)\, l(ds) \qquad \text{as } n \rightarrow \infty . \qquad (3.5)$$

If $u \in \Sigma$ is continuous, then by [4, p. 300, Example 3] a sequence $\{A_n\}$ can be constructed for which (3.5) is satisfied.

Now let $u \in \Sigma$ be arbitrary and fixed. Since $\mathcal{L}_2(I) \supset \mathcal{L}_{\infty}(I)$, and $C(I)$, the space of real-valued continuous functions on I, is dense in $\mathcal{L}_2(I)$ with respect to the $\mathcal{L}_2$-norm, there exists a sequence $\{v_i\} \subset C(I)$ such that

$$\int_I (v_i(s) - u(s))^2\, l(ds) \rightarrow 0 \qquad \text{as } i \rightarrow \infty .$$

This implies the existence of a subsequence, also labelled $\{v_i\}$, such that v_i converges to u a.e. on I with respect to l. If necessary, each v_i can be redefined so that $0 \leq v_i(t) \leq 1$ for all $t \in I$ and v_i remains in $C(I)$.

By Renyi's example, for each $v_p \in C(I)$, there exists a sequence of measurable sets $\left\{ A_{p,k} \right\}_{k=1}^{\infty}$ such that

$$\lim_{k \to \infty} \left| \, l(B \cap A_{p,k}) - \int_B v_p(s) \; l(ds) \right| = 0 \tag{3.6}$$

It is also proved in [4] that it suffices to prove (3.5) for $B = A_i$, $i = 1, 2, \ldots$. Let $\tilde{\tilde{\mathcal{B}}}(I) = \left\{ B_i \right\}_{i=1}^{\infty}$ be all the sets of $\left\{ A_{p,k} \right\}$ in some order. We wish to choose a subsequence $\left\{ A_n \right\}$ from $\tilde{\tilde{\mathcal{B}}}(I)$ such that (3.5) is satisfied for all $B \in \tilde{\tilde{\mathcal{B}}}(I)$. Then, by (3.6), we can choose the sequence of integers $\left\{ k_n \right\}$ as follows:

choose $k_1 \ni \left| \, l(B_1 \cap A_{1,k}) - \int_{B_1} v_1(s) \; l(ds) \right| < 1$ for $k \geq k_1$,

choose $k_2 > k_1 \ni \left| \, l(B_i \cap A_{2,k}) - \int_{B_i} v_2(s) \; l(ds) \right| < \frac{1}{2}$ for $k \geq k_2$, $i = 1, 2$,

....

choose $k_n > k_{n-1} \ni \left| \, l(B_i \cap A_{n,k}) - \int_{B_i} v_n(s) \; l(ds) \right| < \frac{1}{n}$ for $k \geq k_n$, $i = 1, \ldots n$.

Let $A_n = A_{n, k_n}$. Thus,

$$\left| \int_I \chi_{B_i} (\chi_{A_n} - v_n(x)) \; l(dx) \right| < \frac{1}{n}, \quad i = 1, 2, \ldots n, \tag{3.7}$$

for $n \geq 1$. We claim that $\{A_n\}$ satisfies (3.5). Let $B \in \tilde{\mathcal{B}}(I)$ be arbitrary. Then,

$$| \int_I \chi_B \chi_{A_n} \, l(dx) - \int_B u(x) \, l(dx) | \leq | \int_I \chi_B (\chi_{A_n} - v_n(x)) \, l(dx) | \tag{3.8}$$
$$+ | \int_I \chi_B (v_n(x) - u(x)) \, l(dx) | .$$

The first term on the right-hand side of (3.8) goes to 0 as $n \to \infty$ by (3.7), and the second term approaches 0 by the a.e. convergence of $\{v_n\}$ to u. The Dominated Convergence Theorem permits the limiting operation in the second term. Thus, for any $B \in \tilde{\mathcal{B}}(I)$,

$$l(A_n \cap B) \longrightarrow \int_B u(x) \, l(dx)$$

as $n \to \infty$.

Q.E.D.

Remarks :

(i) Lemma 3.1 can be proved in a different manner by using the (deterministic) bang-bang principle [9, p.23].

Lemma 3.2

Given $u \in \Sigma$, there exists $\{u_n\} \subset \Sigma^o$ such that

$$| \int_0^t (u(s) - u_n(s)) \, ds | \longrightarrow 0$$

as $n \to \infty$ for all $t \in I$.

Proof: Subdivide $I = [0, t_f]$ into n equal subintervals each of length Δ. Then, by the bang-bang principle [9, p.23], there exists $\bar{u}_j$ for $1 \le j \le n$ such that $\bar{u}_j(t) = +1$ or 0 for all $t \in I$ and so that

$$| \int_{(j-1)\Delta}^{j\Delta} (u(s) - u_j(s))\, ds \,| = 0 \quad .$$

Define u_n on I by $u_n(t) = u_j(t)$ for $(j-1)\Delta < t \le j\Delta$, $1 \le j \le n$. Then, for any $t \in I$, we have $(j-1)\Delta < t \le j\Delta$ for some j, and hence the following relation:

$$| \int_0^t (u(s) - u_n(s))\, ds \,| = | \int_0^{(j-1)\Delta} (u(s) - u_n(s))\, ds + \int_{(j-1)\Delta}^{t} (u(s) - u_n(s))\, ds \,| \le 2\Delta .$$

Letting $\Delta \to 0$, we obtain the desired result.

Q.E.D.

(ii) We note that if $\{u_n\} \subset \Sigma^o$ converges to $u \in \Sigma$, $u \notin \Sigma^o$, in the $\sigma(\mathcal{L}_\infty(I), \mathcal{L}_1(I))$ topology, there exists no subsequence of $\{u_n\}$ which converges to u in measure (Lebesgue measure on I). To prove this, suppose that there exists a subsequence $\{u_{n_k}\}$ such that $u_{n_k} \to u$ in Lebesgue measure l on I, i.e., given any $\epsilon > 0$, $l(x : |u_{n_k}(x) - u(x)| > \epsilon) \to 0$ as $n_k \to \infty$, then there exists a further subsequence $\{u_{n'_k}\}$ which converges a.e., on I, to u. But this is impossible since $\{u_n\} \subset \Sigma^o$ and $u \in \Sigma$, $u \notin \Sigma^o$.

(iii) In the proof of Lemma 3.1 it is sufficient to work with sequences since $\mathcal{L}_1(I)$ is a separable Banach space, implying that Σ with its $\sigma(\mathcal{L}_\infty(I), \mathcal{L}_1(I))$ topology is a metric topology [1, Theorem V. 5.1, p. 426].

3.3 Stochastic Bang-Bang Principle

It was shown in Section 3.1 that, for the simple control system (3.4), given any $u \in$ and $t_1 \in I = [0, t_f]$ there exists a $\bar{u}$, $\bar{u}(t) = +1$ or 0 for $0 \le t \le t_1$ such that the normal distributions attained at t_1 are the same for u and $\bar{u}$. We shall prove a related result for the stochastic control system (2.4).

Let the coefficients a and b in (2.4) be uniformly Lipschitz continuous in x, and bounded continuous functions of t in I for each $x \in R$. Let u be an arbitrary fixed function in Σ, and let $P_x \equiv P_{0,x}$ be the probability measure on $(\Omega, \mathfrak{F}^0)$ such that $\pi_0 = x$ a.s. P_x and

$$\pi_t = x + \int_0^t (a(s, \pi_s) + u(s))\, ds + \int_0^t b(s, \pi_s)\, dz_s \quad \text{a.s. } P_x .$$

By virtue of Ito's theory [8, Vol. I, Theorem 11.3], such a unique P_x does exist. This is equivalent to saying that there exists a stochastic process x_t such that $x_0 = x$ a.s. with respect to the Wiener measure W on $\Omega = C(I, R)$, the space of continuous functions from I into R, and such that

$$x_t = x + \int_0^t (a(s, x_s) + u(s))\, ds + \int_0^t b(s, x_s)\, dz_s \quad \text{a. s. } W . \tag{3.9}$$

Proposition 3.3

Let $u \in \Sigma$, and let $\{u_n\} \subset \Sigma^0$ converge in the $\sigma(\mathfrak{L}_\infty(I), \mathfrak{L}_1(I))$ topology to u. Let x_t^n and x_t be the stochastic processes defined by (2.4) corresponding to u_n and u, respectively, where a and b are uniformly Lipschitz continuous in x, and for each $x \in R$ are bounded continuous functions of $t \in I$, with common bound k. Then x_t^n converges to x_t in probability, uniformly in $t \in I$, as $n \to \infty$.

Proof : Lemma 3.1 implies that given any $u \in \Sigma$, a sequence $\{u_n\} \subset \Sigma^o$ exists which converges to u in the $\sigma(\mathcal{L}_\infty(I), \mathcal{L}_1(I))$ topology, or, equivalently,

$$v_n(t) \equiv \int_0^t u_n(s)\,ds \rightarrow \int_0^t u(s)\,ds \equiv v(t) \qquad \text{as } n \rightarrow \infty \tag{3.10}$$

for each $t \in I$ [1, p. 342, Example 27].

By (3.9)

$$\begin{aligned} x_t - x_t^n = v(t) - v_n(t) &+ \int_0^t (a(s, x_s) - a(s, x_s^n))\,ds \\ &+ \int_0^t (b(s, x_s) - b(s, x_s^n))\,dz_s \quad . \end{aligned} \tag{3.11}$$

Squaring (3.11), integrating over Ω with respect to the Wiener measure, and using the properties of stochastic integrals, we arrive at

$$\begin{aligned} E^W\left\{x_t - x_t^n\right\}^2 \le \left\{v(t) - v_n(t)\right\}^2 &+ K^2(t_f + 1) \int_0^t E^W\left\{x_s - x_s^n\right\}^2 ds \\ &+ 2\,t_f\,k\,|\,v(t) - v_n(t)\,| \quad , \end{aligned} \tag{3.12}$$

where

$$|a(s,x) - a(s,y)| \le K|x - y| \quad , \quad |b(s,x) - b(s,y)| \le K\,|\,x - y\,|$$

for all $s \in I$. Now, given any $\epsilon > 0$, we can find $N_\epsilon \ge 0$ such that for $n \ge N_\epsilon$, the sum of the first and third terms on the right-hand side of (3.12) is less than ϵ. This follows from (3.10). Thus,

$$E^W\left\{x_t - x_t^n\right\}^2 \le \epsilon + K^2(t_f + 1) \int_0^t E^W\left\{x_s - x_s^n\right\}^2 ds \ ,$$

which implies, by Gronwall's Lemma [11, p.11], that

$$E^W\left\{x_t - x_t^n\right\}^2 \le \epsilon \ \exp\left\{K^2 t_f \ (1 + t_f)\right\} \ . \tag{3.13}$$

Hence,

$$\lim_{n \to \infty} E^W \left\{ x_t - x_t^n \right\}^2 = 0 \quad . \tag{3.14}$$

Moreover, we have :

$$W\left\{ \sup_{t \in I} | x_t - x_t^n | > \epsilon \right\} \leq W\left\{ \sup_{t \in I} \left| \int_0^t (a(s, x_s) - a(s, x_s^n))\, ds \right| > \frac{\epsilon}{3} \right\}$$
$$+ W\left\{ \sup_{t \in I} \left| \int_0^t (b(s, x_s) - b(s, x_s^n))\, dz_s \right| > \frac{\epsilon}{3} \right\}$$
$$+ W\left\{ \sup_{t \in I} | v(t) - v_n(t) | > \frac{\epsilon}{3} \right\} \quad . \tag{3.15}$$

Since

$$W\left\{ \sup_{t \in I} \left| \int_0^t (a(s, x_s) - a(s, x_s^n))\, ds \right| > \frac{\epsilon}{3} \right\} \leq W\left\{ K \int_0^{t_f} | x_s - x_s^n |\, ds > \frac{\epsilon}{3} \right\}$$
$$\leq \frac{9}{\epsilon^2} K^2 E^W \left\{ \int_0^{t_f} (x_s - x_s^n)^2\, ds \right\}$$

by Chebyshev's inequality, the first term on the right-hand side of (3.15) goes to 0 , as $n \to \infty$, by (3.13) and (3.14). The second term approaches 0, as $n \to \infty$, by the martingale inequality, (3.13), and (3.14), while the third term goes to 0, as $n \to \infty$, by the choice of $\{u_n\}$. (The formula (3.13) permits the use of the Dominated Convergence Theorem .)

Q.E.D.

<u>Corollary 3.4</u>.

Let f be a bounded continuous function on R, then $f(x_t^n)$ converges in probability to $f(x_t)$ as $n \to \infty$ for each $t \in I$, and $E^W f(x_t^n) \to E^W f(x_t)$ as $n \to \infty$.

Proof : The first part is a standard result of probability theory, and the second follows from the Dominated Convergence Theorem.

Q.E.D.

Corollary 3.5

If f is uniformly Lipschitz continuous with constant L, then $E^W\{f(x_t^n) - f(x_t)\}^2 \to 0$ as $n \to \infty$ for each $t \in I$.

Proof : $$E^W\{f(x_t) - f(x_t^n)\}^2 \leq L^2 E^W\{x_t - x_t^n\}^2 \to 0 \quad \text{as } n \to \infty$$

by (3.14).

Q.E.D.

Corollary 3.4 implies that for any $t \in I$ and any bounded $f \in C(I, R)$ $E^W f(x_t)$ can be approximated as closely as desired by $E^W f(x_t^n)$ where x_t^n is the solution process associated with (2.4) for the 'bang-bang' control u_n. If we assume that the coefficient b is strictly elliptic, then we can obtain a stronger and more interesting result. The following is an extension of [5, Part II, Theorem 9.1] since the drift coefficient contains a discontinuous term and the convergence of u_n to u is not uniform. The proof depends on techniques developed in [5].

Theorem 3.6 (Stochastic Bang-Bang Principle)*

Assume the coefficients a and b of (2.4) satisfy the following conditions : a and b are bounded continuous functions, and $b^2(t, x) \geq \nu > 0$. Let $x \in R$ and $P_x^n \equiv P_{0,x}^n$, $P_x \equiv P_{0,x}$, be the probability measures on $(\Omega, \mathfrak{F}^o)$** defined by (2.4)

* See Remark (iv) at the end of this section.

** $\Omega = C(I, R)$ and $\mathfrak{F}_t^s$ is the smallest σ-algebra on Ω with respect to which all the coordinate functions $\pi_\tau : \Omega \to R$, $s \leq \tau \leq t \leq t_f$ are measurable. Define $\mathfrak{F}^o \equiv \mathfrak{F}_{t_f}^o$.

for the respective controls u_n and u, where $\{u_n\} \subset \Sigma^o$ converges to $u \in \Sigma$ in the $\sigma(\mathcal{L}_\infty(I), \mathcal{L}_1(I))$ topology. Then, P_x^n converges weakly to P_x.

<u>Proof</u> : Let

$$y_\theta(t) = \exp\left\{\theta(\pi_t - x) - \theta\int_0^t a(s, \pi_s)\,ds - \theta\int_0^t u(s)\,ds - \frac{\theta^2}{2}\int_0^t b^2(s, \pi_s)\,ds\right\}$$

and

$$y_\theta^n(t) = \exp\left\{\theta(\pi_t - x) - \theta\int_0^t a(s, \pi_s)\,ds - \theta\int_0^t u_n(s)\,ds - \frac{\theta^2}{2}\int_0^t b^2(s, \pi_s)\,ds\right\} \tag{3.15}$$

be the martingales associated with u and u_n, respectively.

Using the fact that $y_\theta^n(t)$ is a martingale with respect to P_x^n, we can derive the following estimate, as in [5, Part I, Lemma 3.2]:

$$E^{P_x^n}\{\pi_t - \pi_s\}^4 \le C\,|t-s|^2 \quad , \; t \ge s, \; t \in I$$

for all $n \ge 1$, where C depends on the bounds of a, b, and u. Then, by [15, Theorem 2, p.33], the family $\{P_x^n\}_{n=1}^\infty$ is relatively weakly compact, i.e. there exists a measure Q on $(\Omega, \mathcal{F}^o)$ and a convergent subsequence of $\{P_x^n\}_{n=1}^\infty$, also denoted by $\{P_x^n\}_{n=1}^\infty$, which converges weakly to Q.

Since $y_\theta^n(t)$ is a martingale with respect to P_x^n, we have for $\Gamma \in \mathcal{F}_s^o$ and $t > s \ge 0$, $t \in I$,

$$\int_\Gamma y_\theta^n(t, \omega)\, P_x^n(d\omega) = \int_\Gamma y_\theta^n(s, \omega)\, P_x^n(d\omega)$$

or, for any $\mathcal{F}_s^o$ measurable bounded continuous function $g(\omega)$,

$$\int_\Omega g(\omega)\, y_\theta^n(t, \omega)\, P_x^n(d\omega) = \int_\Omega g(\omega)\, y_\theta^n(s, \omega)\, P_x^n(d\omega) \; . \tag{3.16}$$

We must show that $y_\theta(t)$ is a Q - martingale, i.e.,

$$\int_\Omega g(\omega)\, y_\theta(t,\omega)\, Q(d\omega) = \int_\Omega g(\omega)\, y_\theta(s,\omega)\, Q(d\omega) \quad . \tag{3.17}$$

Let M be a number greater than 0. We claim that the family $\Psi_\omega \equiv \{ g(\omega)\,(y_\theta^n(t,\omega) \wedge M) : n \geq 1 \}$ is uniformly bounded, $\mathfrak{F}_t^o$ measurable, and equicontinuous at each $\omega \in \Omega$, where $\gamma \wedge \lambda = \inf(\gamma, \lambda)$. It is only necessary to show that for fixed t and θ, $\{ y_\theta^n(t,\omega) : n \geq 1 \}$ is equicontinuous at each $\omega \in \Omega$. Fix $\omega \in \Omega$, $t \in I$, and let

$$\zeta(\omega) \equiv \exp\left\{ \theta(\pi_t(\omega) - x) - \theta \int_0^t a(s, \pi_s(\omega))\, ds - \frac{\theta^2}{2} \int_0^t b^2(s, \pi_s(\omega))\, ds \right\} .$$

Then,

$$| y_\theta^n(t,\tilde{\omega}) - y_\theta^n(t,\omega) | \leq e^{\theta t} | \zeta(\tilde{\omega}) - \zeta(\omega) | ,$$

which is independent of n. Since π_t is a continuous function of Ω into R, where Ω has the sup norm topology, and a, b are continuous functions of the space variable, the set $\{ y_\theta^n(t,\omega) : n \geq 1 \}$ is equicontinuous at each $\omega \in \Omega$. Since g is continuous, Ψ_ω is also equicontinuous, i.e., given any $\epsilon > 0$ there exists a δ - ball (in the sup norm topology), $N_\delta(\omega)$, of ω such that for all $\tilde{\omega} \in N_\delta(\omega)$,

$$\sup_{n \geq 1} | g(\tilde{\omega})(y_\theta^n(t,\tilde{\omega}) \wedge M) - g(\omega)(y_\theta^n(t,\omega) \wedge M) | < \epsilon \quad .$$

Thus, since $y_\theta^n(t,\omega) \to y_\theta(t,\omega)$ for each ω as $n \to \infty$, [13, Theorem 6.8, p. 51] implies that

$$\int_\Omega g(\omega)\,(y_\theta(t,\omega) \wedge M)\, Q(d\omega) = \lim_{n \to \infty} \int_\Omega g(\omega)\,(y_\theta^n(t,\omega) \wedge M)\, P_x^n(d\omega) \tag{3.18}$$

for each $M > 0$.

We now show that

$$E^{Q}\left\{ | y_{\theta}(t) - y_{\theta}(t) \wedge M | \right\},$$

and

$$\sup_{n \geqslant 1} E^{P_x^n}\left\{ | y_{\theta}^{n}(t) - y_{\theta}^{n}(t) \wedge M | \right\}$$

tend to 0 as $N \to \infty$: Using [5, Part I, Lemma 3.1] we have that for all $n \geq 1$,

$$P_x^n\left\{ \sup_{t \in I} | \pi_t - x | > l \right\} \leq d \exp\left\{ - v l^2 \right\},$$

where d and v are positive constants, implying that

$$\sup_{n \geqslant 1} E^{P_x^n}\left\{ | y_{\theta}^{n}(t) - y_{\theta}^{n}(t) \wedge M | \right\} \to 0 \tag{3.19}$$

as $M \to \infty$. Now, the set $\left\{ \omega : | \pi_t(\omega) - x | > l \right\}$ is open and $\mathfrak{F}_t^o$ measurable, and therefore, by virtue of [13, Theorem 6.1, p. 40],

$$\begin{aligned} Q\left\{ | \pi_t - x | > l \right\} &\leq \varliminf_{n \to \infty} P_x^n\left\{ | \pi_t - x | > l \right\} \\ &\leq d \exp\left\{ - v l^2 \right\} . \end{aligned}$$

Thus, for any $t \in I$,

$$E^{Q}\left\{ | y_{\theta}(t) - y_{\theta}(t) \wedge M | \right\} \to 0 \tag{3.20}$$

as $M \to \infty$.

Now,

$$E^{Q}\left\{ g\, y_{\theta}(t) \right\} = E^{Q}\left\{ g\,(y_{\theta}(t) - y_{\theta}(t) \wedge M) \right\} + E^{Q}\left\{ g\,(y_{\theta}(t) \wedge M) \right\},$$

and

$$\lim_{n\to\infty} E^{P_x^n}\left\{g\, y_\theta^n(t)\right\} = \lim_{n\to\infty} E^{P_x^n}\left\{g\,(y_\theta^n(t) - y_\theta^n(t)\wedge M)\right\} + \lim_{n\to\infty} E^{P_x^n}\left\{g\,(y_\theta^n(t)\wedge M)\right\}.$$

Therefore, (3.18), (3.19) and (3.20) imply that

$$E^Q\left\{g\, y_\theta(t)\right\} = \lim_{n\to\infty} E^{P_x^n}\left\{g\, y_\theta^n(t)\right\} .$$

Taking the limit on both sides of (3.16), we obtain (3.17), i.e., $y_\theta(t)$ is a Q-martingale. But $y_\theta(t)$ is given as a martingale with respect to P_x, and the solution of the martingale problem is unique [5, Part I, Theorem 5.6]. Therefore, $Q = P_x$ and $P_x^n \Longrightarrow P_x$ as $n \to \infty$.

Q.E.D.

If the drift coefficient a in (2.4) is bounded and measurable, then the proof of Theorem 3.6 becomes more difficult. However, in [5, Part II, Theorem 9.2], it is shown that if u^n converges to u in measure (Lebesgue measure on R) on compact sets of R, then $P_x^n \Longrightarrow P_x$ as $n \to \infty$. In fact, convergence of u_n to u as $n \to \infty$, in the sense that

$$\int_I u_n(s)\, g(s)\, ds \longrightarrow \int_I u(s)\, g(s)\, ds \qquad \text{as} \quad n \to \infty, \tag{3.21}$$

for all g bounded continuous on I, is sufficient. This is remarked, but not proven, in [5, Part II, p. 500]. Therefore, since the convergence of (3.21) is weaker than $\sigma(\mathcal{L}_\infty(I), \mathcal{L}_1(I))$ convergence, we have :

Corollary 3.7

Let a be bounded measurable, b bounded continuous and strictly elliptic. Then given $u \in \Sigma$ there exists $\{u_n\} \subset \Sigma^o$ such that $u_n \to u$ as $n \to \infty$ in the $\sigma(\mathcal{L}_\infty(I), \mathcal{L}_1(I))$ topology, and $P_x^n \Longrightarrow P_x$.

Remarks :

(i) Theorem 3.6 and Corollary 3.7 imply that for each $f \in C(R)$ and $t \in I$,

$$\int_\Omega f(\pi_t(\omega))\, P_x^n(d\omega) \rightarrow \int_\Omega f(\pi_t(\omega))\, P_x(d\omega)$$

as $n \rightarrow \infty$ or, equivalently,

$$\int_R f(y)\, P^n(0, x, t, dy) \rightarrow \int_R f(y)\, P(0, x, t, dy) \tag{3.22}$$

as $n \rightarrow \infty$. On integrating both sides of (3.22) with respect to the probability measure φ, we obtain

$$\eta^{u_n}(t)\,\varphi \Longrightarrow \eta^{u}(t)\,\varphi \tag{3.23}$$

as $n \rightarrow \infty$. This means that the set $\overset{o}{\mathcal{R}}{}_\varphi^t$ is dense in $\mathcal{R}_\varphi^t$ in the weak topology on $\mathcal{M}_1(R)$ for each $t \in I$.

(ii) All the results of this section can be easily extended to an n - dimensional analogue of the stochastic differential equation (2.4).

(iii) If we set $b \equiv 0$ in Proposition 3.3, we have that

$$\lim_{n \rightarrow \infty} x^n(t) = x(t)$$

for each $t \in I$, where $x(t)$ $(x^n(t))$ is the unique solution of the ordinary differential equation.

$$\dot{x}(t) = a(t, x(t)) + u(t) \qquad (\dot{x}^n(t) = a(t, x^n(t)) + u^n(t)), \tag{3.24}$$

which is a result obtained in [10]. The work in [10] employs the (deterministic) bang-bang principle whereas, in this work, we make no recourse to it.

(iv) Calling Theorem 3.6 a 'bang-bang principle' is not strictly in accord with its meaning in the literature. In the usual sense, Theorem 3.6 would be a 'bang-bang principle' if we could show that for $t \in I$ and $u \in \Sigma$, there exists a $\bar{u} \in \Sigma^{o}$ such that $\eta^{u}(t)\varphi = \eta^{\bar{u}}(t)\varphi$. This result appears to be difficult to prove; even for the deterministic system (3.24), the corresponding result has not yet been demonstrated.

(v) Theorem 3.6 does not generalize the result in [10] (where the drift coefficient is required to be uniformly Lipschitz in the state variable), since the diffusion coefficient in (2.4) must be strictly elliptic.

(vi) Rather than use the control classes Σ and Σ^{o}, we could have employed

$$\Upsilon = \{ u : u \text{ measurable on } I, \; |u(t)| \leq \beta, \; t \in I \},$$

and

$$\Upsilon^{o} = \{ u : u \text{ measurable on } I, \; |u(t)| = \beta, \; t \in I \},$$

where $\beta > 0$.

(vii) Given any $u \in \mathcal{L}_{\infty}(I)$, if there exists a sequence $\{u_n\} \subset \mathcal{L}_{\infty}(I)$ such that $u_n \to u$ as $n \to \infty$ in the $\sigma(\mathcal{L}_{\infty}(I), \mathcal{L}_1(I))$ topology, then the results of Theorems 3.6 and Corollary 3.7 remain valid.

(viii) The author is aware of only one other work in the literature, Fleming [23], which deals with bounded controls for stochastic systems. Using the theory of weak solutions of linear parabolic equations, Fleming shows that the probability density function $q^u(t, y)$ of the random variable x_t^u exists, where $u \in \mathcal{L}_\infty(I)$. Also, if $\underline{u_n \to u \text{ a.e. on } I}$, it can be inferred from the work in [23, Appendix 2] that $q^{u_n}(t, y) \to q^u(t, y)$, as $n \to \infty$, uniformly on compact sets, where $q^{u_n}(t, y)$ is the density function of $x_t^{u_n}$. Theorem 3.6 is a stronger result since it requires only $\sigma(\mathcal{L}_\infty(I), \mathcal{L}_1(I))$ convergence of $\{u_n\}$.

(ix) In [51], Fattoroni considered an "approximating bang-bang principle", similar in idea to Theorem 3.6, for the linear control system

$$\dot{x}(t) = A(t)\, x(t) + B(t)\, u(t), \tag{3.25}$$

where $x(t)$ and $u(t)$ take values in a Banach space E, $A(t)$ is an unbounded linear operator from E into E and $B(t) \in \mathcal{L}(E, E)$. Assuming the existence of a solution for (3.25), and making certain assumptions on $B(t)$ and $u(t)$, Fattorini showed that the attainable set of $x(t)$, for bang-bang controls, is dense (in a certain topology) in the attainable set of $x(t)$, for an appropriate larger class of controls. To show that this result does not detract from Theorem 3.6, we observe that even if the two-parameter semi-group $\{ U^u(s, t) : t \geq s \geq 0 \}$, given by (2.5), satisfies the differential equation

$$\frac{\partial U^u(s,t)f}{\partial t} = (A(t) + u(t, \cdot)\, D)\, U^u(s, t)f, \tag{3.26}$$

(see Corollary 5.5) where $A(t)$ is defined by (2.6), (3.26) cannot be considered as a special case of (3.25) since D is an unbounded linear operator on $C_0(R)$, and D operates on the solution itself, not the control function.

(x) We observe that Theorem 3.6 cannot follow from arguments similar to those used in Proposition 3.3 since, although the family $\{P_x^n\}_{n=1}^{\infty}$ is weakly compact, we do not know that the finite dimensional distributions of P_x^n converge to the finite dimensional distributions* of P_x for the case where the coefficients are not Lipschitz continuous.

(xi) The theory of absolute continuity of measures corresponding to diffusions, as developed in [57, Chapters 4 and 5], cannot be used to obtain a result such as Theorem 3.6 if the drift coefficients $\{u_n(t)\}$ converge to $u(t)$ in a topology weaker than that of $\mathcal{L}_2(I)$. This is obvious from Formula (1.4) of [57, Section 5.1].

---

* See [12, Theorem 8.1, p. 54].

3.4 Compactness of Attainable Sets

Let us assume the coefficients of (2.4) satisfy : a is bounded measurable, b is bounded continuous, and $b^2(t,x) \geq \nu > 0$. In Theorem 3.8 it will be shown that for any starting probability measure, the set of probability measures induced by the solution process of (2.4), at each time $t \in I$, forms a weak compact set in $\mathcal{M}_1(R)$ if the control functions are in Σ. This result will be needed in the existence proof for a time-optimal stochastic control in the next chapter.

Theorem 3.8

If a is bounded measurable, b is bounded continuous and strictly elliptic in the stochastic differential equation (2.4), then $\mathcal{R}_\varphi^t$ is weakly compact in $\mathcal{M}_1(R)$ for each $t \in I$ and $\varphi \in \mathcal{M}_1(R)$.

Proof : Consider $\mathcal{L}_\infty(I)$ with its $\sigma(\mathcal{L}_\infty(I), \mathcal{L}_1(I))$ topology where $I = [0, t_f]$, $t_f < +\infty$. We shall show that the map $u \to \eta^u(t)\varphi$, from $\left(\mathcal{L}_\infty(I), \sigma(\mathcal{L}_\infty(I), \mathcal{L}_1(I))\right)$ into $\mathcal{M}_1(R)$ with its weak topology, is continuous. Since $\mathcal{M}_1(R)$ is a metric space [13, Theorem 6.2, p. 43], we may use sequences to prove continuity. Let $\{u_n\} \subset \mathcal{L}_\infty(I)$ converge to $u \in \mathcal{L}_\infty(I)$ in the $\sigma(\mathcal{L}_\infty(I), \mathcal{L}_1(I))$ topology, which implies by [1, p.342, Example 27] that

$$\int_0^t u_n(s)\, ds \to \int_0^t u(s)\, ds \qquad \text{as } n \to \infty$$

for all $t \in I$.

Let $P_x^n \equiv P_{0,x}^n$ and $P_x \equiv P_{0,x}$ be the measures on $\Omega = C(I, R)$ associated with (2.4) for the controls u_n and u, respectively. Then, by Remark (vii) in Section 3.3, $P_x^n \Longrightarrow P_x$ as $n \to \infty$. This implies that for each $f \in C(R)$,

$$\int_\Omega f(\pi_t(\omega))\, P_x^n(d\omega) \to \int_\Omega f(\pi_t(\omega))\, P_x(d\omega) \quad \text{as} \quad n \to \infty$$

or, equivalently,

$$\int_R f(y)\, P^n(0, x, t, dy) \to \int_R f(y)\, P(0, x, t, dy) \quad \text{as} \quad n \to \infty \tag{3.25}$$

for each $x \in R$. On integrating both sides of (3.25) with respect to the probability measure φ (and using the Dominated Convergence Theorem), we get that $\eta^{u_n}(t)\varphi \Longrightarrow \eta^u(t)\varphi$ as $n \to \infty$. Thus, $u \to \eta^u(t)\varphi$ is a continuous map from $\left(\mathcal{L}_\infty(I),\ \sigma(\mathcal{L}_\infty(I), \mathcal{L}_1(I))\right)$ into $\mathcal{M}_1(R)$ with its weak topology.

Since Σ is the translation by the function $\frac{1}{2}$ in $\mathcal{L}_\infty(I)$ of the closed ball of radius $\frac{1}{2}$ in $\mathcal{L}_\infty(I)$, it is $\sigma(\mathcal{L}_\infty(I), \mathcal{L}_1(I))$ compact. Therefore, the image of Σ, $\mathcal{R}_\varphi^t$, is weakly compact.

Q.E.D.

Of some importance in optimal control theory is the convexity of the attainable sets. In general, $\mathcal{R}_\varphi^t$ is not a convex set of $\mathcal{M}_1(R)$ as the following example illustrates. Consider the simple stochastic differential equation

$$dx_t^u = u(t)\, dt + dz_t, \tag{3.26}$$

starting a.s. W at $t = 0$ and $x = 0$. We shall show that $\mathcal{R}_{\epsilon_0}^t$ is not convex, where ϵ_0 is the Dirac measure at the origin. Let $u_1 \not\equiv u_2$ both be in Σ. Then, at time $t > 0$, $t \in I$, $x_t^{u_1}$ has a normal distribution function, $\eta^{u_1}(t)\epsilon_0$ *, with mean $\int_0^t u_1(s)\, ds$ and variance t, while $x_t^{u_2}$ has a normal distribution function, $\eta^{u_2}(t)\epsilon_0$,

* Actually, $\eta^{u_1}(t)\epsilon_0$ is the measure associated with the normal distribution function.

with mean $\int_0^t u_2(s)\,ds$ and variance t. Therefore, the distribution function $\frac{1}{2}\eta^{u_1}(t)\epsilon_0 + \frac{1}{2}\eta^{u_2}(t)\epsilon_0$ cannot possibly be normal. But (3.26) implies that the random variables x_t^u can only have a normal distribution. Thus, $\mathcal{R}_{\epsilon_0}^t$ is not convex. However, we do having the following different convexity result.

Theorem 3.9

Let a and b be as in Theorem 3.8. Let $u \in \Sigma$, $\varphi \in \mathcal{M}_1(R)$, and $f \in C_o^\infty(R)$. Then $\mathcal{R}_\varphi^t(f)$ is convex in R, i.e., an interval.

Proof : By the definition of the weak topology on $\mathcal{M}_1(R)$, the map $\mu \to \mu(f)$ from $\mathcal{M}_1(R)$ into R, for a fixed $f \in C(R)$, is continuous. Since $\mathcal{R}_\varphi^t$ is weak compact, by Theorem 3.8, this implies that $\mathcal{R}_\varphi^t(f)$ is compact in R. Let $f \in C_o^\infty(R)$ and assume $\mathcal{R}_\varphi^t(f)$ is not an interval. Then there exist two compact subsets $K_1, K_2 \subset R$ which are separated and whose union is $\mathcal{R}_\varphi^t(f)$. Let $\{\alpha_1, \beta_1\}$ and $\{\alpha_2, \beta_2\}$, $\alpha_1 \le \beta_1 < \alpha_2 \le \beta_2$, be the end points of K_1 and K_2, respectively. Choose u_1 and u_2 in Σ such that $y_1 = \eta^{u_1}(t)\varphi(f) \in K_1$ and $y_2 = \eta^{u_2}(t)\varphi(f) \in K_2$.

Divide $[0, t]$ into n equal subintervals of length Δ. Thus, using (2.7),

$$y_1 = \varphi(f) + \sum_{i=1}^{n} \int_{(i-1)\Delta}^{i\Delta} \varphi(U^{u_1}(s) A^{u_1}(s) f)\,ds \;,$$

and

$$y_2 = \varphi(f) + \sum_{i=1}^{n} \int_{(i-1)\Delta}^{i\Delta} \varphi(U^{u_2}(s) A^{u_2}(s) f)\,ds \;,$$

where

$$A^u(s) \equiv A(s) + u(s) D \;.$$

Since $y_2 > y_1$, for at least one interval $[j\Delta, (j+1)\Delta] \subset [0, t]$

$$\int_{j\Delta}^{(j+1)\Delta} \varphi\,(U^{u_2}(s)\,A^{u_2}(s)\,f)\,ds \;>\; \int_{j\Delta}^{(j+1)\Delta} \varphi\,(U^{u_1}(s)\,A^{u_1}(s)\,f)\,ds\ .$$

We can choose $\Delta > 0$ so that for some set $\mathcal{A} \subset \{1, \ldots, n\}$

$$\beta_1 < \varphi(f) + \sum_{\substack{i=1\\ i \notin \mathcal{A}}}^{n} \int_{(i-1)\Delta}^{i\Delta} \varphi\,(U^{u_1}(s)\,A^{u_1}(s)\,f)\,ds + \sum_{i \in \mathcal{A}} \int_{i\Delta}^{(i+1)\Delta} \varphi\,(U^{u_2}(s)\,A^{u_2}(s)\,f)\,ds < \alpha_2\ .$$

Define

$$u_3(\tau) = \begin{cases} u_1(\tau) & : \quad 0 \le \tau \le t\ , \quad \tau \notin \bigcup_{j \in \mathcal{A}} [\,j\Delta, (j+1)\Delta\,] \\ u_2(\tau) & : \quad \tau \in \bigcup_{j \in \mathcal{A}} [\,j\Delta, (j+1)\Delta\,] \end{cases} .$$

Obviously, $u_3 \in \Sigma$. But

$$\eta^{u_3}(t)\,\varphi(f) = \varphi(f) + \int_0^t \varphi\,(U^{u_3}(s)\,A^{u_3}(s)\,f)\,ds \notin \mathcal{R}_\varphi^t(f)\,,$$

by the definition of u_3, which is a contradition. Therefore, K_1 and K_2 are not separated. The same procedure as above proves that K_1 and K_2 cannot consist of disjoint bounded intervals themselves. Therefore, $\mathcal{R}_\varphi^t(f)$ is convex in R.

Q.E.D.

Remarks:

(i) Theorem 3.9 simply states that, given any $f \in C_0^\infty(R)$, the expected value of the random variable $f(x_t^u(\omega))$ for fixed $t \in I$, where $\{x_s^u(\omega) : s \in I\}$ is the solution process of (2.4) associated with control u, and starting at $t = 0$ and $\varphi \in \mathcal{M}_1(R)$, takes on all values in some interval as u varies through Σ.

(ii) Theorems 3.8 and 3.9 can be extended to the n-dimensional analogue of (2.4) .

(iii) Using the result of Section 5.2, Theorem 3.9 can be proved for all functions which have compact support and are twice continuously differentiable .

CHAPTER IV

OPTIMAL STOCHASTIC CONTROL

4.1 Definition of Stochastic Control Problem and Controllability

In this section, we define what we mean by a stochastic control problem. First, we focus attention on a dynamic process (flow) in $\mathcal{M}_1(E)$, $E = R^n$. This may arise, for instance, from a diffusion such as a stochastic differential equation or from a more general Markov process. Along with a process in $\mathcal{M}_1(E)$, we assume the existence of an admissible class of controls that can influence the process in question. Thirdly, there is an 'objective' to be achieved by the process using the available controls. In the next section, the objective will be to 'hit' a target in minimum time. Another objective may be to minimize the Prohorov or norm distance between some probability measure ψ and the probability measures of the flow in $\mathcal{M}_1(E)$. If the flow in $\mathcal{M}_1(E)$ is induced by a Markov process and the target is a fixed measure in $\mathcal{M}_1(E)$, the objective to 'hit' the target means that, at some time $t \in I = [0, t_f]$, the probability measure on R^n, induced by the random variable x_t, is identical to the target measure.

We describe a simple example : let the target be a fixed measure ψ on a sphere in R^3, and let the process in R^3 be a nonhomogeneous Markov process with transition function $P^u(s, x, t, \Gamma)$, where u indicates the dependence on the control. Let φ_0 be the initial probability measure of the process. If we can find a control u in the admissible class such that for some starting time t_0 and final time t_f

$$\int_{R^3} P^u(t_0, x, t_f, \Gamma)\, \varphi_0(dx) = \psi(\Gamma)$$

for all $\Gamma \in \mathcal{B}(R^3)$, then u 'hits' the target ψ.

Let $\{\tau(t) : t \in I\}$ be a family of subsets of $\mathcal{M}_1(E)$. We shall call it the <u>moving target</u>. For most applications, $\tau(t)$ will be a single point in $\mathcal{M}_1(E)$ for each $t \in I$. Unless the supports of two or more probability measures are disjoint or 'almost' disjoint in some sense, there is little physical meaning in trying to 'hit' a set consisting of more than one probability measure.

<u>Definition 4.1.</u> A stochastic control problem consists of the following :

(i) a dynamic system in $\mathcal{M}_1(E)$,

(ii) a control class $\mathcal{U}$,

(iii) an initial probability measure ω,

(iv) a target $\{\tau(t) : t \in I\}$,

(v) an objective .

Before discussing the controllability problem in the space of probability measures we shall show that viewing dynamical systems as flows of measures subsumes the theory of ordinary differential equations.

Much of the mathematical theory of deterministic control deals with a system of differential equations, in E, of the form

$$\dot{x}(t) = g(t, x(t), u(t)) \quad , \qquad x(0) = x_0 \quad , \tag{4.1}$$

where the control function $u : I \to E$ takes values in some set $J \subset E$. We define $\mathcal{U}(J)$ to be the set of bounded measurable functions taking values in J, and assume conditions on g, so that for any $u \in \mathcal{U}(J)$ a unique solution $x(\cdot)$ of (4.1) exists. The solution of (4.1) is given by

$$x(t) = x_0 + \int_0^t g(s, x(s), u(s))\, ds . \tag{4.2}$$

The attainable set $\mathcal{A}(t)$, for $t \in I$, is defined to be the set of points attainable at time t by solutions of (4.1) for all possible controls $u \in \mathcal{U}(J)$, i.e.,

$$\mathcal{A}(t) = \left\{ x_0 + \int_0^t g(s, x(s), u(s))\, ds : u \in \mathcal{U}(J) \right\} .$$

Let ϵ_x be the Dirac measure at $x \in E$ (ϵ_x has mass 1 at x, and mass 0 otherwise). Since E is a metric space, it is homeomorphic to the subset $\xi(E) \equiv \{ \epsilon_x : x \in E \}$ of $\mathcal{M}_1(E)$, where $\mathcal{M}_1(E)$ has the weak topology [13, p. 42].

We define, for solutions of (4.1),

$$\epsilon_0 = \{ \text{Dirac measure at } x(0) \}$$

$$\epsilon_t^u = \{ \text{Dirac measure at } x(t, x(0), u(t)) \} .$$

Let $\mathcal{A}^*(t) \equiv \{ \epsilon_t^u : u \in \mathcal{U}(J) \}$. Since R^n and $\xi(E)$ are homeomorphic, so are $\mathcal{A}(t)$ and $\mathcal{A}^*(t)$. Therefore, instead of studying $\mathcal{A}(t)$ in E we could, equivalently, investigate the properties of the set of Dirac measures $\mathcal{A}^*(t)$.

To summarize, by the definition of ϵ_t^u, a trajectory $\{ \epsilon_t^u : t \in I \} \subset \mathcal{M}_1(E)$ is completely equivalent to the corresponding trajectory $\{ x(t, x(0), u) : t \in I \} \subset E$. This implies that deterministic systems can be studied in $\mathcal{M}_1(E)$. For deterministic systems the attainable sets in $\mathcal{M}_1(E)$ are, of course, sets of Dirac measures, whereas, for stochastic systems, the attainable sets are more general sets of probability measures with supports non-trivial closed sets in E [13, Theorem 2.1, p. 27]. This approach, therefore, unifies the theories of deterministic and stochastic systems, since the solutions of both systems are simply flows of probability measures. We mention that for diffusions, the stochastic system reduces to deterministic systems only if the diffusion coefficient is not constrained to be strictly elliptic.

We now discuss controllability in $\mathcal{M}_1(E)$. Let $\left\{ \Gamma^u(t)\,\omega_{t_0} : t \geq t_0,\ t \in I \right\}$ represent a flow of probability measures starting at φ_{t_0} at time $t_0 \in I$.

Definition 4.2. Let $\mathcal{U}(J)$ be as above. Fix $t_1 \in I$, and for $0 \leq t \leq t_1$, define

$$\mathcal{K}_{t_1}(t) = \left\{ \varphi_t \in \mathcal{M}_1(R^n) : \Gamma^u(t_1)\,\varphi_t \in \tau(t_1) \text{ for some } u \in \mathcal{U}(J) \right\},$$

where $\tau(t_1)$ is the target set of probability measures at t_1 and φ_t is the starting probability measure at time t. $\mathcal{K}_{t_1}(t)$ is called the controllable set of probability measures at time t with respect to $\tau(\cdot)$ at time t_1. (For deterministic systems, the target set is often the Dirac measure at the origin of R^n, and the controllable sets are obviously sets of Dirac measures.)

For our purposes, we let $n = 1$ and $\tau(t) = \nu$ for all $t \in I$, where $\nu \in \mathcal{M}_1(R)$. Also, let J be a compact subset of R. Thus,

$$\mathcal{K}_{t_1}(t) = \left\{ \omega_t \in \mathcal{M}_1(R) : \Gamma^u(t_1)\,\omega_t = \nu \text{ for some } u \in \mathcal{U}(J) \right\}. \tag{4.3}$$

For a non-stationary Markov process in R with transition function $P^u(s, x, t, B)$, $\varphi_t \in \mathcal{K}_{t_1}(t)$ if and only if these exist a $u \in \mathcal{U}(J)$ such that for all $B \in \mathcal{B}(R)$

$$\int_R P^u(t, x, t_1, B)\,\varphi_t(dx) = \nu(B).$$

For the dynamic system (2.7), $\omega \in \mathcal{K}_{t_1}(0)$ if and only if there exists a $u \in \mathcal{U}(J)$ such that

$$\omega(f) + \int_0^t \omega(U^u(s)\,A^u(s)\,f)\,ds = \nu(f)$$

for all $f \in C_0^\infty(R)$, where $A^u(s) = A + u(s)\,D$.

Let

$$C_\nu = \left\{ f \in C_0^\infty (R) : \text{supp}\, f \cap \text{supp}\, \nu = \emptyset \right\}^* .$$

Then a simple necessary condition for φ to be in $\mathcal{K}_{t_1}(0)$ is that there exist $u \in \mathcal{U}(J)$ such that for all $f \in C_\nu$,

$$\int_0^{t_1} \varphi \left(U^u (s) A^u (s) f \right) ds = -\varphi (f) .$$

Intuitively, controllability with respect to ν means that there exists an admissible control u such that the flow of probability measures $\{\eta^u(t)\varphi : t \in I\}$ (φ in the controllable set at time 0), 'hits' ν at time t_1, i.e., $\eta^u(t_1)\varphi = \nu$. If ν has compact support, then controllability with respect to ν implies that for some $u \in \mathcal{U}(J)$, $x^u(t_1, \omega) \in \text{supp}\, \nu$ a.s.W. This is also a finite time stochastic stability result.

4.2 Continuity of Attainable Sets and Time-Optimal Stochastic Control

In this section, we shall study the continuity properties of the attainable set of probability measures, $\mathcal{R}_\varphi^t$, of the stochastic control system (2.4) starting at $t = 0$ with the probability measure φ. As in the deterministic case, continuity of the attainable sets, in an appropriate topology, is an essential requirement for the existence proof of time-optimal controls.

---

* supp $(\cdot)$ is the support of $(\cdot)$.

We shall show that the map $t \to \mathcal{R}_{\varphi}^{t}$ from $I = [0, t_f]$ into the non-empty compact subsets $\mathcal{K}$ of $\mathcal{M}_1(R)$ is continuous with respect to a metric h, where $\mathcal{M}_1(R)$ has the weak topology. The weak topology is a metric topology with the Prohorov metric ρ [12].

Definition 4.3. For $(V, W) \in \mathcal{K} \times \mathcal{K}$, define

$$\rho(\mu, W) = \inf\left\{\rho(\mu, \varphi) : \varphi \in W\right\}$$

$$d(V, W) = \sup\left\{\rho(\mu, W) : \mu \in V\right\}$$

$$h(V, W) = \max\left\{d(V, W), d(W, V)\right\}.$$

h is a metric on $\mathcal{K}$. We shall call h the Prohorov - Hausdorff metric. Let $N(V, \epsilon)$ be an ϵ ρ - ball of V. Then, $h(V, W) \leq \epsilon$ if and only if $V \subset N(W, \epsilon)$ and $W \subset N(V, \epsilon)$ [16, p. 205].

Lemma 4.1

Let x_t^u be the solution process of the stochastic differential equation (2.4) where $u \in \Sigma$, and the random variable $x_0 \equiv x_0^u$ has initial probability measure $\varphi \in \mathcal{M}_1(R)$. Then, $x_t^u \to x_{t'}^u$ in probability as $|t - t'| \to 0$, uniformly with respect to $u \in \Sigma$.

Proof : From (2.4), for $t \in I$, $u \in \Sigma$,

$$x_t^u = x_0 + \int_0^t (a(s, x_s^u) + u(s))\, ds + \int_0^t b(s, x_s)\, dz_s ,$$

where x_0 induces the probability measure φ. Let $t' \in I$, and α be the upper bound of a and b. Then,

$$x_t^u - x_{t'}^u = \int_{t'}^t (a(s, x_s^u) + u(s))\, ds + \int_{t'}^t b(s, x_s^u)\, dz_s$$

and using the properties of stochastic integrals and Hölder's Inequality,

$$E^W (x_t^u - x_{t'}^u)^2 \le (\alpha+1)^2 |t-t'|^2 + \alpha^2 |t-t'| ,$$

which is independent of u. Therefore,

$$E^W (x_t^u - x_{t'}^u)^2 \to 0 ,$$

uniformly for $u \in \Sigma$ as $|t-t'| \to 0$. Chebyshev's Inequality implies that

$$W\left\{ |x_t^u - x_{t'}^u| \ge \epsilon \right\} \le \frac{E^W \left\{ x_t^u - x_{t'}^u \right\}^2}{\epsilon^2} ,$$

which yields the result.

Q.E.D.

<u>Lemma 4.2</u>

Let x_t^u be as defined in Lemma 4.1. Let $\eta^u(t)\varphi$ be the probability measure induced by x_t^u, starting with $\varphi \in \mathcal{M}_1(R)$. Then the map $t \to \mathcal{R}_\varphi^t$ from I into $\mathcal{K}$ is continuous with respect to the Prohorov-Hausdorff metric h.

<u>Proof :</u> Let $F_t^u(\cdot)$ be the distribution function induced by x_t^u starting at $t = 0$ with probability measure φ. We know from Lemma 4.1 that, given any $\epsilon > 0$, there exists $\delta > 0$ such that $|t-t'| < \delta$ implies

$$\sup_{u \in \Sigma} W\left\{ |x_t^u - x_{t'}^u| \ge \epsilon \right\} \le \epsilon , \tag{4.4}$$

Now, for any $y \in R$,

$$\begin{aligned} F_t^u(y-\epsilon) - F_{t'}^u(y) &= W\left\{x_t^u \le y-\epsilon\right\} - W\left\{x_{t'}^u \le y\right\} \\ &\le W\left\{x_t^u \le y-\epsilon\right\} - W\left\{x_t^u \le y-\epsilon,\ x_{t'}^u \le y\right\} \\ &= W\left\{x_t^u \le y-\epsilon,\ x_{t'}^u > y\right\} \\ &\le W\left\{|x_t^u - x_{t'}^u| > \epsilon\right\} \le \epsilon \end{aligned}$$

for all $u \in \Sigma$, by (4.4). Interchanging x_t^u and $x_{t'}^u$ and replacing y by $y + \epsilon$, we obtain

$$F_{t'}^u(y) - F_t^u(y+\epsilon) \leq \epsilon$$

for all $u \in \Sigma$. Thus, x_t^u converges in distribution to $x_{t'}^u$ uniformly in $u \in \Sigma$ as $|t - t'| \to 0$. This is equivalent to saying that

$$\eta^u(t)\varphi \Longrightarrow \eta^u(t')\varphi \; ,$$

uniformly with respect to $u \in \Sigma$ as $|t - t'| \to 0$. Hence, given any $\epsilon > 0$, we find $\delta > 0$ such that

$$\mathcal{R}_\varphi^t \subset N(\mathcal{R}_\varphi^{t'}, \epsilon) \; ,$$

and

$$\mathcal{R}_\varphi^{t'} \subset N(\mathcal{R}_\varphi^t, \epsilon)$$

for $|t - t'| < \delta$. Thus,

$$h(\mathcal{R}_\varphi^t, \mathcal{R}_\varphi^{t'}) < \epsilon$$

for $|t - t'| < \delta$.

Q.E.D.

We now consider the stochastic control problem, as defined in Section 4.1, for the following system :

(i) $n = 1$ and the flow in $\mathcal{M}_1(R)$ is given by (2.7) ,

(ii) the control class is Σ ,

(iii) the initial probability measure is φ ,

(iv) the moving target $\tau(\cdot)$ is a continuous map from I into with respect to the metric h ,

(v) the objective is to have the flow $\{\eta^u(t)\varphi : t \in I\}$, associated with (2.4) , hit $\tau(\cdot)$ in minimum time.

We prove the following existence theorem which is similar in form to its deterministic counterpart (see [14], for instance).

<u>Theorem 4.3</u> <u>(Existence of Time-Optimal Stochastic Control)</u>

If there exists a control $u \in \Sigma$ which steers $\varphi \in \mathcal{M}_1(R)$, for the stochastic system (2.4), to the target $\tau(\cdot)$, then there exists $u^* \in \Sigma$ which steers φ to the target in minimum time . (u^* is said to be stochastically time-optimal.)

<u>Proof</u> : We are given the existence of a $u \in \Sigma$ such that $\eta^u(t)\varphi \in \tau(t)$ for some $t \in I$. This means that

$$\tau(t) \cap \mathcal{R}_\varphi^t \neq \emptyset \quad .$$

Define

$$t^* = \inf\left\{t \in I : \tau(t) \cap \mathcal{R}_\varphi^t \neq \emptyset\right\} \quad .$$

We claim that $\tau(t^*) \cap \mathcal{R}_\varphi^{t^*} \neq \emptyset$. If not,

$$h(\tau(t^*), \mathcal{R}_\varphi^{t^*}) = \delta > 0$$

since $\tau(t)$ is ρ - compact by assumption, and $\mathcal{R}_\varphi^t$ is ρ- compact by Theorem 3.8. Thus, there exists an $\alpha > 0$ such that

$$h(\mathcal{R}_\varphi^t, \mathcal{R}_\varphi^{t^*}) < \frac{\delta}{2},$$

and

$$h(\tau(t), \tau(t^*)) < \frac{\delta}{2}$$

for all $t^* < t < t^* + \delta$, since the map $s \to \tau(s)$ and $s \to \mathcal{R}_\varphi^s$, from I into $\mathcal{K}$, are continuous in the h metric, by the hypothesis and Lemma 4.2, respectively. For such a t

$$\delta = h(\tau(t^*), \mathcal{R}_\varphi^{t^*}) \leq h(\tau(t^*), \tau(t)) + h(\tau(t), \mathcal{R}_\varphi^t) + h(\mathcal{R}_\varphi^t, \mathcal{R}_\varphi^{t^*})$$
$$< \delta + h(\tau(t), \mathcal{R}_\varphi^t)$$

i.e.

$$h(\tau(t), \mathcal{R}_\varphi^t) > 0,$$

which means that $\tau(t) \cap \mathcal{R}_\varphi^t = \emptyset$, contradicting the choice of t^*. Therefore, there exists $u^* \in \Sigma$ such that $\eta^{u^*}(t^*) \varphi \in \tau(t^*)$ and no $t < t^*$ has this property.

Q.E.D.

We now discuss a necessary condition for a control $u \in \Sigma$ to be time-optimal for the stochastic system (2.4) starting at $\varphi \in \mathcal{M}_1(R)$.

Let

$$\text{supp}\,\tau(t) = \bigcup \left\{ \text{supp}\,\varphi : \varphi \in \tau(t) \right\},$$

and

$$Z = \cup \left\{ \operatorname{supp} \tau(t) : t \in I \right\} .$$

Define

$$F(Z) = \left\{ f \in C_0^{\infty}(R) : \operatorname{supp} f \cap Z = \emptyset \right\} .$$

If $Z = R$, then obviously $F(Z) = \emptyset$. For the important case where $\tau(t) = \nu$ for all $\tau \in I$ and $\nu \in \mathfrak{M}_1(R)$ has compact support, $F(Z) \neq \emptyset$. Assuming $F(Z) \neq \emptyset$, take $f \in F(Z)$. Then, since a time-optimal control exists, $\eta^{u^*}(t^*)\varphi \in \tau(t^*)$ for some $u^* \in \Sigma$, $t^* \in I$, and (2.7) becomes

$$\varphi(f) + \int_0^{t^*} \varphi(U^{u^*}(s)\, A^{u^*}(s)\, f)\, ds = 0 . \tag{4.5}$$

So, if for some $u^* \in \Sigma$, (4.5) is satisfied for all $f \in F(Z)$, it may be a time-optimal control. This condition is in general far from sufficient for optimality, but may serve to disquality suspected time-optimal controls.

We saw in Chapter III that in general $\mathcal{R}_{\varphi}^{t}$ is not convex. However, for certain state-dependent control systems, where $\mathcal{R}_{\varphi}^{t}$ does not consist only of probability measures whose associated distributions functions are normal, as for the system (3.26), $\mathcal{R}_{\varphi}^{t}$ may be convex. In this case, we can carry the time-optimal stochastic control theory further.

<u>Definition 4.4.</u> We define the <u>reachable cone of probability measures</u> of the stochastic system (2.4) by

$$\mathcal{R}(\varphi) \equiv \left\{ (t, \mathcal{R}_{\varphi}^{t}) : t \in I \right\} .$$

<u>Definition 4.5.</u> We say that $u \in \Sigma$ is a <u>boundary control</u> on $[0, t_1] \subset I$ if $\eta^u(\cdot)\varphi$ lies on the boundary of $\mathcal{R}(\varphi)$, $\partial\mathcal{R}(\varphi)$, on $[0, t_1]$, i.e. for each $0 \le t \le t_1$, $\eta^u(t)\varphi \in \partial\mathcal{R}(\varphi)$.

Assume that $\tau(t)$, the target set, is continuous with respect to h. Then, if $\mathcal{R}^t_\varphi$ is convex, one would expect as in the deterministic case that, if a control u^* is time-optimal on $[0, t^*] \subset I$, then $\eta^{u^*}(t^*)\varphi \in \partial\mathcal{R}^{t^*}_\varphi$. We can actually prove:

<u>Theorem 4.4.</u> (Necessary condition for time-optimal stochastic control)

Let $u^* \in \Sigma$ be time-optimal on $[0, t^*] \subset I$. Then u^* is a boundary control on $[0, t^*]$ if $\mathcal{R}^t_\varphi$ is convex for all $t \in I$.

<u>Proof:</u> The proof is exactly as in the deterministic case [14, p. 65-67], except that the Prohorov-Hausdorff metric is used in place of the Hausdorff metric on R^n. The proof requires the continuity, convexity, and compactness of $\mathcal{R}^t_\varphi$ as well as the continuity of $\tau(t)$.

For $\mathcal{R}^t_\varphi$ convex, we can obtain another type of necessary condition for a control to be time-optimal. Let $u^* \in \Sigma$ be time-optimal and t^* the minimum time. Then, by Theorem 4.4, $\varphi_t \equiv \eta^{u^*}(t)\varphi \in \partial\mathcal{R}^t_\varphi$ for $t \in [0, t^*]$. Let u be any other control in Σ and $\nu \equiv \eta^u(t)\varphi$. On $\mathcal{M}_1(R)$, the weak topology and weak*topology $(\sigma(\mathcal{M}_1(R), C_o(R)))$ are identical [18, Theorem 4.4.4, p. 81]. Hence, by Theorem 3.8, $\mathcal{R}^t_\varphi$ is weak*compact which implies that $\mathcal{R}^t_\varphi$ is weak*closed [1, Corollary V.4.3, p. 424]. Therefore, $\mathcal{R}^t_\varphi$ is a convex, weak*closed, subset of $\mathcal{M}(R)$ for all $t \in I$. Then, in the light of [54, Theorem 1], for each $t \in [0, t^*]$, given any $\epsilon > 0$, (i) there exists $\varphi^\epsilon_t \in \partial\mathcal{R}^t_\varphi$ such that

$\| \varphi_t - \varphi_t^\epsilon \| < \epsilon$, and (ii) there exists $f_t^\epsilon \in C_o(R)$, $f_t^\epsilon \neq 0$, such that $\nu(f_t^\epsilon) \leq \varphi_t^\epsilon(f_t^\epsilon)$ for all $\nu \in \mathcal{R}_\varphi^t$. Thus, for all $\nu \in \mathcal{R}_\varphi^t$,

$$|\nu(f_t^\epsilon)| \leq |\varphi_t(f_t^\epsilon)| + |\varphi_t^\epsilon(f_t^\epsilon) - \varphi_t(f_t^\epsilon)| \leq |\varphi_t(f_t^\epsilon)| + \epsilon \| f_t^\epsilon \|$$

or

$$|\nu(g_t^\epsilon)| \leq |\varphi_t(g_t^\epsilon)| + \epsilon ,$$

where $g_t^\epsilon \equiv \dfrac{f_t^\epsilon}{\| f_t^\epsilon \|}$. Hence, given any $\delta > 0$, there exists $g_t^\delta \in C_o(R)$, $g_t^\delta \neq 0$, such that for all $\nu \in \mathcal{R}_\varphi^t$

$$|\nu(g_t^\delta)| \leq |\varphi_t(g_t^\delta)| + \delta .$$

In terms of (2.8), this condition becomes

$$\left| \int_0^t \varphi(U^u(s) A^u(s) g_t^\delta) ds \right| \leq \left| \int_0^t \varphi(U^u(s) A^u(s) g_t^\delta) ds \right| + \delta \tag{4.6}$$

for all $u \in \Sigma$. Therefore, a necessary condition for u^* to be time-optimal is that for all $\delta > 0$ and $t \in [0, t^*]$, (4.6) is satisfied.

4.3 Existence of Minimum Cost Feedback Stochastic Controls

In this section we shall be dealing with the feedback stochastic control system (2.4) where $a(t, x)$ is bounded measurable, and $b(t, x)$ is bounded continuous and strictly elliptic.

Let the system (2.4) start at $t = 0$ with $\varphi \in \mathcal{M}_1(R)$ and suppose we are given a cost function $V \in C(R)$. The problem is to prove the existence of a control u in a specified admissible class which minimizes the expected value of $V(x_{t_1}^u(\omega))$ at some time $t_1 \in I$. Such a control, if it exists, will be referred to as optimal. If we let $\mathcal{G} = \mathcal{L}_\infty(I \times R)$, the space of bounded measurable functions from $I \times R$ into R, be the admissible control class, then we wish to prove the existence of a $\bar{u} \in \mathcal{Y}$ such that

$$\eta^{\bar{u}}(t_1)\,\varphi\,(V) \;\leq\; \eta^{u}(t_1)\,\varphi\,(V)$$

for all $u \in \mathcal{Y}$.

We say that $\{u_n\} \subset \mathcal{Y}$ converges to $u \in \mathcal{Y}$ in the $\sigma(\mathcal{L}_\infty(I \times R), \mathcal{L}_1(I \times R))$ topology if and only if

$$\int_I \int_R u_n(t,x)\, g(t,x)\, dt\, dx \;\to\; \int_I \int_R u(t,x)\, g(t,x)\, dt\, dx$$

for all $g \in \mathcal{L}_1(I \times R)$, the space of integrable functions on $I \times R$.

Define

$$\tilde{\Sigma} \equiv \left\{ u \in \mathcal{L}_\infty(I \times R) \;:\; 0 \leq u(t,x) \leq 1,\; t \in I,\; x \in R \right\}.$$

<u>Theorem 4.5</u>

Let a be bounded measurable, b bounded continuous and strictly elliptic in (2.4). Let $t_1 \in I$ and $V \in C(R)$. Then there exists a $\bar{u} \in \tilde{\Sigma}$ such that

$$\eta^{\bar{u}}(t_1)\,\varphi\,(V) \;\leq\; \eta^{u}(t_1)\,\varphi\,(V)$$

for all $u \in \tilde{\Sigma}$, where $\varphi \in \mathcal{M}_1(R)$.

<u>Proof :</u> Let $\{u_n\} \subset \tilde{\Sigma}$ converge to $u \in \tilde{\Sigma}$ in the $\sigma(\mathcal{L}_\infty(I \times R), \mathcal{L}_1(I \times R))$ topology. Then, using an extended version of Corollary 3.7,* it can be shown that $\eta^{u_n}(t)\,\varphi \Rightarrow \eta^{u}(t)\,\varphi$ as $n \to \infty$. Therefore, the map $u \to \eta^{u}(t)\,\varphi$, from $\left(\mathcal{L}_\infty(I \times R),\, \sigma(\mathcal{L}_\infty(I \times R), \mathcal{L}_1(I \times R))\right)$ into $\mathcal{M}_1(R)$ with its weak topology, is continuous.

Since $\tilde{\Sigma}$ is the translation by the function $\frac{1}{2}$ of the ball with radius $\frac{1}{2}$ in $\mathcal{L}_\infty(I \times R)$, it is $\sigma(\mathcal{L}_\infty(I \times R), \mathcal{L}_1(I \times R))$ compact, implying that the image

* Extended in the sense that the control is also a function of the state.

of $\tilde{\Sigma}$, $\tilde{\mathcal{R}}_{\varphi}^{t} \equiv \{ \eta^{u}(t)\,\varphi : u \in \tilde{\Sigma} \}$, is weakly compact in $\mathcal{M}_1(R)$.

By the definition of the weak topology on $\mathcal{M}_1(R)$, the map $\mu \to \mu(V)$ is a continuous function from $\mathcal{M}_1(R)$ into R. Thus, $\tilde{\mathcal{R}}_{\varphi}^{t}(V)$ is a compact set in R for each $t \in I$. Therefore, there exists a control $\bar{u} \in \tilde{\Sigma}$ such that

$$\eta^{\bar{u}}(t_1)\,\varphi(V) \le \eta^{u}(t_1)\,\varphi(V)$$

for all $u \in \tilde{\Sigma}$.

Q.E.D.

Remarks:

(i) In Section 5.3 we shall obtain a necessary and sufficient condition for a control to be optimal.

(ii) Theorem 4.5 is similar in form to Theorem 3 of [23], but the proof here seems to be somewhat easier.

CHAPTER V

EXISTENCE OF UNIQUE QUASI-DIFFUSIONS WITH APPLICATION TO OPTIMAL STOCHASTIC CONTROL THEORY

5.1 Introduction

The aim of this chapter is to study one-dimensional processes associated with the coefficients $[a(x), b(x)]$, where a is the drift coefficient and b is the diffusion coefficient. If the coefficients are smooth, a unique transition function $P(t, x, \Gamma)$, $t \geq 0$, $x \in R$, $\Gamma \in \mathcal{B}(R)$ can be associated with $[a, b]$, where the density of the measure $P(t, x, \cdot)$ is the solution of the Kolmogorov forward equation. If both a and b are bounded and uniformly Hölder continuous and b is strictly elliptic, then we can still associate a unique transition density function $p(t, x, y)$ with $[a, b]$ as the solution of the Kolmogorov backward equation. When a and b are not Hölder continuous, the classical theory of parabolic differential equations does not imply the existence of a fundamental solution to the backward equation. To relate a Markov process to $[a, b]$, one must therefore resort to other methods. For $a(x)$ bounded measurable, $b(x)$ bounded uniformly continuous and strictly elliptic, Tanaka [47] and Krylov [48] were able to construct a quasi-diffusion (see (2.2)) corresponding in some sense to $[a, b]$. However, they were unable to show that the resulting semi-group is unique. Therefore, they could not uniquely identify the quasi-diffusion with $[a, b]$. This difficulty is overcome in [5], by using the martingale approach, where it is shown that to each $[a, b]$ there corresponds a unique probability measure P_x on $(\Omega, \mathfrak{F}^o)$ which solves the martingale problem starting at $x \in R$. If one can also associate a semi-group with $[a, b]$, as is done for instance in [47] and [48], then it can be shown [5, Part II, Theorem 11.1] that this semi-group must be unique. The uniqueness of P_x, however, does not ensure the pathwise uniqueness which results from Ito's formulation of stochastic differential equations.

The major portion of the work in [47] and [48] is devoted to showing that the Markov process constructed for the poorly behaved coefficients is a quasi-diffusion. The approach presented in [47], [48] and [5] is highly probabilistic in its nature. In the next section, we shall derive a formula similar to (2.2), using only functional analytic methods, for the case where the drift coefficient is bounded and integrable on R. With this formula, the uniqueness of the semi-group, generated by $[a, b]$ (in the sense of [48]), can be readily established.

The main difficulty in proving uniqueness from (2.2) is that the transition function which appears is that associated with the poorly behaved coefficients, and since very little is known about this transition function (it emerges from the Riesz-Markov Theorem), uniqueness is difficult to prove. The crux of our approach is to find a representation for the semi-group $\{U(t) : t \geq 0\}$, associated with the poorly behaved coefficients, in terms of a semi-group $\{T(t) : t \geq 0\}$ associated with well-behaved coefficients. Since a great deal is known about the unique transition density function $p(t, x, y)$ corresponding to $\{T(t) : t \geq 0\}$, uniqueness of $\{U(t) : t \geq 0\}$ follows from the properties of $p(t, x, y)$.

Finally, we remark that the techniques of the next section do not appear to be as general as those in [5]. The main contribution is that the uniqueness of the semi-group, generated by a bounded integrable drift coefficient, is demonstrated from purely functional analytic considerations. The analysis is done for a one-dimensional system, but the extension to n-dimensions is straight forward. In Section 5.3, we apply the results of Section 5.2 to obtain a necessary and sufficient condition for a control, in a certain admissible class, to minimize the average of a cost functional.

5.2 Existence of Unique Quasi-Diffusions

From now on, let a and b be bounded uniformly Hölder continuous, and let b also be strictly elliptic. Let $u(x)$ be a real-valued bounded, measurable, and integrable function on R.

Define

$$h_\lambda(\beta) = \frac{2}{\pi} \frac{\lambda}{\lambda^2 + \beta^2} , \qquad \lambda > 0 .$$

Then,

$$u_\lambda(x) \equiv h_\lambda * u(x) \equiv \int_{-\infty}^{\infty} h_\lambda(y)\, u(x-y)\, dy$$

is infinitely differentiable in x. Since u is integrable over R, it follows from [20, Problem 13, pp. 196-197] that

$$\lim_{\lambda \downarrow 0} u_\lambda(x) = u(x) \qquad \text{a.e. on } R .$$

Since u_λ is continuously differentiable, the mean value theorem implies that u_λ is uniformly Lipschitz continuous. Thus, there exists a sequence $\{u_n\}$ of uniformly Lipschitz continuous functions such that $\lim_{n \to \infty} u_n(x) = u(x)$ a.e. on R.

We now construct a Markov process for the coefficients $[a+u, b]$. Let

$$A_n = (a(x) + u_n(x)) \frac{\partial}{\partial x} + \frac{b^2(x)}{2} \frac{\partial^2}{\partial x^2} .$$

Then, for each n, [8, Vol. 1, Theorem 5.11] implies the existence of a unique diffusion process $(\Omega, \mathfrak{F}_t^o, P_x^n, \pi_t, x \in R)$, where the transition density function $p_n(t, x, y)$ is the fundamental solution of

$$\frac{\partial v}{\partial t} = A_n v .$$

Let $P_n(t, x, \Gamma) = \int_\Gamma p_n(t, x, y)\, dy$ and

$$U_n(t) f(x) \equiv \int_R f(y) P_n(t, x, dy)$$

for $n > 0$, where $P^n(t, x, \Gamma) = P_x^n\{\pi_t \in \Gamma\}$, $\Gamma \in \mathcal{B}(R)$, and $f \in B(R)$. Define for $t \geq 0$

$$U(t) f(x) = \lim_{n \to \infty} U^n(t) f(x) ,$$

where f has compact support and is three times continuously differentiable. Proceeding exactly as in [48, Section 1], it can be shown that $U(t) f$ is representable as

$$U(t) f(x) = \int_R f(y) P^U(t, x, dy) \tag{5.1}$$

for any $f \in C(R)$, where $P^U(t, x, \Gamma)$ is a transition function. Employing the properties of $P^U(t, x, \Gamma)$ proved in [48, Section 1], we can conclude on the basis of [5, Part 1, Theorem 3.14] that there corresponds a continuous Markov process to $P^U(t, x, \Gamma)$. We denote this Markov process by $(\Omega, \mathfrak{F}_t^o, P_x^U, \pi_t, x \in R)$.

Now that we have associated a Markov process with $[a+u, b]$, we shall obtain a representation for $\left\{ U(t) f : t \geq 0 \right\}$ which is different from (5.1). First, let $(\Omega, \mathfrak{F}_t^0, P_x, \pi_t, x \in R)$ be the Markov process generated by the coefficients $[a, b]$, which satisfy the conditions of the first paragraph of this section. In view of [8, Vol. 1, Theorem 5.11], the transition density function of this Markov process is the fundamental solution of the parabolic differential equation

$$\frac{\partial v}{\partial t} = A v , \tag{5.2}$$

where

$$A = a(x) \frac{\partial}{\partial x} + \frac{b^2(x)}{2} \frac{\partial^2}{\partial x^2} . \tag{5.3}$$

The transition density function $p(t, x, y)$ induces a unique semi-group $\left\{ T(t) : t \geq 0 \right\}$ of bounded linear operators on $B(R)$ through the relation

$$T(t) f(x) = \int_R f(y) \, p(t, x, y) \, dy .$$

A is the s-infinitesimal generator of $\left\{ T(t) : t \geq 0 \right\}$ with domain $\mathfrak{D}(A)$ containing

$$\mathfrak{D} \equiv \left\{ f \in C(R) : f \text{ has compact support }, f', f'' \in C(R) \right\} .$$

Integrating (5.2), we get Dynkin's formula :

$$T(t) f - f = \int_0^t A \, T(s) f \, ds$$

for $f \in \mathfrak{D}$.

We are concerned with defining a unique semi-group for the (formal) stochastic differential equation

$$d\pi_t = (a(\pi_t) + u(\pi_t))\,dt + b(\pi_t)\,dz_t, \qquad t \geq 0, \text{ a.s. } P^u, \tag{5.4}$$

where P^u is some probability measure on $(\Omega, \mathfrak{F}^o)$, and a, b, u are as above.

In view of the estimates in [8, Vol. II, Theorem 0.5], it can be shown that for each $n > 0$ and $f \in \mathcal{D}$, both $T(t)\,f(x)$ and $U_n(t)\,f(x)$ are differentiable in t, and twice continuously differentiable in x and satisfy

$$\frac{\partial T(t)\,f}{\partial t} = A\,T(t)\,f \quad \text{and} \quad \frac{\partial U_n(t)\,f}{\partial t} = A_n\,U_n(t)\,f, \tag{5.5}$$

respectively, with the respective initial conditions,

$$\operatorname*{s-lim}_{t \downarrow 0} T(t)\,f = f \quad \text{and} \quad \operatorname*{s-lim}_{t\downarrow 0} U_n(t)\,f = f. \tag{5.6}$$

We claim that the solution of the second equation in (5.5) can be represented as

$$U_n(t)\,f = T(t)\,f + \int_0^t T(t-s)\,u_n(\cdot)\,D\,U_n(s)\,f\,ds \;^* \tag{5.7}$$

for any $n > 0$ and $f \in \mathcal{D}$. For $t > 0$ we write the incremental ratio as

$$\begin{aligned} \frac{1}{\Delta}(U_n(t+\Delta)f - U_n(t)f) &= \frac{1}{\Delta}\Big(T(t+\Delta)\,f - T(t)\,f\Big) \\ &+ \left(\frac{T(\Delta)-I}{\Delta}\right)\int_0^t T(t-s)\,u_n(\cdot)\,D\,U_n(s)\,f\,ds \\ &+ \frac{1}{\Delta}\int_t^{t+\Delta} T(t+\Delta-s)\,u_n(\cdot)\,D\,U_n(s)\,f\,ds. \end{aligned} \tag{5.8}$$

---

*$D \equiv \frac{\partial}{\partial x}$ is a closed linear operator on $C_o(R)$ with domain, $\mathcal{D}(D)$, containing $\mathcal{D}_1 = \Big\{ f \in C(R):\ f \text{ has compact support},\ f' \in C(R)\Big\}$.

The first term on the right-hand side of (5.8) approaches $AT(t)f$, since $f \in \mathfrak{D}$ and $T(t)f \in \mathfrak{D}(A)$. The third term approaches $u_n(\cdot)\, D\, U_n(t) f$ since the map $s \to T(t+\Delta-s)\, u_n(\cdot)\, D\, U_n(s) f$ from $[0, t+\Delta]$ into $C_o(R)$ is s-continuous. This follows from $\| D\, U_n(s) \| < \infty$ for $s > 0$ (Lemma 5.1), and $\| U_n(\delta) f - f \| \to 0$ as $\delta \to 0$ [8, Vol. II, Equation (5.69)]. Since the left-hand side of (5.8) goes to $\dfrac{\partial U_n(t) f}{\partial t}$, which exists by (5.5), we have

$$\int_0^t T(t-s)\, u_n(\cdot)\, D\, U_n(s) f\, ds \in \mathfrak{D}(A) ,$$

and

$$\frac{\partial U_n(t) f}{\partial t} = A\,T(t) f + A \int_0^t T(t-s)\, u_n(\cdot)\, D\, U_n(s) f\, ds + u_n(\cdot)\, D\, U_n(t) f$$

$$= A_n\, U_n(t) f .$$

The following lemmas are presented in preparation for Theorem 5.4.

Lemma 5.1

The linear operator $D\,T(t)$ is bounded on $\mathfrak{D}$ and, since $\overline{\mathfrak{D}} = C_o(R)$, can be extended to a unique bounded linear operator with the same norm and symbol on $C_o(R)$ for any $t > 0$, and

$$\int_0^\tau \| D\,T(t) \|\, dt < \infty$$

for any $0 < \tau < \infty$.

Proof: Let $f \in \mathfrak{D}$, and let $p(t, x, y)$ be the transition density function associated with $\{T(t) : t \geq 0\}$. By [8, Vol. II, Theorem 0.5], we know that $p(t, x, y)$ is continuously differentiable in x and

$$\left| \frac{\partial p(t, x, y)}{\partial x} \right| \leq \frac{M}{t} \exp\left\{ - \frac{\alpha (y-x)^2}{t} \right\} ,$$

where M and α are positive constants. Thus,

$$| D T(t) f(x) | = \left| \int_R f(y) \frac{\partial p(t, x, y)}{\partial x} dy \right| .$$

$$\leq \| f \| \ M \sqrt{\frac{2\pi}{\alpha t}} \int_R \frac{1}{\sqrt{2\pi t}} \exp\left\{ - \frac{z^2}{t} \right\} dz ,$$

where $z = \frac{y-x}{\alpha}$. Hence,

$$\| D T(t) \| \leq M \sqrt{\frac{2\pi}{\alpha t}} \equiv \psi(t) ,$$

and

$$\int_0^\tau \| DT(t) \| \, dt < \infty$$

for $0 \leq \tau < \infty$.

Q.E.D.

From now on let u, bounded by γ, integrable on R, and vanishing at $\pm\infty$, be fixed, and let $\{u_n\}$ be a sequence of infinitely differentiable functions (as described earlier) converging a.e. to u on R .

Lemma 5.2

For $t > 0$, $n > 0$, $D U_n(t)$ is a bounded linear operator on $\mathfrak{D}$, and therefore can be extended to a unique bounded linear operator on $C_o(R)$ with the same symbol and norm, and $\| D U_n(t) \|$ is uniformly integrable with respect to n over $[0, \tau]$, $\tau < \infty$.

Proof: By the construction of u_n, it is bounded by γ and uniformly Lipschitz continuous. Therefore, Lemma 5.1 implies that $\| D U_n(t) \| < \infty$ for $t > 0$ and

$$\int_0^{\tau} \| D U_n(t) \| \, dt < \infty \tag{5.9}$$

for each $n > 0$, and $\tau < \infty$.

Now, for each $f \in \mathscr{D}$, $T(t-s) u_n(\cdot) D U_n(s) f \in \mathscr{D}(D)$, where $0 \le s \le t < \infty$. Therefore,

$$\int_0^t \| D T(t-s) u_n(\cdot) D U_n(s) f \| \, ds \le \gamma \| f \| \int_0^t \| D T(t-s) \| \, \| D U_n(s) \| \, ds$$

$$< \infty \quad , \tag{5.10}$$

Since the convolution of two integrable functions is itself integrable [1, Lemma VIII.1.24]. (The existence of the right-hand side is assured by Lemma 5.1 and (5.9), for each $n > 0$.). Hence, invoking [1, Theorem III.6.20, p. 153], we can operate on both sides of (5.7) with D to get

$$D U_n(t) f = D T(t) f + \int_0^t D T(t-s) u_n(\cdot) D U_n(s) f \, ds \ . \tag{5.11}$$

Then,

$$\| D U_n(t) \| \le \psi(t) + \gamma \int_0^t \psi(t-s) \| D U_n(s) \| \, ds < \infty \ . \tag{5.12}$$

Let $I = [0, t_f]$, and $\rho = \int_0^{t_f} \psi(s)\, ds$. Then, integrating both sides of (5.12) from 0 to $\tau \in I$, $\tau \geq t$, we get

$$h_n(\tau) \leq \rho + \gamma \rho \int_0^{\tau} h_n(s)\, ds ,$$

where $h_n(t) \equiv \int_0^t \| D U_n(s) \| ds$, which implies, by Gronwall's Lemma [11, p. 11], that for all $n > 0$

$$h_n(\tau) \leq \rho e^{\gamma \rho \tau} .$$

Q.E.D.

Let

$$\lambda(t) = \psi(t) + \gamma \rho^2 e^{\gamma \rho t} .$$

Then,

$$\| D U_n(t) \| \leq \lambda(t)$$

for all $n > 0$, $t \in I$, and $\lambda(t)$ is integrable over $[0, \tau]$, $\tau \in I$.

<u>Lemma 5.3</u>

For each $t \geq 0$, $\{ D U_n(t) \}_{n \geq 0}$ converges in the uniform operator topology on $\mathcal{L}(C_o(R), C_o(R))$* to $V(t) \in \mathcal{L}(C_o(R), C_o(R))$ as $n \to \infty$, uniformly on every finite interval.

<u>Proof:</u> For each $t \geq 0$, and $n, m > 0$, $D U_n(t) - D U_m(t)$ can be extended uniquely from $\mathfrak{D}$ to a bounded linear operator on $C_o(R)$ with the same symbol and norm.

---

* The space of bounded linear operators from $C_o(R)$ into $C_o(R)$.

Using (5,11), we have, for $f \in \mathcal{D}$

$$D U_n (t) f - D U_m (t) f = \int_0^t D T (t-s) (u_n(\cdot) - u_m(\cdot)) D U_n (s) f ds \tag{5.13}$$
$$+ \int_0^t D T (t-s) u_m(\cdot) (D U_n (s) f) ds .$$

Consider the integrand of the first term on the right-hand side of (5.13). Let

$$\phi^{t,s}_{n,m}(x) \equiv \int_R | u_n(y) - u_m(y) | \frac{1}{\sqrt{2\pi(t-s)}} \exp\left\{- \frac{(y-x)^2}{t-s}\right\} dy , \; x \in R . \quad \text{Then}$$

$$| D T(t-s) (u_n(x) - u_m(x)) D U_n (s) f (x) | \le \int_R | u_n(y) - u_m(y) | \, | DU_n(s) f(y) | \, \left| \frac{\partial}{\partial x} p(t-s,x,y) \right| d \tag{5.14}$$
$$\le \| f \| \lambda(s) \psi(t-s) \phi^{t,s}_{n,m}(x) .$$

By definition of u_n and the fact that $| u (x) | \to 0$ as $x \to \pm \infty$, there exists $k > 0$ such that $| u_n(x) | < \frac{\epsilon}{8}$ for all $n \ge$ some N if $| x | > k$. Therefore

$$\phi^{t,s}_{n,m}(x) \le \frac{\epsilon}{2} + \int_{-k}^{k} | u_n(y) - u_m(y) | \frac{\exp}{\sqrt{2\pi(t-s)}} \left\{ - \frac{(y-x)^2}{t-s} \right\} dy$$
$$\le \frac{\epsilon}{2} + \frac{1}{\sqrt{2\pi(t-s)}} \int_{-k}^{k} | u_n(y) - u_m(y) | \, dy \le \epsilon$$

for sufficiently large n, m. Hence, $\| \phi^{t,s} \| \to 0$ as $n, m \to \infty$ for $0 \le s < t \in I$.

Returning to (5.13),

$$\| D U_n(t) f - D U_m(t) f \| \le \| f \| \int_0^t \lambda(s) \psi(t-s) \| \phi^{t,s}_{n,m} \| ds \tag{5.15}$$
$$+ \int_0^t \psi(t-s) \| D U_n(s) f - D U_m(s) f \| ds .$$

Now, for any $t \in I$, for almost every $s \in [0, t]$, $\| \phi^{t,s}_{n,m} \| \to 0$ as $n, m \to \infty$. Then, since $\lambda(s) \psi(t-s) \| \phi^{t,s}_{n,m} \| \le 2 \nu \lambda(s) \psi(t-s)$, which is integrable on $[0, t]$, the first term on the right-hand side of (5.13) approaches 0 as $n, m \to \infty$. Given any $\beta > 0$, we can choose n, m large enough so that for all $t \in I$ (choose n, m for $t = t_f$)

$$\| D U_n(t) - D U_m(t) \| \leq \beta + \int_0^t \psi(t-s) \| D U_n(s) - D U_m(s) \| \, ds .$$

Using Laplace Transforms, we can show that

$$\| D U_n(t) - D U_m(t) \| \leq \beta \left(1 + \int_0^t h(t-s) \, ds\right),$$

where $h(t)$ is the inverse Laplace Transform of $\dfrac{N}{\sqrt{s} - N}$, $N = M \pi \sqrt{\dfrac{2}{\alpha}}$.

It can be readily shown that $h(t)$ is an integrable function over finite intervals.

Hence, $\lim_{n,m \to \infty} \| D U_n(t) - D U_m(t) \| = 0$, uniformly on I. Since $\mathcal{L}(C_o(R), C_o(R))$ is a Banach space, there exists $V(t) \in \mathcal{L}(C_o(R), C_o(R))$ such that

$$\lim_{n \to \infty} \| D U_n(t) - V(t) \| = 0$$

uniformly on I.

Q.E.D.

<u>Theorem 5.4</u> (Existence of Unique Quasi-Diffusions)

The semi-group of operators $\{ U(t) : t \geq 0 \}$ defined by (5.1) can be represented as

$$U(t) f = T(t) f + \int_0^t T(t-s) \, u(\cdot) \, D U(s) f \, ds, \qquad t \geq 0, \tag{5.16}$$

for $f \in \mathfrak{D}$, and $\{ U(t) : t \geq 0 \}$ is unique.

Proof: For $f \in \mathcal{D}$, $t \geq 0$, $n, m > 0$, (5.7) yields

$$U_n(t)f - U_m(t)f = \int_0^t T(t-s)(u_n(\cdot) - u_m(\cdot))\, D U_n(s) f \, ds$$

$$+ \int_0^t T(t-s)\, u_m(\cdot)\,(D U_n(s) f - D U_m(s) f)\, ds \, .$$

Since $\{ T(t) : t \geq 0 \}$ is induced by a conservative diffusion process, $\| T(t) \| = 1$ for all $t \geq 0$. Then, exactly as in Lemma 5.3, we can show that

$$\lim_{n,\, m \to \infty} \| U_n(t) - U_m(t) \| = 0 \, ,$$

where we use the fact that $\| D U_n(s) \|$ is uniformly integrable with respect to $n \geq 0$ over $[0, t]$, and $\| D U_n(s) - D U_m(s) \| \to 0$ as $n, m \to \infty$. Hence, $\{ U_n(t) \}_{n \geq 0}$ converges in the uniform operator topology to $U(t) \in \mathcal{L}(C_o(R), C_o(R))$ as $n \to \infty$, i.e.,

$$\lim_{n \to \infty} \| U_n(t) - U(t) \| = 0 \quad .$$

We now show that for $f \in \mathcal{D}$, $V(t) f = D U(t) f$ for any $t \geq 0$. Since $D U_n(t) f(x)$ is a continuously differentiable function of x, and $D U_n(t) f \xrightarrow{s} V(t) f$ by Lemma 5.3,

$$U_n(t) f(x) - U_n(t) f(y) = \int_y^x D U_n(t) f(\xi)\, d\xi$$

$$\to \int_y^x V(t) f(\xi)\, d\xi \qquad \text{as } n \to \infty \, ,$$

where $\| D U_n(t) f \| \leq \lambda(t) \| f \|$ for all $n > 0$. Also, $U_n(t) f \xrightarrow{s} U(t) f$ as $n \to \infty$. Thus,

$$U(t) f(x) - U(t) f(y) = \int_y^x V(t) f(\xi)\, d\xi,$$

which implies that for $f \in$

$$D U(t) f = V(t) f .$$

For the remainder of this proof, let $t \geq 0$ be fixed. We claim that for $0 \leq s \leq t$ and $f \in$

$$q_n(s) \equiv T(t-s) u_n(\cdot) D U_n(s) f \xrightarrow{s} T(t-s) u(\cdot) D U(s) f \equiv q(s)$$

as $n \to \infty$. From Lemma 5.3, we know that

$$g_n(s, \cdot) \equiv D U_n(s) f(\cdot) \xrightarrow{s} D U(s) f(\cdot) \equiv g(s, \cdot) \tag{5.17}$$

as $n \to \infty$. Now,

$$| T(t-s) (u_n(x) g_n(s, x) - u(x) g(s, x))| \leq \int_R | u_n(y) - u(y)| \, | g_n(s, y)| \, p(t-s, x, y)\, dy$$

$$+ \int_R u(y) \, | g_n(s, y) - g(s, y) | \, p(t-s, x, y)\, dy$$

$$\leq \| f \| \lambda(s) \int_R | u_n(y) - u(y) | \frac{M}{\sqrt{t-s}} \exp\left\{ - \frac{a (y-x)^2}{t-s} \right\} dy$$

$$+ \gamma \| g_n(s, \cdot) - g(s, \cdot) \| \tag{5.18}$$

where the estimate for $p(t, x, y)$ is obtained from [8, Vol. II, Theorem 0.5]. The first term in the right-hand side of (5.18) goes to 0 uniformly in x as $n \to \infty$ by the a.e. convergence of $u_n(y)$ to $u(y)$ and by the same argument as in the proof of Lemma 5.3. The second term goes to 0, as $n \to \infty$, by (5.17). Hence,

$$q_n(s) \xrightarrow{s} q(s) \tag{5.19}$$

as $n \to \infty$, for each $0 \leq s \leq t$.

Let $\mu \in \mathfrak{M}(R)$ be arbitrary. Then,

$$\mu(q_n(s)) = \int_R \int_R u_n(y)\, D\, U_n(s)\, f(y)\, p(t-s, x, y)\, dy\ \mu(dx)$$

is a continuous function from $[0, t]$ into R, since $DU_n(\tau) f(y)$ and $p(\tau, x, y)$ are continuous in τ, by virtue of the fact that the coefficients which generate $\{U_n(t) : t \geq 0\}$ and $\{T(t) : t \geq 0\}$ satisfy the conditions of [8, Vol. II, Theorem 0.5], ensuring the existence of fundamental solutions. Thus, the function $q_n(\cdot) : [0, t] \to C_o(R)$ is weakly measurable for each $n > 0$. Since $C_o(R)$ is a separable Banach space, the theorem in [6, p. 131] implies that $q_n(s)$ is strongly measurable in the Bochner sense [6, p. 130]*. Therefore, since $\| q_n(s) \| \leq \gamma \lambda(s)$ for all $n > 0$ and $\lambda(s)$ is integrable over $[0, t]$, the theorem in [6, p. 133] implies that $\{q_n(s)\}_{n \geq 0} \subset B([0, t], C_o(R))$, the space of Bochner integrable functions from $[0, t]$, with the Lebesgue measure, into the separable Banach space $C_o(R)$. Since $q_n(s) \xrightarrow{s} q(s)$ as $n \to \infty$ for each $s \in [0, t]$, and $\| q_n(s) \| \leq \gamma\lambda(s)$ for all $n > 0$, [3, Theorem 3.7.9, p. 83] implies that $q(s) \in B([0, t], C_o(R))$ and

$$\lim_{n \to \infty} \int_0^t q_n(s)\, ds = \int_0^t q(s)\, ds .$$

Thus, letting $n \to \infty$ in (5.7), we get for $f \in \mathfrak{D}$

$$U(t) f = T(t) f + \int_0^t T(t-s)\, u(\cdot)\, D\, U(s) f\, ds . \tag{5.20}$$

---

* Recall that every subset of a separable metric space is separable [16, Theorem 7.3, p. 176].

It remains only to show the uniqueness of $\{U(t) : t \geq 0\}$. We shall need the following result: for $f \in \mathfrak{D}$, $DU(t)f$ is s-continuous on $[0, t_f]$. In order to prove this, we first show that for $n > 0$, $DU_n(t)f$ in s-continuous on $[0, t_f]$. For $t \in (0, t_f]$,

$$D U_n(t+\Delta)f - D U_n(t)f = D T(t+\Delta)f - D T(t)f$$

$$+ \int_0^{t+\Delta} D T(t+\Delta-s)\, u_n(\cdot)\, D U_n(s)\, f\, ds - \int_0^t D T(t-s)\, u_n(\cdot)\, D U_n(s) f\, ds,$$

and

$$\| D U_n(t+\Delta) f - D U_n(t) f \| \leq \| D T(t) \| \, \| T(\Delta) f - f \|$$

$$+ \int_0^t \| D T(t-s) \| \, \| (T(\Delta) - I)\, u_n(\cdot)\, D U_n(s)\, f \| \, ds$$

$$+ \int_t^{t+\Delta} \gamma \| D T(t+\Delta-s) \| \, \| D U_n(s) f \| \, ds .$$

The first term on the right side goes to 0 as $\Delta \to 0$ by the s-continuity of $\{T(t): t \geq 0\}$ on $C_o(R)$ [8, Vol. I, Equation (5.69), p. 163]. Since $u_n(x)$ is a continuous function, $u_n(\cdot)\, D U_n(s)\, f \in C_o(R)$ for all $s \in I$, and the integrand $\| D T(t-s) \| \, \|(T(\Delta) - I)\, u_n(\cdot)$ $D U_n(s)\, f \|$ is uniformly bounded by $2 \gamma\, \psi(t-s)\, \lambda(s) \| f \|$, which is integrable on $[0, t]$. Therefore, the second term goes to 0 as $\Delta \to 0$. The third term approaches 0 as $\Delta \to 0$ since $\psi(t+\Delta-s)\, \lambda(s)$ is integrable. Therefore, $D U_n(t) f$ is s-continuous on $(0, t_f]$. To prove continuity at the origin, we need only show that $\| D T(\Delta) f - D f \| \to 0$ as $\Delta \to 0$. This follows from

$$\overline{\lim_{x \to \infty}} \sup_{0 \leq \tau < t_f} | D T(\tau) f(x) | = 0 ,$$

which can be proved by an argument similar to 2° of [8, Vol. 1, p. 163] using the estimate (0.41) of [8, Vol. II, Theorem 0.5] in place of the estimate (0.40), and the fact that $D\,T(\Delta)\,f(x) \to f(x)$ as $\Delta \to 0$ uniformly on compact sets [8, Vol. 1, bottom of p. 164 and top of p. 165]. Thus, for each $n > 0$, $D\,U_n(t)f$ is s-continuous on $[0, t_f]$.

Now, Lemma 5.3 shows that $D\,U_n(t)f \overset{s}{\to} D\,U(t)f$ as $n \to \infty$ uniformly on $[0, t_f]$. Hence, $D\,U(t)f$ is s-continuous on $[0, t_f]$ by a standard argument, which implies that

$$\sup \left\{ \| D\,U(t)f \| : t \in [0, t_f] \right\} < \infty \quad . \tag{5.21}$$

Returning to the uniqueness proof, let $\{U(t)f : t \geq 0\}$ and $\{Z(t)f : t \geq 0\}$ both satisfy (5.20). Then $W(t)f \equiv U(t)f - Z(t)f$ satisfies

$$W(t)f = \int_0^t T(t-s)\; u(\cdot)\; D\,W(s)\,f\; ds \quad . \tag{5.22}$$

For $x \in R$,

$$D\,W(t)\,f(x) = \int_0^t \int_R u(y)\,D\,W(s)\,f(y)\,\frac{\partial p(t-s, x, y)}{\partial x}\,dy\,ds \quad ,$$

and

$$|D\,W(t)\,f(x)| \leq \gamma \int_0^t \|D\,W(s)\,f\| \left\{ \int_R \frac{M}{(t-s)} \exp\left\{ -\alpha \frac{(y-x)^2}{t-s} \right\} dy \right\} ds \quad .$$

Thus,

$$\|D\,W(t)\,f\| \leq \gamma \int_0^t \psi(t-s)\,\|D\,W(s)\,f\|\,ds \quad .$$

Let $c_t = \sup \left\{ \| D\, W(\tau) f \| : 0 \le \tau \le t \right\}$, which is finite by (5.21). Then

$$c_t \le c_t \, \gamma \int_0^t \psi(s)\, ds .$$

Since $\psi(s)$ is integrable over any interval $[0, t]$, we can find $t' > 0$ small enough so that

$$\gamma \int_0^{t'} \psi(s)\, ds < 1 .$$

Hence, $c_{t'} = 0$. Since there is nothing special about the origin in this argument, it follows that $D\, W(t) f = 0$ for all $t \ge 0$. Substituting this into (5.22) we get $W(t) f = 0$ for all $t \ge 0$. (The uniqueness proved here is with respect to all competitors which satisfy (5.21).)

Q.E.D.

Corollary 5.5

If $A\, T(t-s)\, u(\cdot)\, D\, U(s) f \in B([0, t], C_o(R))$ for each $t \ge 0$, then for $f \in \mathcal{D}$, $U(t)f$ satisfies

$$\frac{\partial U(t) f}{\partial t} = (A + u(\cdot) D)\, U(t) f \qquad \text{a.e. on } [0, \infty) .$$

Proof: Using (5.20), for $t \ge 0$ fixed,

$$\frac{U(t+\Delta)f - U(t)f}{\Delta} - \frac{T(t+\Delta)f - T(t)f}{\Delta} = \int_0^t \frac{(T(\Delta) - I)}{\Delta} T(t-s)\, u(\cdot)\, D\, U(s)\, f\, ds$$
$$+ \frac{1}{\Delta} \int_t^{t+\Delta} T(t + \Delta - s)\, u(\cdot)\, D\, U(s) f\, ds . \tag{5.23}$$

It was shown earlier that $T(t-s)\,u(\cdot)\,D\,U(s)f \in B([0, t], C_o(R))$. Therefore, [3, Corollary 2, p. 88] implies that the second term on the right-hand side of (5.23) goes to $u(\cdot)\,D\,U(t)f$ a.e. on $[0, \infty)$ as $\Delta \to 0$. For all $0 < \Delta \le t$,

$$\left\|\left(\frac{T(\Delta - I)}{\Delta}\right) q(s)\right\| = \left\| \frac{1}{\Delta} \int_0^\Delta T(\tau)\,A\,q(s)\,d\tau \right\|$$

$$\le \frac{1}{\Delta} \int_0^\Delta \|A\,q(s)\|\,d\tau = \|A\,q(s)\| \in B([0, t], C_o(R)),$$

where $q(s) \equiv T(t-s)\,u(\cdot)\,D\,U(s)f$. Therefore, the first term on the right-side of (5.23) goes to $\int_0^t A\,T(t-s)\,u(\cdot)\,D\,U(s)\,f\,ds$ as $\Delta \to 0$. Since $f \in \mathfrak{D}$ and $T(t)\mathfrak{D} \subset \mathfrak{D}(A)$, the second term on the left side of (5.23) goes to $A\,T(t)f$ as $\Delta \to 0$. Thus, $\dfrac{\partial U(t)f}{\partial f}$ exists a.e. on $[0, \infty)$, and

$$\frac{\partial U(t)f}{\partial t} = A\,T(t)f + A \int_0^t T(t-s)\,u(\cdot)\,D\,U(s)f\,ds + u(\cdot)\,D\,U(t)\,f$$

$$= (A + u(\cdot)\,D)\,U(t)\,f \qquad \text{a.e. on } [0, \infty).$$

Q.E.D.

The method for integration employed in this section seems to be restricted to unbounded operators of the form

$$A^u = (a(x) + u(x))\,D + \frac{b^2(x)}{2}\,D^2 ,$$

where $D^2 \equiv \dfrac{\partial^2}{\partial x^2}$. The critical point is the integrability of $\|D\,T(t)\|$ over bounded intervals $[0, \tau]$. If we consider a perturbation term in the diffusion coefficient as well as in the drift coefficient, the operator A^u is of the form

$$A + u_1(x)\, D + u_2(x)\, D^2 .$$

To use the techniques of this section, we would require the integrability of $\| D^2 T(t) \|$ over $[0, \tau]$. This, however, is not implied by the estimates in [8, Vol. II, Theorem 0.5]. In fact, we can show that

$$\int_0^{\tau} \| D^2 T(t) \| \, dt = \infty \tag{5.24}$$

for any $0 < \tau < \infty$. To see this, let

$$\tilde{A} = D + D^2 ,$$

which is certainly an infinitesimal generator, and let $\{ \tilde{T}(t) : t \geq 0 \}$ be the semi-group generated by $\tilde{A}$. Now, suppose that

$$\int_0^{\tau} \| D^2 \tilde{T}(t) \| \, dt < \infty$$

for any $0 < \tau < \infty$. Then, by the integrability of $\| D \tilde{T}(t) \|$ and Proposition 2.2(b),

$$\| \tilde{T}(t) - I \| \leq \int_0^t \| \tilde{A} \tilde{T}(s) \| \, ds \, *$$

exists, and goes to 0 with t. This shows that $\{ \tilde{T}(t) : t \geq 0 \}$ is uniformly continuous, which implies that $\tilde{A}$ is a bounded linear operator [1, Theorem VIII.1.2, p. 614]. But, this is impossible since $\tilde{A}$ is a differential operator on the Banach space $C_o(R)$.

<u>Remarks:</u>

(i) For $u(t, x)$ bounded, and integrable as a function of x for each t, the methods of this section can be used to associate a unique two-parameter

---

* $\tilde{A}\, \tilde{T}(s)$ is actually defined only on the domain of A; the $\tilde{A}\, \tilde{T}(s)$ in this relation is the natural extension to the entire space $C_o(R)$.

family $\{U(s, t) : 0 \leq s \leq t,\ t \in I\}$ with $[a(x) + u(t, x), b(x)]$.

For $f \in \mathcal{D}$, $U(s, t) f$ is represented as

$$U(s, t) f = T(t - s) f + \int_s^t T(t - \tau)\, u(\tau, \cdot)\, D\, U(s, \tau)\, f\, ds\ . \tag{5.25}$$

The extension is accomplished by applying [8, Vol. II, Theorem 0.4] in place of [8, Vol. II, Theorem 0.5], to ensure the existence of fundamental solutions of

$$\frac{\partial U_n(s, t) f}{\partial t} = A_n(t)\, U_n(s, t) f\ ,$$

where

$$A_n(t) = A + u_n(t, x) D\ ,$$

and $\{U_n(s, t) : t \geq s \geq 0\}$ is generated by $A_n(t)$; $u_n(t, x)$ is continuous and bounded on $[0, t_f] \times R$, uniformly Hölder continuous in x for all t, and $u_n(t, x) \rightarrow u(t, x)$ a.e. on $[0, t_f] \times R$.

(ii) Since $u(x)$ is bounded and integrable, we can assume that $a(x)$ has the same properties, and that the semi-group $\{T(t) : t \geq 0\}$ is generated by $\frac{b^2(x)}{2} \frac{\partial^2}{\partial x^2}$.

5.3 Application to Optimal Stochastic Control Theory

Let $\mathcal{U}$ be the set of control functions $u(x)$ which are bounded and integrable over R. Suppose we are given a cost function $V \in \mathfrak{D}$, and we wish to find conditions which will ensure that, at a certain time $t_1 > 0$, $E_x^{P^u}\left\{ V(\pi_{t_1}) \right\}$ is minimized, i.e., we want to find a condition such that for some $\bar{u} \in \mathcal{U}$

$$\eta^{\bar{u}}(t_1)\,\varphi(V) \;\le\; \eta^{u}(t_1)\,\varphi(V) \quad , \qquad \varphi \in \mathfrak{M}_1(R),$$

for all $u \in \mathcal{U}$. In the light of Theorem 5.4, for each $u \in \mathcal{U}$, the flow $\{\eta^u(t)\,\varphi : t \ge 0\} \subset \mathfrak{M}_1(R)$ is unique. Integrating both sides of (5.20) with respect to φ, we get for $f \in \mathfrak{D}$

$$\varphi(U^u(t)f) \;=\; \varphi(T(t)f) + \int_0^t \varphi(T(t-s)\,u(\cdot)\,D\,U^u(s)f)\;ds\,.$$

Let A^* be the adjoint operator of A with domain $\mathfrak{D}(A^*)$. (See Remark (v).) If $\varphi \in \mathfrak{M}_1(R) \cap \mathfrak{D}(A^*)$, then exactly as in Proposition 2.5, it can be shown that

$$\eta(t)\,\varphi(f) \;=\; m(t)\,\varphi(f) + \int_0^t \eta(s)\,D^*\,u(\cdot)\,m(t-s)\,\varphi(f)\;ds$$

for any $f \in C_o(R)$, where $\eta(t) \equiv U^u(t)^*$, $m(t) \equiv T^*(t)$ and $D^* \equiv -\frac{\partial}{\partial x}$.

Now, a necessary and sufficient condition for $\bar{u} \in \mathcal{U}$ to minimize $\eta^u(t_1)\,\varphi(V)$, $\varphi \in \mathfrak{M}_1(R)$, is that

$$\int_0^{t_1} \varphi(T(t_1-s)\,\bar{u}(\cdot)\,D\,U^{\bar{u}}(s)\,V)\,ds \;\le\; \int_0^{t_1} \varphi(T(t_1-s)\,u(\cdot)\,D\,U^u(s)\,V)\,ds \tag{5.26}$$

for all $u \in \mathcal{U}$. If $V \in C_o(R)$ and $\varphi \in \mathfrak{M}_1(R) \cap \mathfrak{D}(A^*)$ then a necessary and sufficient condition for $\bar{u} \in \mathcal{U}$ to be optimal is that

$$\int_0^{t_1} \eta^{\bar{u}}(s)\, D^* \bar{u}(\cdot)\, m(t-s)\, \varphi(V)\, ds \;\le\; \int_0^{t_1} \eta^{u}(s)\, D^* u(\cdot)\, m(t-s)\, \varphi(V)\, ds$$

for all $u \in \mathcal{U}$.

Remarks:

(i) If we wish to utilize the theory in [5], we can obtain a stronger version of (5.26) for $V \in C_0^\infty(R)$. Let $a(t, x)$ and $u(t, x)$ be bounded measurable on $[0, \infty) \times R$, and let $b(t, x)$ be bounded continuous on $[0, \infty) \times R$ and strictly elliptic. Then, there exists a unique family $\{ U^u(s, t) : 0 \le s \le t \}$ associated with $[a(t, x) + u(t, x)]$, and

$$U^u(t)\, f = f + \int_0^t U^u(s)\, A^u(s)\, f\, ds \quad , \tag{5.27}$$

where $U^u(t) \equiv U^u(0, t)$, $f \in C_0^\infty(R)$, and

$$A^u(s) = (a(s, x) + u(s, x)) \frac{\partial}{\partial x} + \frac{b^2(s, x)}{2} \frac{\partial^2}{\partial x^2} \quad .$$

Let $\tilde{\mathcal{U}}$ be the class of bounded measurable controls from $[0, \infty) \times R$ into R. Then, by (5.27), $\bar{u} \in \tilde{\mathcal{U}}$ minimizes the average of $V \in C_0^\infty(R)$, at time $t_1 > 0$, if and only if

$$\int_0^{t_1} \varphi(U^{\bar{u}}(s)\, A^{\bar{u}}(s)\, V)\, ds \;\le\; \int_0^{t_1} \varphi(U^u(s)\, A^u(s)\, V)\, ds$$

for all $u \in \tilde{\mathcal{U}}$, where $\varphi \in \mathcal{M}_1(R)$.

(ii) A sufficient condition for optimality in (5.26) is that for each $s \in [0, t_1]$

$$\varphi(T(t_1 - s)\, \bar{u}(\cdot)\, D\, U^{\bar{u}}(s)\, V) \;\le\; \varphi(T(t_1 - s)\, u(\cdot)\, D\, U^u(s)\, V)$$

for all $u \in \mathcal{U}$.

(iii) To further stress the significance of Theorem 5.4, and the related results in [5], we will suggest a formulation of an optimal stochastic control problem and conjecture a method of solution. Let $s = 0$, $U(t) = U(0, t)$, and $u(t) = u(t, x)$ be a bounded function of t only, in (5.26). Then, for $V \in \mathfrak{D}$, in view of Theorem 5.4 and Remark (i) of Section 5.2,

$$\varphi(U(t)V) = \varphi(T(t)V) + \int_0^t \varphi(T(t-s)\,u(s)\,D\,U(s)\,V)\,ds \tag{5.28}$$

uniquely defines the trajectory of the average of the random variable $V(x_t)$, where x_t starts at $t = 0$ with probability measure φ. Letting $y(t) = \varphi(U(t)V)$, $f(t) = \varphi(T(t)V)$ and $K(t, s, u(s), V) = \varphi(T(t-s)\,u(s)\,D\,U(s)\,V)$, (5.28) is rewritten as

$$y(t) = f(t) + \int_0^t K(t, s, u(s), V)\,ds, \tag{5.29}$$

where K is related to $y(t)$ in an implicit manner. We now consider the cost functional

$$J(y, u) = \int_0^{t_1} K_o(y(s), u(s), s)\,ds,$$

where K_o satisfies certain continuity and differentiability conditions. The problem is to find an admissible control $u(t)$ such that $J(y, u)$ is minimized. As formulated, this optimal stochastic control problem resembles the problems studied in [52]. Unfortunately, the integrand in (5.28) is considerably more complicated than that in Equation (1.1) of [52, Part I]; the difficulty is due to the unbounded operator D which affects K in a complex manner, and the fact that $y(t)$ does not appear explicitly. However, the methods employed in [52] may prove useful in deducing a maximum principle for the optimal stochastic control problem formulated above.

(iv) We shall now discuss another type of optimal stochastic control problem. Given the final time t_1, we are interested in minimizing some functional of the final probability measure induced by the random variable $x^u_{t_1}(\omega)$, for the open-loop control system (2.4). To be more specific, if φ is the initial probability measure, we wish to find the control u which minimizes

$$f_o(\eta^u(t_1)\varphi) ,$$

where f_o is a continuous map from $\mathcal{M}_1(R)$ with its relative norm topology, $(\mathcal{M}_1(R), \| \ \|)$, into $[0, \infty)$.

Consider the following space of control functions, which is used in [53]: Let $L_p\{R, [0, t_1]\}$ be the space of measurable functions $u(t)$ with range in R such that

$$\int_0^{t_1} | u(t) |^p \, dt < \infty$$

for some $1 < p < \infty$. Let

$$\Psi = \left\{ u \in L_p\{R, [0, t_1]\} , \sup_{0 \le t \le t_1} | u(t) | \le M < \infty \right\} .$$

Then, Ψ is a closed bounded convex set of $L_p\{R, [0, t_1]\}$ which is reflexive. Define the mapping

$$f(u) = \int_0^{t_1} \left(\eta^u(s)\, D^*\, u(s)\, m(t-s)\, \varphi \right)(\cdot)\ ds$$

from $L_p \{ R, [0, t_1] \}$ into $(\mathfrak{M}(R),\ \|\cdot\|)$, where $\varphi \in \mathfrak{M}_1(R) \cap \mathfrak{D}(A^*)$. Then,

$$\eta^u(t_1)\,\varphi = m(t_1)\,\varphi + f(u) .$$

Let

$$f_o(u) = \| m(t_1)\,\varphi + f(u) - \psi \| ,$$

where $\psi \in \mathfrak{M}_1(R)$ is a target measure. Now, if $f(\Psi)$ is a convex set in $\mathfrak{M}(R)$ and $f_o(u)$ is a continuous convex functional on Ψ, then [53, Theorem 2.1] shows that there exists an optimal control $\bar{u} \in \Psi$ which minimizes f_o. The difficulty in the above formulation is that, in general $f(\Omega)$ is not convex, and f_o is not a convex function of u. (Observe that f is not a linear operator on u.)

Actually, we would prefer treating this problem in the metric space $\mathfrak{M}_1(R)$ with its weak topology. That is, we wish to know if there exists a $\bar{u} \in \Psi$, or in some other control class, such that

$$f_1(u) \equiv \rho(\eta^u(t_1)\,\varphi,\ \psi)$$

is minimized, where ρ is the metric of weak convergence and $\varphi \in \mathfrak{M}_1(R)$. This problem is very difficult to handle, because there are no techniques available (to the author's knolwedge) for optimization in a general metric space. The most general optimization theory seems to require at least a linear space.

(v) From Proposition 2.5, we know that $\mathfrak{D}(A^*)$ is weak*dense $\big(\sigma(\mathcal{M}(R), C_o(R))$ topology$\big)$ in the space $\mathcal{M}(R)$, i.e., given any $\mu \in \mathcal{M}(R)$, we can find $\{\mu_n\} \subset \mathfrak{D}(A^*)$ such that $\lim_{n\to\infty} \mu_n(f) = \mu(f)$ for all $f \in C_o(R)$. We now wish to characterize $\mathfrak{D}(A^*)$ more completely. We shall consider the situation where $a(x)$ and $b(x)$ are twice continuously differentiable on R. Let $\Phi(R)$ be the funadmental space consisting of infinitely differentiable functions on R with compact support. Define the linear operator A on $\Phi(R)$ by

$$A = a(x)\frac{\partial}{\partial x} + h(x)\frac{\partial^2}{\partial x^2} ,$$

where $h(x) = \dfrac{b^2(x)}{2}$. Let $\mu \in \mathcal{M}(R)$. Then μ defines a distribution T_μ on $\Phi(R)$ through

$$T_\mu(\varphi) = \int_{-\infty}^{\infty} \varphi(x)\ \mu(dx) .$$

Integrating by parts, we get

$$\mu(A\varphi) = A^*\mu(\varphi) ,$$

where

$$A^*\mu = -\frac{\partial}{\partial x}(a\mu) + \frac{\partial^2}{\partial x^2}(h\mu) .$$

Since $\mu \in \mathcal{M}(R)$, and a, h are continuous, $a\mu$, $h\mu \in \mathcal{M}(R)^*$. This means that $a\mu$, $h\mu \in \Phi'(R)$, the space of distributions on R. By Leibniz's formula,

* See [17, Section 13].

$$\frac{\partial a\mu}{\partial x} = \frac{\partial a}{\partial x}\mu + a\frac{\partial \mu}{\partial x} \quad , \tag{5.30}$$

and

$$\frac{\partial^2 h\mu}{\partial x^2} = \left(\frac{\partial^2 h}{\partial x^2}\right)\mu + 2\frac{\partial h}{\partial x}\frac{\partial \mu}{\partial x} + h\cdot\frac{\partial^2 \mu}{\partial x^2} \quad . \tag{5.31}$$

We define

$$\mathfrak{D}(A^*) = \left\{ \mu \in \mathfrak{M}(R) \ : \ A^*\mu \in \mathfrak{M}(R) \right\} \quad .$$

Since $\frac{\partial a}{\partial x}$, $\frac{\partial h}{\partial x}$, $\frac{\partial^2 h}{\partial x^2}$ are continuous, a necessary and sufficient condition that $A^*\mu \in \mathfrak{M}(R)$ is that $\frac{\partial \mu}{\partial x}$, $\frac{\partial^2 \mu}{\partial x^2} \in \mathfrak{M}(R)$, for then all the terms of (5.30) and (5.31) are in $\mathfrak{M}(R)$, since $\mathfrak{M}(R)$ is a linear space. Thus,

$$\mathfrak{D}(A^*) = \left\{ \mu \in \mathfrak{M}(R) : \frac{\partial \mu}{\partial x}, \frac{\partial^2 \mu}{\partial x^2} \in \mathfrak{M}(R) \right\} \quad .$$

In other words, $\mu \in \mathfrak{D}(A^*)$ if $\frac{\partial \mu}{\partial x}$ and $\frac{\partial^2 \mu}{\partial x^2}$ are functions of bounded variation on R, equal to 0 at $-\infty$, and finite at $+\infty$.

CHAPTER VI

DYNAMICAL SYSTEMS IN THE SPACE OF MEASURES

6.1 Examples

To motivate the work that follows, we present two examples.

Example I

Let $E = R^n$ and let x_t be a Feller process taking values in E with transition density function $p(t, x, y)$ such that for all $f \in C(E)$*,

$$\lim_{t \downarrow 0} \int_E f(y)\, p(t, x, y)\, dy = f(x) \text{ uniformly on compacts.} \tag{6.1}$$

By [8, Vol. I, Theorem 5.11], diffusions whose coefficients are bounded uniformly Hölder continuous, and the diffusion coefficient strictly elliptic, satisfy the above conditions.**

Let φ be an initial probability measure and define for $\Gamma \in \mathcal{B}(E)$, the Borel σ-algebra generated by the open sets of E, $t \geq 0$,

$$m(t, \varphi)(\Gamma) \equiv \int_E \int_\Gamma p(t, x, y)\, dy\, \varphi(dx). \tag{6.2}$$

From (6.2) we have that

$$m(0, \varphi) = \varphi \quad ,$$

and by the Chapman-Kolmogorov Equation,

$$m(t+s, \varphi) = m(t, m(s, \varphi)) \qquad t, s \geq 0 \quad .$$

---

* $C(E)$ is the Banach space of real bounded continuous functions on E with the supremum norm.

** This is also true if the coefficients are Lipschitz continuous (see [8, Chapter II]).

We claim that the map $(t, \varphi) \longrightarrow m(t, \varphi)$ from $R^+ \times \mathcal{M}_1(E)$ into $\mathcal{M}_1(E)$ is continuous where $R^+ = [0, \infty)$ and $\mathcal{M}_1(E)$ has the weak topology: Let $(t + \frac{1}{n}, \varphi_n) \to (t, \varphi)$ as $n \to \infty$; then, for any $f \in C(E)$,

$$m(t + \frac{1}{n}, \varphi_n)(f) - m(t, \varphi)(f) = \int_E \int_E f(y)\, p(t + \frac{1}{n}, x, y)\, dy\ \varphi_n(dx) - \int_E \int_E f(y)\, p(t, x, y)\, dy\ \varphi(dx) \tag{6.3}$$

$$= \int_E \int_E f(y) \Big(p(t + \frac{1}{n}, x, y) - p(t, x, y) \Big) dy\ \varphi_n(dx) + \int_E T_t f(x) \left(\varphi_n(dx) - \varphi(dx) \right) , \tag{6.4}$$

where $T_t f(x) = \int_E f(y)\, p(t, x, y)\, dy$. The first term on the right-hand side of (6.4) goes to 0 as $n \to \infty$ by (6.1) and the Dominated Convergence Theorem, while the second terms approaches 0 in virtue of the fact that $T_t f \in C(E)$ and $\varphi_n \Longrightarrow \varphi$ as $n \to \infty$. Hence, $(t, \varphi) \to m(t, \varphi)$ is continuous.

To summarize, we have shown that for $\varphi \in \mathcal{M}_1(E)$,

(1) $m(0, \varphi) = \varphi$

(2) $m(t + s, \varphi) = m(t, m(s, \varphi))$ for all $s, t \geq 0$ (6.5)

(3) the map $R^+ \times \mathcal{M}_1(E) \to \mathcal{M}_1(E)$ defined by $(t, \varphi) \to m(t, \varphi)$ is continuous, where $\mathcal{M}_1(E)$ has the weak topology.

Example II

Let E be a metric space.

<u>Definition 6.1.</u> A map $\nu : R^+ \times E \to E$ is said to define a (positive) <u>dynamical system</u> (E, R^+, ν), or continuous flow, on E if it has the following properties:

(1) $\nu(0, x) = x$ for all $x \in E$

(2) $\nu(t, \nu(s, x)) = \nu(t+s, x)$ for all $x \in E$, $t, s \geq 0$

(3) ν is continuous .

We define probability measures on E in the following manner: for $x \in E$,

$$\varphi(\Gamma) = \begin{cases} 1 & \text{if } x \in \Gamma \\ 0 & \text{otherwise} \end{cases},$$

and

$$m(t, \varphi)(\Gamma) = \begin{cases} 1 & \text{if } \nu(t, x) \in \Gamma \\ 0 & \text{otherwise} \end{cases},$$

where $\Gamma \in \mathcal{B}(E)$. Note that $m(t, \varphi)$ and $\nu(t, x)$ are completely equivalent; specifying one determines the other. The family $\{ m(t, \varphi) : t \geq 0 \}$ has the following properties:

(i) $$m(0, \varphi)(\Gamma) = \begin{cases} 1 & \text{if } x \in \Gamma \\ 0 & \text{otherwise} \end{cases} = \varphi(\Gamma) .$$

Therefore,

$$m(0, \varphi) = \varphi .$$

(ii) $$m(t+s, \varphi)(\Gamma) = \begin{cases} 1 & \text{if } \nu(t+s, x) \in \Gamma \Longleftrightarrow \nu(t, \nu(s, x)) \in \Gamma \\ 0 & \text{otherwise} \end{cases}$$

$$= m(t, m(s, \varphi))\ (\Gamma) .$$

Hence,

$$m(t+s, \varphi) = m(t, m(s, \varphi)) \qquad \text{for } t, s \in R^+ .$$

(iii) Weak continuity of $m(t, \varphi)$ in t and φ follows directly from the continuity of ν .

Therefore, any dynamical system (Definition 6.1) defines flows of probability measures $\{m(t, \varphi) : t \in R^+\}$ which satisfy (6.5).

6.2 Definition of Non-Deterministic Dynamical System

Let E be a complete separable metric space, and $\mathcal{M}(E)$ the space of real signed measures on E . $\mathcal{M}(E)$ is a Banach space with the variation norm. Let $\mathcal{M}_1(E)$ be the set of probability measures on E , and let ρ be the Prohorov metric on $\mathcal{M}_1(E)$ [12]. A sequence $\{\mu_n\} \subset \mathcal{M}_1(E)$ converges to $\mu \in \mathcal{M}_1(E)$ in ρ if and only if $\mu_n(f) \to \mu(f)$ for all $f \in C(E)$, i.e., $\{\mu_n\}$ converges weakly to μ . In this section we shall interpret some of the results of [32, Chapter I] for the metric space $(\mathcal{M}_1(E), \rho)$. We could work in $\mathcal{M}(E)$ or $\mathcal{M}^+(E) \equiv \{\varphi \in \mathcal{M}(E) : \varphi \geq 0\}$ rather than in $\mathcal{M}_1(E)$, but often results in these spaces do not lend themselves to physical interpretation, so we restrict ourselves to the intuitive space $\mathcal{M}_1(E)$.

Definition 6.2. A transformation $m : R^+ \times \mathcal{M}_1(E) \to \mathcal{M}_1(E)$ is said to define a non-deterministic dynamical system (NDDS), or a weak continuous flow, on $\mathcal{M}_1(E)$ if it has the following properties:

(1) $m(0, \varphi) = \varphi$ for all $\varphi \in \mathfrak{M}_1(E)$

(2) $m(s, m(t, \varphi)) = m(t+s, \varphi)$ for all $\varphi \in \mathfrak{M}_1(E)$, $t, s \in R^+$

(3) the map $(t, \varphi) \to m(t, \varphi)$ from $R^+ \times \mathfrak{M}_1(E) \to \mathfrak{M}_1(E)$ is continuous, where $\mathfrak{M}_1(E)$ has the weak topology.

(The usual definition of a dynamical system is for R rather than R^+.)

In the previous section we saw that diffusions, which are special Markov processes, and deterministic dynamical systems are examples of NDDS. (Note that not all semi-groups of probability measures originate from Markov processes [40, p. 340].)

A NDDS is a flow on $\mathfrak{M}_1(E)$. Knowing the probability measure at the present time permits the prediction of the probability measure at any future time. For Markov processes either the transition function or the adjoint of Dynkin's Formula [8, Vol. 1, p. 23] explicitly describes the flow of probability measures.

For every $\varphi \in \mathfrak{M}_1(E)$, the mapping m induces a weak continuous map $m^{\varphi}: R^+ \to \mathfrak{M}_1(E)$ such that $m^{\varphi}(t) = m(t, \varphi)$. The mapping m^{φ} is called the <u>motion</u> of probability measures starting at φ. For every $t \in R^+$, m induces a weak continuous map $m_t : \mathfrak{M}_1(E) \to \mathfrak{M}_1(E)$ such that $m_t(\varphi) = m(t, \varphi)$. The map m_t is called the <u>transition</u>. A NDDS may be visualized as the law with which the probability measure $m(t, \varphi)$ moves along $m(R^+, \varphi) \equiv \{ m(t, \varphi) : t \in R^+ \}$.

6.3 Dynamical Systems in the Space of Measures : Some Results

We commence with the following standard definition:

Definition 6.3. If $\varphi \in \mathcal{M}_1(E)$ has the property that

$$m(t, \varphi) = \varphi$$

for all $t \in R^+$, it is called a stationary (invariant or equilibrium) measure.

In form, the following results are standard, but reveal new information when interpreted for Markov processes.

Proposition 6.1

The set of stationary probability measures of a NDDS is weak closed in $\mathcal{M}_1(E)$.

Proof: Since E is a separable metric space, so is $\mathcal{M}_1(E)$ [13, Lemma 6.3, p. 43]. Thus, we can work with sequences rather than with nets.

We must show that the weak limit of a sequence $\{\varphi_n\} \subset \mathcal{M}_1(E)$, of stationary measures, is itself stationary. From Definition 6.3, $m(t, \varphi_n) = \varphi_n$ for all n, and for all $t \geq 0$. Since m is weak continuous in t and φ, $\varphi_n \Longrightarrow \varphi$ as $n \to \infty$ implies that $m(t, \varphi_n) \Longrightarrow m(t, \varphi)$ as $n \to \infty$. But $m(t, \varphi_n) = \varphi_n$ for all n. Hence, $m(t, \varphi) = \varphi$ for all $t \geq 0$.

Q.E.D.

The proof of the following proposition can be found in [32, p. 15].

Proposition 6.2

If $m([a, b], \varphi) \equiv \{ m(t, \varphi) : t \in [a, b] \} = \varphi$ for $b > a > 0$, $\mathcal{M}_1(E)$, then φ is a stationary probability measure.

A closed ϵ ball of $\varphi \in \mathcal{M}_1(E)$, in the Prohorov metric, is denoted by $S(\varphi, \epsilon) \equiv \left\{ \psi \in \mathcal{M}_1(E) : \rho(\varphi, \psi) \leq \epsilon \right\}$.

Proposition 6.3

If, for every $\epsilon > 0$, there exists at least one $\psi \in S(\varphi, \epsilon)$ such that $m(t, \psi) \subset S(\varphi, \epsilon)$ for all $t \in R^+$, then φ is a stationary probability measure.

Proof: Suppose φ is not a stationary probability measure. Then there exists $\tau > 0$ such that $m(t, \varphi) \neq \varphi$ for $0 \leq t \leq \tau$, otherwise Proposition 6.2 implies that φ is stationary. Let $\epsilon > 0$ be such that $\varphi \notin S(m(t, \varphi), \epsilon)$. By the weak continuity of m, there exists $\delta > 0$ such that $\psi \in S(\varphi, \delta)$ implies $m(\tau, \psi) \in S(m(\tau, \varphi), \epsilon)$. We can also assume that $S(\varphi, \delta) \cap S(m(t, \varphi), \epsilon) = \emptyset$ (empty set). This implies that $m(t, \psi) \notin S(\varphi, \delta)$ for all $t \in R^+$. Thus, if φ is not a stationary measure, then φ has a ρ- neighbourhood which contains no (positive) trajectory of probability measures. This contradicts the hypothesis.

Q.E.D.

Corollary 6.4

Let $\varphi, \psi \in \mathcal{M}_1(E)$. If $m(t, \psi) \Rightarrow \varphi$ as $t \to \infty$, then φ is a stationary probability measure.

Proof: By the definition of the Prohorov metric,

$$m(t, \psi) \Longrightarrow \varphi \text{ as } t \to \infty \Longleftrightarrow \lim_{t \to \infty} \rho(m(t, \psi), \varphi) = 0 .$$

Therefore, for every $\epsilon > 0$ there exists $t_0(\epsilon) > 0$ such that $m(t, \psi) \in S(\varphi, \epsilon)$ for $t > t_0(\epsilon)$. We claim that for $t > t_0(\epsilon)$ there exists $\psi_1 \in S(\varphi, \epsilon)$ such that $m(t, \psi_1) \in S(\varphi, \epsilon)$ for all $t \in R^+$: let $\psi_1 = m(t_0(\epsilon), \psi)$, then

$$m(t, \psi_1) = m(t, m(t_0(\epsilon), \psi)) = m(t + t_0(\epsilon), \psi) ,$$

which implies that $m(t, \psi_1) \in S(\varphi, \epsilon)$ for all $t > 0$. Proposition 6.2 then implies that φ is a stationary probability measure.

Q.E.D.

Definition 6.4. A motion of probability measures for all $t \in R^+$ and some $\tau > 0$ satisfies the condition $m(t + \tau, \varphi) = m(t, \varphi)$, $\varphi \in \mathfrak{M}_1(E)$, is called periodic.

By definition of a NDDS it follows that

$$m(t + n\tau, \varphi) = m(t, m(n\tau, \varphi)) = m(n\tau, \varphi) .$$

The smallest positive number $\tau > 0$ satisfying $m(t + \tau, \varphi) = m(t, \varphi)$ is called the period of $m(t, \varphi)$. If a periodic motion of probability measures does not have a least period τ, then $m(t, \varphi)$ is a stationary probability measure.

Proposition 6.5

If there exists at least one $s \geq 0$ and one $\tau \geq 0$ such that $m(s + \tau, \varphi) = m(s, \varphi)$, $\varphi \in \mathfrak{M}_1(E)$, then $m(t, \varphi)$ is periodic.

Proof: The proof is exactly as in [32, p. 18].

The concept of a stationary probability measure can be imbedded in the concept of an invariant set of probability measures.

Definition 6.5 A set $M \subset \mathfrak{M}_1(E)$ is called (positively) invariant if under all transformations of the semi-group $\{ m_t : t \geq 0 \}$ it is transformed into itself. That is, for all $t \geq 0$, $m_t(M) \equiv \{ m_t \varphi : \varphi \in M \} \subset M$.

The proof of the following proposition can be found in [32 , p. 21].

Proposition 6.6

A set $M \subset \mathfrak{M}_1(E)$ is invariant if and only if $\varphi \in M$ implies $m(t, \varphi) \in M$ for all $t \geq 0$.

Proposition 6.6 is equivalent to saying that invariant sets of probability measures consist of entire trajectories of probability measures.

The following is a standard definition in the theory of dynamical systems [32, p. 28].

Definition 6.6. $\psi \in \mathfrak{M}_1(E)$ is called an ω-limit measure of $\varphi \in \mathfrak{M}_1(E)$ if there exists a sequence $\{ t_n \} \to +\infty$ such that $m(t_n, \varphi) \Longrightarrow \psi$. The set of all ω limit measures of φ is called the ω limit set of φ and denoted by $\overset{+}{\Lambda}(\varphi)$. Thus,

$$\overset{+}{\Lambda}(\varphi) = \left\{ \psi \in \mathfrak{M}_1(E) : \exists \ \{ t_n \} \to +\infty \quad \text{such that } m(t_n, \varphi) \Longrightarrow \psi \right\}.$$

The set of all ω limit measures of all $\varphi \in N \subset \mathfrak{M}_1(E)$ is called the ω limit set of N. Thus,

$$\overset{+}{\Lambda}(N) = \cup \left\{ \overset{+}{\Lambda}(\varphi) : \varphi \in N \right\} .$$

Proposition 6.7

For every $\varphi \in \mathcal{M}_1(E)$, $\Lambda^+(\varphi)$ is weak closed in $\mathcal{M}_1(E)$ and invariant.

Proof: Consider the sequence $\{\psi_k\} \subset \Lambda^+(\varphi)$ such that $\psi_k \Longrightarrow \psi$. We must show that $\psi \in \Lambda^+(\varphi)$. For each ψ_k there exists a sequence $\{t_n^k\} \to +\infty$ such that $m(t_n^k, \varphi) \Longrightarrow \psi_k$ as $n \to \infty$. We may assume without loss of generality that $\rho(m(t_k^k, \varphi), \psi_k) < \frac{1}{k}$ for all k, $t_k^k > k$. Then, letting $t_n = t_n^n$, we have $t_n \to +\infty$ and $m(t_n, \varphi) \Longrightarrow \psi$ since

$$\rho(m(t_n, \varphi), \psi) \leq \rho(m(t_n, \varphi), \psi_n) + \rho(\psi, \psi_n).$$

$$\leq \frac{1}{n} + \rho(\psi, \psi_n) \to 0 \qquad \text{as } n \to \infty.$$

Thus, $\psi \in \Lambda^+(\varphi)$ and $\Lambda^+(\varphi)$ is weak closed in $\mathcal{M}_1(E)$.

To show that $\Lambda^+(\varphi)$ is invariant, consider the sequence $\{t_n\} \to \infty$ such that $m(t_n, \varphi) \Longrightarrow \psi \in \Lambda^+(\varphi)$. We must show that $m(t, \psi) \in \Lambda^+(\varphi)$ for all $t \geq 0$. Consider the point $m(\tau, \psi)$ where $\tau \geq 0$ is arbitrary and fixed. From the weak continuity axiom of NDDS,

$$m(t_n + \tau, \varphi) = m(\tau, m(t_n, \varphi)) \Longrightarrow m(\tau, \psi),$$

which implies that $m(\tau, \psi) \in \Lambda^+(\varphi)$. This can be proved for all $\tau \geq 0$.

Q.E.D.

If E is a compact space, then $(\mathcal{M}_1(E), \rho)$ is a compact space [13, Theorem 6.4, p. 45], which implies that the weak closure of

$m(R^+, \varphi) \equiv \{m(t, \varphi) : t \geq 0\}$, $\overline{m(R^+, \varphi)}$, and $\Lambda^+(\varphi)$, are weak compact, where $\varphi \in \mathcal{M}_1(E)$.

Proposition 6.8

Let E be a complete separable metric space. If $\varphi \in \mathcal{M}_1(E)$ and $\overline{m(R^+, \varphi)}$ is weak compact, then

$$\lim_{t \to \infty} \rho(m(t, \varphi), \Lambda^+(\varphi)) = 0 .$$

Proof: The proof proceeds exactly as in the deterministic case. Suppose the conclusion is false, then there could be found a sequence $\{t_n\} \to +\infty$ and an $\alpha > 0$ such that for all n

$$\rho(m(t_n, \varphi), \Lambda^+(\varphi)) \geq \alpha > 0 . \qquad (6.6)$$

The sequence $\{m(t_n, \varphi)\} \subset m(R^+, \varphi)$, and contains a subsequence $\{m(t_n', \varphi)\}$ such that $m(t_n', \varphi) \Longrightarrow \psi \in m(R^+, \varphi)$ as $t_n' \to \infty$, since $\overline{m(R^+, \varphi)}$ is weak compact by hypothesis. Thus, $\psi \in \Lambda^+(\varphi)$ and

$$\lim_{n \to \infty} \rho(m(t_n', \varphi), \Lambda^+(\varphi)) = \rho(\psi, \Lambda^+(\varphi)) = 0 ,$$

contradicting (6.6).

Q.E.D.

Proposition 6.9

If, for some $\varphi \in \mathcal{M}_1(E)$, $\Lambda^+(\varphi)$ is nonempty and has a weak compact neighbourhood which strictly contains it, then $\overline{m(R^+, \varphi)}$ is weakly compact. If the

NDDS $\{m_t : t \in R^+\}$ is also defined by a Feller process, then there exists an invariant measure for the process.

<u>Proof:</u> Let N be the weak compact neighbourhood strictly containing $\Lambda^+(\varphi)$. We claim that there exists a $\tau > 0$ such that $m(R^+, m(\tau, \varphi)) \subset N$. Otherwise, there exists a sequence $\{t_n\} \to \infty$ with $m(t_n, \varphi) \in \partial N$, since $\Lambda^+(\varphi)$ is nonempty.* Therefore, since the boundary of any set is closed, ∂N is weak compact. This implies the existence of a subsequence $\{t'_n\} \to \infty$, such that

$$m(t'_n, \varphi) \Longrightarrow \psi \in \partial N \quad .$$

Hence, $\psi \in \Lambda^+(\varphi)$. But $\psi \in \partial N$ and N strictly contains $\Lambda^+(\varphi)$, which is impossible. So there exists $\tau > 0$ such that $m(R^+, m(\tau, \varphi)) \subset N$.

Now,

$$m(R^+, \varphi) = m([0, \tau], \varphi) \cup m(R^+, m(\tau, \varphi)) \quad ,$$

and

$$\overline{m(R^+, \varphi)} = m([0, \tau], \varphi) \cup \overline{m(R^+, m(\tau, \varphi))} \, .$$

$m([0, \tau], \varphi)$ is weak compact since $m(\cdot, \varphi)$ is a continuous map from the compact interval $[0, \tau]$, and $\overline{m(R^+, m(\tau, \varphi))}$ is weak compact since it is closed and contained in N. Therefore, $\overline{m(R^+, \varphi)}$ is weak compact.

If $\{m_t : t \in R^+\}$ is defined by a Feller process, and the hypothesis of the theorem is satisfied, then from conditions (i) and (iii) of [43, p. 204] it follows that there exists an invariant measure for the process.

Q.E.D.

---

* This statement requires the continuity of $m(t, \varphi)$.

Remarks

(i) Weak compactness of $\Lambda^+(\varphi)$ is not sufficient to ensure the existence of a weak compact neighbourhood containing $\Lambda^+(\varphi)$.

(ii) In [35, Theorem 3] it is shown that if, for a process satisfying certain conditions, $m(R^+, \varphi)$ is weak compact, then $\Lambda^+(\varphi)$ is weak compact. This result is a direct consequence of [32, Theorem 2.2.13, p. 119]. Actually, the result in [35, Theorem 3] is slightly more general in that it considers a two-sided flow, i.e., on $(-\infty, \infty)$, but for practical purposes this extension is of little significance. Also, the assumption of the weak compactness of $m(R^+, \varphi)$ is tantamount to assuming stability in the sense of Lagrange [32, Section I.5.1].

(iii) The theory of NDDS can be carried much further than is done here. For instance, we can study minimal sets, prolongations, attractors, and many other implements of dynamical system theory on metric spaces [32, Chapter 2].

6.4 Some Results Concerning $\Lambda^+(\varphi)$

In this section we present some results concerning the weak limit set of a probability measure φ. Let $E = R^n$ and let $\{m_t : t \in R^+\}$ be a NDDS.

Proposition 6.10

Let $\varphi \in \mathcal{M}_1(E)$ and $g \in C(E)$. If $m_t\varphi(g)$ is either non-increasing or non-decreasing for t sufficiently large, then $\mu(g)$ = constant for all $\mu \in \Lambda^+(\varphi)$.

Proof: We shall prove the proposition only in the case that $m_t\,\varphi(g)$ is non-decreasing for t large. For any $\mu \in \overset{+}{\Lambda}(\varphi)$, let $\{t_n\} \to \infty$ be such that

$$\mu_n \equiv m_{t_n}\,\varphi \Longrightarrow \mu \quad .$$

Then,

$$\limsup_{t\to\infty} m_t\,\varphi(g) = \lim_{n\to\infty} m_{t_n}\,\varphi(g)$$

$$= \lim_{n\to\infty} \mu_n(g) = \mu(g)\ .$$

Since $\limsup_{t\to\infty} m_t\,\varphi(g)$ is independent of $\mu \in \overset{+}{\Lambda}(\varphi)$, $\mu(g)$ is constant for all $\mu \in \overset{+}{\Lambda}(\varphi)$.

The proof for the non-increasing case is similar.

Q.E.D.

Proposition 6.11

Let $g(x)$ be a $\mathcal{B}(E)$ measurable real-valued function such that $|g(x)| \to +\infty$ as $\|x\| \to \infty$, where $\|\ \|$ is the euclidean norm on $E = R^n$. Then $m_t\,\varphi(|g|) \to \infty$ as $t \to \infty$, $\varphi \in \mathfrak{M}_1(E)$, if $\overset{+}{\Lambda}(\varphi) = \emptyset$ (empty set).

Proof: Suppose $\overset{+}{\Lambda}(\varphi) = \emptyset$ but $m_t\,\varphi(|g|) < K < \infty$ for all $t \geq 0$. Then, the family $\Xi = \{m_t\,\varphi : t \geq 0\}$ is realtively weakly compact : If not, there exists an $\epsilon > 0$ such that for every compact set $J \subset E$,

$$\sup_{\mu\in\Xi} \mu(J^c) > \epsilon\ ^{*},$$

---

* J^c is the complement of the set J in E.

which implies that

$$\sup_{\mu \in \Xi} \mu(|g|) = \infty ,$$

contradicting the fact that $m_t \varphi(|g|) < K$ for all $t \geq 0$. Thus, Ξ is relatively weakly compact [13, Theorem 6.7, p. 47].

Hence, given any sequence $\{ m_{t_n} \varphi : t_n \geq 0 , t_n \to \infty \}$, there exists a convergent subsequence $\{ m_{t'_n} \varphi : t'_n \geq 0 , t'_n \to \infty \}$ and a $\psi \in \mathcal{M}_1(E)$ such that

$$m_{t'_n} \varphi \Rightarrow \psi ,$$

which implies that $\psi \in \Lambda^+(\varphi)$, contradicting $\Lambda^+(\varphi) = \emptyset$. Thus,

$$m_t \varphi \; (|g|) \longrightarrow +\infty \qquad \text{as } t \to +\infty .$$

Q.E.D.

In particular, for a Markov process which is a NDDS with φ the starting probability measure, $|g(x)| \to \infty$ as $\|x\| \to \infty$, and $\Lambda^+(\varphi) = \emptyset$, we have $E_\varphi |g(x_t)| \to \infty$ as $t \to \infty$.

<u>Proposition 6.12</u>

Let A be the s-infinitesimal generator of a $(\mathcal{C}_o)$ Markov semi-group on $C_o(E)$* and let A^* be its adjoint operator. If $\varphi \in \mathcal{M}_1(E) \cap \mathcal{D}(A^*)$, then $\Lambda^+(\varphi) \subset \mathcal{M}_1(E) \cap \mathcal{D}(A^*)$ if $Z \equiv \{ m(t, \varphi) : t \geq 0 \}$ is weak * closed.

---

* $C_o(E)$ is the Banach space, with the supremum norm, of all real bounded continuous functions vanishing at ∞.

Proof: Let $\varphi \in \mathcal{M}_1(E) \cap \mathfrak{D}(A^*)$ and $\psi \in \Lambda^+(\varphi)$. Then, for some $\{t_n\} \to +\infty$,

$$m(t_n, \varphi) \Longrightarrow \psi .$$

By Proposition 2.5, $Z \subset \mathfrak{D}(A^*)$. Since Z is weak * closed, $\psi \in Z$ and thus $\psi \in \mathfrak{D}(A^*)$. Also, $\varphi \in \mathcal{M}_1(E)$ implies that $\| m_t \varphi \| = 1$ for all $t \geq 0$. Therefore, $\psi \in \mathcal{M}_1(E)$ and

$$\Lambda^+(\varphi) \subset \mathcal{M}_1(E) \cap \mathfrak{D}(A^*) .$$

Q.E.D.

The operator A^* is defined by

$$A^* \varphi = \underset{h \downarrow 0}{\text{weak*lim}} \frac{m(h, \varphi) - \varphi}{h}$$

and, iteratively, we can define A^{*^n}. Then, we have

Corollary 6.13

If $\varphi \in \mathcal{M}_1(E) \cap \mathfrak{D}(A^{*^n})$, then $\Lambda^+(\varphi) \in \mathcal{M}_1(E) \cap \mathfrak{D}(A^{*n})$.

Proposition 6.14

Let $\varphi \in \mathcal{M}_1(E) \cap \mathfrak{D}(A^*)$ and $g \in C_o(E)$ such that $m_t \varphi(g)$ is either non-increasing or non-decreasing as $t \to \infty$. Then, for any $\psi \in \Lambda^+(\varphi)$, $(A^* m_t \psi)(g) = 0$ for all $t \geq 0$ if $\{m(t, \varphi) : t \geq 0\}$ is weak * closed.

Proof: Let $\varphi \in \mathcal{M}_1(E) \cap \mathfrak{D}(A^*)$. Then, by Proposition 6.12, $\Lambda^+(\varphi) \subset \mathcal{M}_1(E) \cap \mathfrak{D}(A^*)$. Let $\psi \in \Lambda^+(\varphi)$, and using Proposition 2.5, we have

$$m_t \psi(g) - \psi(g) = \int_0^t A^* m_s \psi(g)\, ds . \qquad (6.7)$$

The formula (6.7) implies that the real-valued function $m_t\psi(g)$ is differentiable in t and

$$\frac{d}{dt}\, m_t\, \psi(g) \;=\; (A^*\, m_t\, \psi)(g) \qquad \text{for all } t \geq 0 \quad .$$

Now, from Proposition 6.1, we know that $\mu(g)$ = constant for all $\mu \in \Lambda^+(\varphi)$. Thus, since $\Lambda^+(\varphi)$ is an invariant set of probability measures, $m_t\, \psi(g)$ = constant for all $t \geq 0$, implying that

$$(A^*\, m_t\, \psi)(g) \;=\; 0 \quad .$$

Q.E.D.

<u>Corollary 6.15</u>

For $\varphi \in \mathfrak{M}_1(E) \cap \mathfrak{D}(A^{*n})$, $(A^{*n}\, m_t\, \psi)(g) = 0$ for all $t \geq 0$, where $\psi \in \Lambda^+(\varphi)$ and $g \in C_o(E)$ is such that $m_t\, \varphi(g)$ is either non-increasing or non-decreasing.

<u>Remark</u>:

In [35], Kushner applied $\Lambda^+(\varphi)$ to the investigation of stability properties of stochastic processes. The goal of his work was to obtain a set in E to which the process converges in probability and which is sometimes smaller than the sets obtained by the theory of stochastic Lyapunov functions [34, Chapter II, Theorems 2 and 3].

CHAPTER VII

STABILITY OF NON-DETERMINISTIC DYNAMICAL SYSTEMS

7.1 Stability in the Space of Measures

We make the following definitions for the metric space $(\mathfrak{M}_1(E), \rho)$ and fixed NDDS $\{ m_t : t \geq 0 \}$, $E = R^n$.

Definition 7.1. An invariant weak closed set $M \subset \mathfrak{M}_1(E)$ is stable in the Lyapunov sense if, for any $\epsilon > 0$, there exists $\delta > 0$ such that $\rho(\varphi, M) < \delta$ implies that $\rho(m(t, \varphi), M) < \epsilon$ for all $t \geq 0$. If, furthermore, $\rho(m(t, \varphi), M) \to 0$ as $t \to \infty$, M is said to be asymptotically stable in the Lyapunov sense.

Definition 7.2. A function $V \in C(E)$ is called a D-function for a weak closed invariant set $M \subset \mathfrak{M}_1(E)$ if V has the following properties for some small $r > 0$:

(1) For any sufficiently small $c_1 > 0$, it is possible to find $c_2 > 0$ such that $\varphi(V) > c_2 \left(\varphi(V) \equiv \int_E V(x) \, \varphi(dx) \right)$ for all φ which satisfy $\rho(\varphi, M) > c_1$, $\varphi \in S(M, r)$.

(2) For any $\gamma_2 > 0$ there exists a $\gamma_1 > 0$ such that $\rho(\varphi, M) < \gamma_1$ implies that $\varphi(V) < \gamma_2$.

(3) $m(t, \varphi)(V) \leq \varphi(V)$ for all $\varphi \in S(M, r)$ and for all $t \geq 0$.

It is important to note that V is not a Lyapunov function for M as defined, for instance, in [32, Chapter 2.7]; it cannot be, since its domain of definition is not even in $\mathfrak{M}_1(E)$. But, as we shall see, V acts like a Lyapunov function, enabling us to prove stability theorems in $\mathfrak{M}_1(E)$.

Using the following lemma, Definition 7.2 can be simplified.

Lemma 7.1

(1) and (2) of Definition 7.2 is equivalent to the condition (4):

$\rho(\varphi_n, M) \to 0$ as $n \to \infty \Longleftrightarrow \varphi_n(V) \to 0$ as $n \to \infty$.

Proof: (4) $\Longrightarrow$ (2) since $\rho(\varphi_n, M) \to 0 \Longrightarrow \varphi_n(V) \to 0$ means that for all $\gamma_2 > 0$ there exists $\gamma_1 > 0$ such that $\varphi_n(V) < \gamma_2$, whenever $\rho(\varphi_n, M) < \gamma_1$.

(4) $\Longrightarrow$ (1). Assume (1) is not true, i.e., there doesn't exist $c_2 > 0$ such that $\varphi_n(V) > c_2$ for all φ_n for which $\rho(\varphi_n, M) > c_1$, $\{\varphi_n\} \subset S(M, r)$. Then, there exists a subsequence $\{\varphi_{n_k}\}$ such that $\varphi_{n_k}(V) \to 0$. By (4), $\varphi_{n_k}(V) \to 0$ implies $\rho(\varphi_{n_k}, M) \to 0$. This is a contradiction since we started with $\rho(\varphi_n, M) > c_1$ for all n. Thus (4) $\Longrightarrow$ (1) and (2).

Now (1) implies that given $c_1 > 0$ there exists $c_2 > 0$ such that $\varphi(V) < c_2$ for all φ such that $\rho(\varphi, M) < c_1$ and $\varphi \in S(M, r)$ i.e.,

$$\rho(\varphi_n, M) \to 0 \Longrightarrow \varphi_n(V) \to 0 .$$

Similarly, (2) implies that

$$\varphi_n(V) \to 0 \quad \Longrightarrow \quad \rho(\varphi_n, M) \to 0 .$$

Thus, $\rho(\varphi_n, M) \to 0 \Longleftrightarrow \varphi_n(V) \to 0$, and the proof is complete.

Q.E.D.

We can now rewrite Definition 7.2 as follows:

<u>Definition 7.2'</u> Given a weak closed invariant set $M \subset \mathfrak{M}_1(E)$. A function $V \in C(E)$ is called a <u>D-function for M</u> if, for some small $r > 0$,

(1) $\rho(\varphi_n, M) \to 0 \Longleftrightarrow \varphi_n(V) \to 0$ for $\{\varphi_n\} \subset S(M, r)$,

(2) $m(t, \varphi)(V) \leq \varphi(V)$ for all $\varphi \in S(M, r)$ and for all $t \geq 0$.

There may be situations where (1) and (2) can only be proved for certain subsets of $S(M, r)$. For instance, using the integral representation in Proposition 2.5(b) $\left(\text{where } \mathfrak{X} = C_o(E) \text{ and } \mathfrak{X}^* = \mathfrak{M}(E)\right)$ may facilitate the proof of (1) and (2) for points in the subset $\mathfrak{M}_1(E) \cap \mathfrak{D}(A^*)$ of $\mathfrak{M}_1(E)$. In this spirit, the following proposition presents a sufficient condition for the existence of a D-function for the invariant set $M \subset \mathfrak{M}_1(E)$.

<u>Proposition 7.2</u>

Let $\mathcal{H}$ be a weak dense subset of $\mathfrak{M}(E)$. If there exists a $V \in C(E)$ such that for some small $r > 0$,

(1) $\rho(\varphi_n, M) \to 0 \Longleftrightarrow \varphi_n(V) \to 0$ for $\{\varphi_n\} \subset \mathcal{H} \cap S(M, r)$,

(2) $m(t, \varphi)(V) \leq \varphi(V)$ for all $\varphi \in \mathcal{H} \cap S(M, r)$ and all $t \geq 0$,

then V is a D-function for M.

<u>Proof:</u> Let $\{\varphi_n\} \subset S(M, r)$. Since $\mathcal{H} \cap S(M, r)$ is weak dense in $S(M, r)$, for each φ_n there exists a sequence $\{\psi_{n,j}\} \subset \mathcal{H} \cap S(M, r)$ such that $\rho(\varphi_n, \psi_{n,j}) \to 0$ as $j \to \infty$. Letting $\psi_n = \psi_{n,n}$, we have that $\rho(\varphi_n, \psi_n) \to 0$ as $n \to \infty$.

Now, (1) implies that

$$\text{(a)} \quad \lim_{n\to\infty} \rho(\psi_n, M) = 0 \iff \lim_{n\to\infty} \psi_n(V) = 0 \quad \text{(b)} \quad .$$

Under the condition $\lim_{n\to\infty} \rho(\varphi_n, \psi_n) = 0$, (a) is equivalent to $\lim_{n\to\infty} \rho(\varphi_n, M) = 0$ since the map $\mu \to \rho(\mu, G)$ from $\mathfrak{M}_1(E)$ into R is continuous, where G is any subset of $\mathfrak{M}_1(E)$ [16, Theorem 4.3, p. 185], and (b) is equivalent to $\lim_{n\to\infty} \varphi_n(V) = 0$. Thus,

$$\lim_{n\to\infty} \rho(\varphi_n, M) = 0 \iff \lim_{n\to\infty} \varphi_n(V) = 0 \quad ,$$

where $\{\varphi_n\} \subset S(M, r)$.

It remains to prove (2) of Definition 7.2'. Let $\varphi \in S(M, r)$. Then there exists a sequence $\{\psi_i\} \subset \mathcal{H} \cap S(M, r)$ such that

$$\lim_{i\to\infty} \rho(\psi_i, \varphi) = 0 \quad .$$

By virtue of (2), for all $t \geq 0$,

$$m(t, \psi_i)(V) \leq \psi_i(V) \qquad \text{for all } i \quad . \tag{7.1}$$

Now, under the condition $\lim_{i\to\infty} \rho(\psi_i, \varphi) = 0$, the right-hand side of (7.1) approaches $\psi(V)$ as $i \to \infty$, and the left-hand side approaches $m(t, \varphi)(V)$, by the weak continuity axiom of NDDS. Hence,

$$m(t, \varphi)(V) \leq \varphi(V)$$

for all $t \geq 0$ and $\varphi \in S(M, r)$.

Q.E.D.

We now present the major theorem of this section. In form it is similar to the sufficiency part of the deterministic theorem [33, Theorem 12, p. 41], but conceptually it is quite different. The interesting fact is that the existence of a specific function in C(E) assures stability, in the sense of Definition 7.1, of flows in a subset of the dual space of C(E).

Theorem 7.3

In order for a weak closed set $M \subset \mathfrak{M}_1(E)$ to be stable, it is sufficient that there exist a D-function for M.

Proof: Let there exist a D-function V in a certain neighbourhood $S(M, r)$ of M. Let $r > \epsilon > 0$ and let

$$\lambda = \inf \left\{ \varphi(V) : \text{ for all } \varphi \text{ such that } \rho(\varphi, M) = \epsilon \right\} . \qquad (7.2)$$

By (1) of Definition 7.2, $\lambda > 0$. By (2) of Definition 7.2, it is possible to find for λ, $\delta > 0$ such that $\rho(\varphi, M) < \delta$ implies that $\varphi(V) < \lambda$. We would now like to show that $\rho(\varphi, M) < \delta$ implies $\rho(m(t, \varphi), M) < \epsilon$ for all $t \geq 0$. Assume the opposite, i.e., that there exists $\varphi \in S(M, \delta)$ such that at some $t_1 > 0$, $\rho(m(t_1, \varphi), M) = \epsilon$ holds true. Then, by (7.2), $m(t_1, \varphi)(V) \geq \lambda$. But, by (3) of Definition 7.2, $\varphi(V) < \lambda$ implies $m(t, \varphi)(V) < \lambda$ for all $t \geq 0$ implying that $m(t_1, \varphi)(V) < \lambda$ which gives the contradiction. Hence, M is istable.

Q.E.D.

Theorem 7.4

In order for a weak closed invariant set $M \subset \mathfrak{M}_1(E)$ to be asymptotically stable it is sufficient that there exist a D-function V for M and that

$m(t, \varphi)(V) \to 0$ as $t \to \infty$ for any $\{ m(t, \varphi) : t \geq 0 \} \subset S(M, r)$.

Proof: By Theorem 7.1 it follows that the invariant set M is stable, i.e., for any $\epsilon > 0$ there exists $\delta > 0$ such that

$$\rho(\varphi, M) < \delta \implies \rho(m(t, \varphi), M) < \epsilon \qquad \text{for all } t \geq 0 . \tag{7.2}$$

We show that δ can be chosen so that $\rho(m(t, \varphi), M) \to 0$ as $t \to \infty$, and $\rho(\varphi, M) < \delta$. For the δ obtained in (7.2), we can find δ_1 such that

$$\rho(\varphi, M) < \delta_1 \implies \rho(m(t, \varphi), M) < \delta \qquad \text{for all } t \geq 0 .$$

We claim that $\lim_{t \to \infty} \rho(m(t, \varphi), M) = 0$. Assume the contrary, i.e., there exists at least one probability measure $\psi \in S(M, \delta_1)$ such that

$$\rho(m(t, \psi), M) > \lambda_1 > 0 \qquad \text{for all } t \geq 0 .$$

Then, by (1) of Definition 7.2, $m(t, \psi)(V) > \lambda_2$ for some $\lambda_2 > 0$ and for all $t \geq 0$ which contradicts the condition

$$m(t, \psi)(V) \to 0 \qquad \text{as } t \to \infty .$$

Q.E.D.

Remark:

(i) Definition 7.2 can be extended in the following manner: Let $\mathfrak{X}$ be a separable Banach space and $\mathfrak{X}^*$ its dual space. Then, the closed unit sphere, $\mathfrak{X}_1^*$, of $\mathfrak{X}^*$ with the weak*topology is a metric space [1, Theorem V.5.1, 1. 426] having metric d. If a dynamical system (Definition 6.1) is defined on $(\mathfrak{X}_1^*, d)$, the existence of an element $x_V \in \mathfrak{X}$, satisfying properties (1) - (3) of Definition 7.2, where $\varphi \in \mathfrak{M}_1(E)$ is replaced by $x^* \in \mathfrak{X}_1^*$ and $V \in C(E)$ is replaced by $x_V \in \mathfrak{X}$, is sufficient to prove stability of invariant subsets of $\mathfrak{X}_1^*$, exactly as in Theorem 7.3.

7.2 Application to Stochastic Stability Theory

We shall be concerned with the stability properties of the following n-dimensional stochastic differential equation on $E = R^n$:

$$dx_t = a(x_t)\,dt + b(x_t)\,dz_t \quad , \tag{7.3}$$

where each of the components of the vector $a(x)$ and matrix $b(x)$ are Lipschitz continuous. The s-infinitesimal generator of (7.3) is

$$A = \sum_{i=1}^{n} a_i(\cdot)\frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^{n} c_{ij}(\cdot)\frac{\partial^2}{\partial x_i\,\partial x_j} \quad ,$$

where $c = b^T b$. The process associated with (7.3) is a diffusion process with transition density function $p(t, x, y)$.

We assume that $a(0) = 0$ and $b(0) = 0$. Therefore, $x = 0$ is a stationary point, i.e., if the process starts there, it remains there forever. Let ϵ_o be the Dirac measure at $x = 0$. Then, ϵ_o is a stationary probability measure. Since, $\{\epsilon_o\}$ is also weak compact, it is a weak compact invariant subset of $\mathfrak{M}_1(E)$.

Definition 7.3. A stochastic Lyapunov function for (7.3) is a continuous function $V \in \mathfrak{D}(A)$ such that

(i) $V : E \rightarrow [0, \infty)$

(ii) $V(0) = 0\,,\ V(x) > 0$ for $x \neq 0$

(iii) $(AV)(x) \leq 0$ for all $x \in E$.

Proposition 7.4

A stochastic Lyapunov function $V(x)$ for (7.3) is a D-function for the weak compact invariant set $\{\epsilon_o\} \subset \mathcal{M}_1(E)$ for the NDDS $\{m_t : t \geq 0\}$ generated by the transition density function $p(t, x, y)$ of (7.3).

Proof: Let $S(\epsilon_o, r)$ be a small ρ-neighbourhood of ϵ_o. Let $\{\varphi_n\} \subset S(\epsilon_o, r)$. First we show that $\rho(\varphi_n, \epsilon_o) \to 0$ if and only if $\varphi_n(V) \to 0$. By the definition of ρ, $\rho(\varphi_n, \epsilon_o) \to 0$ implies that $\varphi_n(f) \to \epsilon_o(f)$ for each $f \in C(E)$. In particular, $\varphi_n(V) \to \epsilon_o(V)$. But $\epsilon_o(V) = 0$ since $V(0) = 0$, implying that $\varphi_n(V) \to 0$ as $n \to \infty$.

To show the reverse implication, suppose $\varphi_n(V) \to 0$ but $\rho(\varphi_n, \epsilon_o) \not\to 0$. Then, since $V(x) > 0$ for $x \neq 0$, $\varphi_n(V) \not\to 0$ which is a contradiction. Therefore, (1) of Definition 7.2' is satisfied.

To prove (2) of Definition 7.2', we proceed as follows. For any $x \in E$, by condition (iii) of Definition 7.3, we have

$$T_t V(x) - V(x) = \int_0^t T_s \, A \, V(s) \, ds \leq 0 \qquad \text{for all } t \geq 0, \qquad (7.4)$$

where we used Dynkin's Formula and the definition

$$T_t \, V(x) = \int_E V(y) \, p(t, x, y) \, dy \qquad .$$

Since

$$m(t, \varphi)(\Gamma) \equiv \int_E \int_\Gamma p(t, x, y) \; dy \; \varphi(dx) \, ,$$

(7.4) implies that for all $x \in E$,

$$m(t, \epsilon_x)\,(V) \leq \epsilon_x(V) \qquad \text{for all } t \geq 0 . \tag{7.5}$$

Let $\varphi \in S(\epsilon_o, r)$, and integrate both sides of (7.5) with respect to φ to obtain

$$m(t, \varphi)\,(V) \leq \varphi(V) \qquad \text{for all } t \leq 0 .$$

Hence the proof is complete.

Q.E.D.

The map $y \to \epsilon_y$ from E into the space of Dirac measures on E is a homeomorphism [13, Lemma 6.1, p. 42]. Therefore, given any $\delta_1 > 0$, there exists $\delta > 0$ such that

$$\| x \| < \delta \implies \rho(\epsilon_x , \epsilon_o) < \delta_1 .$$

Hence, if there exists a D-function for ϵ_o, Theorem 7.3 implies that for any $\alpha > 0$ there exists $\delta > 0$ such that

$$\rho(m(t, \epsilon_x), \epsilon_o) < \alpha$$

for all $t \geq 0$, if $\| x \| < \delta$.

If ϵ_o is asymptotically stable, then the solution process of (7.3) converges in probability to 0 as $t \to \infty$. To see this, let $F_t^x(y) = m(t, \epsilon_x)\,((-\infty, y])$ be the distribution function at time t of the process associated (7.3) starting at $t = 0$ and $x \in E$. In [12, pp. 17-18] it is shown that weak convergence of probability measures is the same as convergence in distribution of the corresponding distribution functions. Then, since ϵ_o is asymptotically stable,

$$\lim_{t \to \infty} \rho(m(t, \epsilon_x), \epsilon_o) = 0$$

for x in some euclidean neighbourhood $\mathcal{O} \equiv \{ y : \| y \| < \delta \}$ of the origin. This implies that

$$F_t^x(\cdot) \xrightarrow{\mathcal{D}} F_o(\cdot)$$

as $t \to \infty$, where $\mathcal{D}$ indicates that the convergence is in distribution, and

$$F_o(y) = \begin{cases} 0 & \text{if } y < 0 \\ 1 & \text{if } y \geq 0 \end{cases} .$$

Convergence in distribution to a Dirac measure implies convergence in probability to its support [18, Exercise 10, p. 86]. Thus, for $x \in \mathcal{O}$ and for all $\alpha > 0$,

$$W\left\{ \| x_t^x(\omega) \| > \alpha \right\} \to 0$$

as $t \to \infty$, where $W\left\{ x_t^x(\omega) = x \right\} = 1$. This result is similar in form to those in [34] and other works on stochastic stability theory.

To summarize, we have shown that certain stochastic stability problems can be reformulated in the framework of NDDS, i.e., we can study the stability of stochastic processes such as diffusions (in the sense of Definition 7.1) by examining the induced flows of probability measures on the range space, rather than by investigating the sample path behaviour.

<u>Remarks</u>:

(i) Let the NDDS $\{ m_t : t \geq 0 \}$ be induced by a Markov process. Then, since $\{ m_t : t \geq 0 \}$ is a contraction semi-group on $\mathcal{M}(E)$, given any $\epsilon > 0$, $\varphi, \psi \in \mathcal{M}_1(E)$,

$$\| \varphi - \psi \| < \epsilon \Longrightarrow \| m(t, \varphi) - m(t, \psi) \| < \epsilon \qquad (7.6)$$

for all $t \geq 0$. (This is a rather weak stability result.) In [46, Theorem 12.2, pp. 317 - 318], it is shown that there exists a countable family $\{f_m\} \subset B(R)$, $\| f_m \| \leq 1$ for all m, such that $\lim_{n \to \infty} \rho(\mu_n, \mu) = 0$ if and only if $\lim_{n \to \infty} \mu_n(f_m) = \mu(f_m)$ for all m, and the metric ρ is equivalent to the metric $\tilde{\rho}$, where $\tilde{\rho}(\mu_n, \mu) \equiv \sum_{m=1}^{\infty} \frac{1}{2^m} | \mu_n(f_m) - \mu(f_m) | \leq \| \mu_n - \mu \|$.

Therefore, given $\epsilon > 0$, it follows from (7.6) that

$$\| \varphi - \psi \| < \epsilon \Longrightarrow \tilde{\rho}(m(t, \varphi), m(t, \psi)) < \epsilon \qquad (7.7)$$

for all $t \geq 0$. Fix $\psi \in \mathcal{M}_1(E)$. Then (7.7) implies that the map $\mu \xrightarrow{\beta} \tilde{\rho}(m(t, \mu), m(t, \psi))$ from $(\mathcal{M}_1(E), \| \cdot \|)$ into R, is uniformly continuous with respect to $t \geq 0$. Moreover, if β remains uniformly continuous for $t \geq 0$ when $\mathcal{M}_1(E)$ has the weak topology then we have stability of the flow $\{m(t, \psi) : t \geq 0\}$ in the sense of Definition 7.1, i.e. given $\epsilon > 0$ there exists $\delta > 0$ such that

$$\tilde{\rho}(\varphi, \psi) < \delta \Longrightarrow \sup_{t \geq 0} \tilde{\rho}(m(t, \varphi), m(t, \psi)) < \epsilon .$$

(ii) The work in this chapter can be extended to non-homogeneous processes by considering time-varying D-functions, in a manner analogous to that of [33, Chapter IV, Section 2].

CHAPTER VIII

CONCLUDING REMARKS

In this dissertation, some problems in the stochastic control and stability theories have been studied by considering the flows of probability measures associated with the stochastic process. Existing work on these subjects deals mainly with the process sample path behaviour, based on methods which depend, to a large degree, on the rather technical theory of Markov processes. In this thesis, the problems are studied in a setting which is closely related to that usually employed for deterministic systems on a euclidean space. The main difference, however, is that the stochastic systems are studied here from the point of view of how they induce flows in the infinite dimensional space of measures.

We now present some general comments and a brief summary of the main contributions of this thesis.

(i) In Chapter II, by using the martingale approach to the study of stochastic differential equations [5], we have proved a 'stochastic bang-bang principle' for the stochastic control system (2.4). We leave the possible extension of this and related results in Chapters III and IV, to more general control systems than (2.4), for future work.

(ii) The functional analytic approach, presented in Chapter V for the integration of the Kolmogorov backward equation, where the drift coefficient is merely bounded and integrable, offers an alternative method to that used in [5], and results in a formula, (5.20), which holds on a larger domain than (2.2). This formula yields a necessary and sufficient condition for a control to minimize a cost functional, for a very large class of admissible controls. Much of the earlier work in stochastic control theory restricts the class of controls to be Lipschitz continuous in the state variable, in order to ensure that the resulting optimal stochastic differential equation is mathematically meaningful.

(iii) We were not concerned with the problem of observability since, obviously, the one-dimensional control system (2.4) is fully observable. However, in the corresponding n-dimensional systems, where one does not necessarily have access to all the states of the process, the problem of observability is an important one (see [23]).

(iv) The problem of controllability is given only a cursory treatment in Chapter IV. Further work in this area is warranted. An important first problem is to find necessary and/or sufficient conditions on the n x n matrices A, B and C of the linear n-dimensional stochastic differential equation

$$dx_t = (Ax_t + B\,u(t))\,dt + C dz_t \quad ,$$

where x_t, $u(t)$, and z_t are n-dimensional vectors, such that for a fixed target $\psi \in \mathcal{M}_1(R)$, the controllable set

$$K = \bigcup_{0 \leq t \leq t_1} \mathcal{K}_{t_1}(t) \quad ,$$

with respect to t_1, contains a Prohorov neighbourhood of ψ. Another problem is to determine under what conditions, if any, $K = \mathcal{M}_1(R)$.

The condition for controllability specified by (4.3) may be too stringent. The following is a weaker definition of controllability: Given $\psi \in \mathcal{M}_1(R^n)$ and $\epsilon > 0$, $\varphi \in \mathcal{M}_1(R^n)$ is ϵ-controllable at time t, with respect to ψ at time t_1, if there exists an admissible control u such that

$$\rho\left(\psi, \int_R P^u(t, x, t_1, \cdot)\varphi(dx)\right) < \epsilon \quad .$$

If the Markov process specified by $P^u(t,x,t_1,\Gamma)$ is a diffusion, generated by an infinitesimal generator whose coefficients satisfy the conditions of [8, Vol. II , Theorem 0.4] , then the problem of stochastic controllability bears a close similarity to problems in the theory of parabolic differential equations (see [55, Chapter 3, Sectiom 10] and [56 ,Chapter 1, Section 14]);however, for quasi-diffusions and more general Markov processes, probabilistic methods have to be used. We remark, more generally, that the stochastic control and stability theories can be related to the corresponding theories for distributed parameter systems [56] only for diffusions (in the sense of [8, Vol.I, Chapter V]); in other words, only when the process is generated by a partial differential equation.

Still another approach to stochastic controllability is described in the following example. Consider the scalar stochastic differential equation

$$dx_t^u = a\,x_t^u\,dt + bu(t)\,dz_t \quad , \tag{8.1}$$

where a, b are constants and $u \in \mathcal{W} \equiv \left\{ u \in \mathcal{L}_\infty(R) : -\sigma \le u(t) \le \sigma \right\}$. Using Ito's Lemma [49, p. 32] , it can easily be verified that the solution of (8.1) is

$$x_t^u = e^{a(t-t_o)} x_o + \int_{t_o}^{t} e^{a(t-s)} bu(s)\,dz_s \quad , \tag{8.2}$$

where x_t^u starts at time t_o with the constant x_o . Let the map $\Gamma_{s,t}^u : R \to R$ be defined by $c_s \to E^W x_t^u$ for the solution process of (8.1) , where c_s is the starting point at time $s < t$. The object is to control the mean $E^W x_t^u$: Given $E^W x_t^u$, and $t_1 > t_o > 0$, we wish to characterize the controllable set

$$\mathcal{X}_{t_o}(t_1) = \left\{ c \in R : \Gamma_{t_o,t_1}^u c = E^W x_{t_1}^u \text{ for some } u \in \mathcal{W} \right\}.$$

Using the isometry property of stochastic integrals [49, p.25] , we have

$$E^{w}\left\{ x_{t_1}^{u} - e^{a(t_1 - t_o)} x_o \right\}^2 = \int_{t_o}^{t_1} e^{2a(t_1 - s)} b^2 u^2(s)\, ds \quad . \tag{8.3}$$

Letting $E^{w} x_{t_1}^{u} = \alpha$, $E^{w}(x_{t_1}^{u})^2 = \beta$, and $\tau = t_1 - t_o$, (8.3) is expanded to

$$e^{2a\tau} x_o^2 - 2\alpha e^{a\tau} x_o + \beta = \int_{t_o}^{t_1} e^{2a(t_1 - s)} b^2 u^2(s) ds \quad .$$

Completing the square, we get

$$(e^{a\tau} x_o - \beta_1)^2 + \beta_2 = \int_{t_o}^{t_1} e^{2a(t_1 - s)} b^2 u^2(s) ds \quad , \tag{8.4}$$

where β_1, β_2 are constants. Given an x_o, we can now determine if there exists a $u \in \mathcal{W}$ such that (8.4) is satisfied, by numerical methods if necessary. To obtain an estimate of the controllable set $\mathcal{X}_{t_o}(t_1)$, we observe that

$$0 \le (e^{a\tau} x_o - \beta_1)^2 \le \int_{t_o}^{t_1} e^{2a(t_1 - s)} b^2 \sigma^2 ds - \beta_2 \, ,$$

i.e.

$$\beta_1 e^{-a\tau} \le x_o \le e^{-a\tau}\left(2a b^2 \sigma^2 (e^{2a\tau} - 1) - \beta_2 \right)^{1/2} + \beta_1 e^{-a\tau} \quad .$$

(v) One of the main features of the dynamical system theory approach of Chapters VI and VII is the facility with which it permits the study of stochastic stability properties with respect to initial sets of probability measures.

(vi) An important aspect of the control theory in Chapters III and IV is the extremely weak condition on the drift coefficient and control term of (2.4) .

(vii) We now briefly summarize* the main contributions of this thesis .

(a) Lemma 3.1 presents a new proof of a well-known result in control theory, without using the powerful (deterministic) bang-bang principle.

---

* See also Section 2.3 .

(b) Theorem 3.6 shows that a solution process of (2.4) for a bounded measurable control can be approximated arbitrarily closely (in the sense of weak convergence on Ω) by a solution process of (2.4) for a bang-bang control. This type of result does not exist in the literature for stochastic control systems.

(c) The weak compactness of the attainable sets of (2.4), proved in Theorem 3.8, is also a new result.

(d) The definition of controllability given in Section 4.1 appears to be the natural extension of the conventional definition of controllability for deterministic systems.

(e) Theorem 4.3 treats a time-optimal stochastic control problem in an original manner, employing attainable sets; this necessitates proving the continuity of the attainable sets in an appropriate topology (Lemma 4.2) .

(f) With the aid of Theorem 3.8, Theorem 4.5 establishes the existence of a control, in a very large admissible control class, which minimizes the average of a cost functional. Theorem 4.5 is similar to Theorem 3 of [23] , but the proof here is completely different, and simpler .

(g) The uniqueness result of Section 5.2 is basically of mathematical interest. However, its implementation in Section 5.3 yields a necessary and sufficient condition,(5.26), for a control u, in a certain admissible class, to minimize the average of a cost functional.

(h) The definition of a dynamical system on the metric space $(\mathcal{M}_1(E), \rho)$ (Definition 6.2) is new to stochastic stability theory, although a theory based on such an approach is hinted at in [35] .

(i) The main results of Section 6.3 are Propositions 6.8 and 6.9, whose proofs rely on fairly standard methods in dynamical system theory but, nevertheless, the results are new to stochastic stability theory.

(j) Section 6.4 is entirely new. Propositions 6.10 and 6.11 relate certain properties of $\overset{+}{\Lambda}(\varphi)$ to properties of the flow $\{ m_t\varphi(g) : t \geq 0 \} \subset R$, where g is a certain real-valued function.. For φ in specified subsets of $\mathcal{M}_1(E)$, $\overset{+}{\Lambda}(\varphi)$ is characterized further in Proposition 6.12 and 6.14.

(k) Definition 7.1 is standard, but Definition 7.2 is new, in spite of its similarity to the usual definition of a Lyapunov function. We remark again that a D-function is not a Lyapunov function ; it does not even operate on the metric space ($\mathcal{M}_1(E), \rho$) .

(l) Proposition 7.2 gives a sufficient condition for a real-valued function to be a D-function.

(m) Theorem 7.3 is the main result of Section 7.1 . In form, it resembles the sufficiency part of [33, Theorem 12, p. 41]; nevertheless, it is a completely different result, since V is not a Lyapunov function.

(n) The work of Section 7.2 has no claims on originality ; it merely shows that the preceding theory reduces to some familiar, more probabilistic results of stochastic stability theory.

Most of the results presented in this dissertation are restatements, in the abstract space $\mathcal{M}_1(E)$, of familiar results in the deterministic theory. This is in contrast to most studies on stochastic control and stability theory, where probabilistic phrases adorn the deterministic statements.

The work in this study has been directed at the theoretical aspects of the stochastic control and stability theories, and did not consider any specific problems of practical interest. It is hoped that the methods and results of this dissertation will provide a basis for such application.

BIBLIOGRAPHY

[1] N. Dunford and J. Schwartz, "Linear Operators", Part I , John Wiley , New York, (Fourth Printing) 1967.

[2] P.L. Butzer and H. Berens, "Semi-Groups of Operators and Approximation", Springer-Verlag, 1967.

[3] E. Hille and R. Phillips, "Functional Analysis and Semi-Groups", Amer. Math. Soc. Coll. Publ. Vol. XXXI , 1957.

[4] A. Renyi, "On Stable Sequences of Events", Sankhya, pp. 293-302, 1963.

[5] D.W. Stroock and S.R.S. Varadhan, "Diffusion Process with Continuous Coefficients", I , Comm. Pure and Appl. Math., Vol. XXII , pp. 345-400, 1969; II , Comm. Pure and Appl. Math., Vol. XXII , pp. 479-530, 1969.

[6] K. Yosida, "Functional Analysis", Second Edition, Springer-Verlag, 1968.

[7] E.B. Dynkin, "Theory of Markov Processes", Prentice-Hall Inc., Englewood Cliffs, N.J., Pergamon Press, 1961.

[8] E.B. Dynkin, "Markov Processes", Vol. I, II , Springer-Verlag, 1965.

[9] H. Hermes and J.P. Lasalle, "Functional Analysis and Time-Optimal Control", Academic Press, New York, 1969.

[10] H. Hermes and G.W. Haynes, "On the Nonlinear Control Problem With Control Appearing Linearly", SIAM J. CONTROL, Vol. I, pp. 85-108 , 1963.

[11] G. Sansone and R. Conti, "Nonlinear Differential Equations", Macmillan (Pergamon), New York, 1964.

[12] P. Billingsley, "Convergence of Probability Measures", John Wiley, New York, 1968.

[13] K.R. Parthasarathy, "Probability Measures on Metric Spaces", Academic Press, New York, 1968.

[14] A. Strauss, "An Introduction to Optimal Control Theory", Lecture Notes in Operations Research and Mathematical Economics, No. 8, Springer-Verlag, New York, 1968.

[15] S.R.S. Varadhan, "Stochastic Processes", Courant Institute of Mathematical Sciences, New York University, 1968.

[16] J. Dugundji, "Topology", Allyn and Bacon Inc., Boston, 1966.

[17] G. Choquet, "Lectures in Analysis", Vol. I, Benjamin Press, 1969.

[18] K.L. Chung, "A Course in Probability Theory", Harcourt, Brace and World, 1968.

[19] I.I. Gikhman and A.V. Skorokhod, "Introduction to the Theory of Random Processes", W.B. Saunders Co., 1969.

[20] W. Rudin, "Real and Complex Analysis", McGraw-Hill, New York, 1966.

[21] A.V. Balakrishnan and L.W. Neustadt, eds., "Mathematical Theory of Control", Proc. Symp., University of Southern California, 1967, Academic Press, New York, 1967.

[22] W. Fleming, "Stochastic Lagrange Multipliers", in [21], p. 433.

[23] W. Fleming, "Optimal Control of Partially Observable Diffusions", SIAM J. CONTROL, Vol. 6, pp. 194-214, 1968.

[24] W. Fleming and N. Nisio, "On the Existence of Optimal Stochastic Controls", J. Math. and Mech., Vol. 15, pp. 777-794, 1966.

[25] W. Fleming, "Optimal Continuous-Parameter Stochastic Control", SIAM Review, Vol. 11, No. 4, pp. 470-500, 1969.

[26] H. Kushner, "On the Stochastic Maximum Principle: Fixed Time of Control", J. Math. Anal. Appl., Vol. 11, pp. 78-92, 1965.

[27] H. Kushner, "On the Stochastic Maximum Principle with Average Constraints", J. Math. Anal. Appl., Vol. 12, pp. 13-26, 1965.

[28] A.A. Andronov, L.S. Pontryagin and A.A. Witt, "On the Statistical Investigation of Dynamical Systems", J. Exp. Theor. Phys., Vol. 3, p. 165, 1933.

[29] R.Z. Khasminskii, "On the Stability of the Trajectories of Markov Processes", J. Appl. Math. and Mech. (PPM), Vol. 26, pp. 1554-1565, 1962.

[30] R.E. Kalman and S.E. Bertram, "Control System Analysis and Design via the Second Method of Lyapunov", J. Basic Eng. , Amer. Soc. Mech. Engrs., Vol. 82, pp. 371-393, 1960.

[31] F. Kozin, "A Survey of Stability of Stochastic Systems", Automatica, Vol. 5, pp. 95-112, 1969.

[32] N.P. Bhatia and G.P. Szegö, "Dynamical Systems: Stability Theory and Applications", Lecture Notes in Mathematics, No. 35, Springer-Verlag, New York, 1967.

[33] V.I. Zubov, "Methods of A.M. Lyapunov and Their Application", P. Noordhoff Ltd., 1964.

[34] H. Kushner, "Stability of Stochastic Control Systems", Academic Press, 1967.

[35] H. Kushner, "The Concept of Invariant Set for Stochastic Dynamical Systems and Application to Stochastic Stability", pp. 47-57 in "Stochastic Optimization and Control", Ed. by H. Karreman, John Wiley, 1968.

[36] W. Feller, "An Introduction to Probability Theory and its Applications", Vol. 11, John Wiley, 1966.

[37] E. Roxin, "Stability in General Control Systems", J. Diff. Eqns., Vol. 1, pp. 115-150, 1965.

[38] S.R.S. Varadhan, "Stochastic Processes", Courant Institute of Mathematical Sciences, New York, 1968.

[39] N.P. Bhatia and O. Hajek, "Local Semi-Dynamical Systems", Vol. 90, "Lecture Notes in Mathematics", Springer-Verlag, New York, 1969.

[40] W. Feller, "The Parabolic Differential Equations and the Associated Semi-Groups of Transformations", Ann. Math., Vol. 55, pp. 468-519, 1952.

[41] R.S. Phillips, "The Adjoint Semi-Group", Pacific Jour. Math., pp. 269-283, 1954.

[42] P.R. Halmos, "Measure Theory", Van Nostrand, 1950.

[43] V.E. Benes, "Finite Regular Invariant Measures for Feller Processes", J. Appl. Prob., Vol. 5, pp. 203-209, 1968.

[44] L. Markus, "Controllability of Nonlinear Processes", J. SIAM CONTROL, Ser. A., Vol. 3, No. 1, pp. 78-90, 1965.

[45] W. Rudin, "Principles of Mathematical Analysis", Second Edition, McGraw-Hill, New York, 1964.

[46] J.F.C. Kingman and S.J. Taylor, "Introduction to Measure and Probability", Cambridge University Press, 1966.

[47] H. Tanaka, "Existence of Diffusions with Continuous Coefficients", Mem. Fac. Sci. Kyushu Univ. Ser. A., Vol. 18, pp. 89-103, 1964.

[48] N.V. Krylov, "On Quasi-Diffusions", Theor. Prob. Appl., Vol. 11, pp. 373-389, 1966.

[49] H.P. McKean Jr., "Stochastic Integrals", Academic Press, New York, 1969.

[50] R.E. Mortensen, "Stochastic Optimal Control with Noisy Observations", Int. J. Control, Vol. 4, No. 5, pp. 455-464, 1966.

[51] H.O. Fattorini, "A Remark on the "Bang-Bang Principle for Linear Control Systems in Infinite-Dimensional Space", SIAM J. CONTROL, Vol. 6, No. 1, 1968.

[52] V.R. Vinokurov, "Optimal Control of Processes Described by Integral Equations", I, SIAM J. CONTROL, Vol. 7, No. 2, pp. 324-336, 1969; II, SIAM J. CONTROL, Vol. 7, No. 2, pp. 337-345, 1969; III, SIAM J. CONTROL, Vol. 7, No. 2, pp. 346-355, 1969.

[53] A.V. Balakrishnan, "Optimal Control Problems in Banach Spaces", J. SIAM CONTROL, Ser. A, Vol. 3, No. 1, pp. 152-180, 1965.

[54] R.S. Phelps, "Weak* Support Points of Convex Sets in E^*", Israel Jour. of Math., Vol. 2, 1964.

[55] J.L. Lions, "Contrôle Optimal de Systèmes Gouvernés par des Equations aux Derivées Partielles", Dunod Gauthier-Villars, Paris, 1968.

[56] A.G. Butkovsky, "Distributed Control Systems", American Elsevier Publishing Company, Inc., New York, 1969.

[57] A.V. Skorokhod, "Studies in the Theory of Random Processes", Addison-Wesley Publishing Co., 1965.