December 1986 LIDS-P-1633

THE PROBABILISTIC STRUCTURE OF CONTROLLED DIFFUSION PROCESSES

by

Vivek S. Borkar

Tata Inst. of Fundamental Research
Bangalore Center, P.O. Box 1234
Bangalore 560012, India
and
Laboratory for Information and Decision Systems
M.I.T.
Cambridge, MA 02139, U.S.A.

ABSTRACT

This paper surveys those aspects of controlled diffusion processes wherein the control problem is treated as an optimization problem on a set of probability measures on the path space. This includes: (i) existence results for optimal admissible or Markov controls (both in nondegenerate and degenerate cases), (ii) a probabilistic treatment of the dynamic programming principle, (iii) the corresponding results for control under partial observations, (iv) a probabilistic approach to the ergodic control problem. The paper is expository in nature and aims at giving a unified treatment of several old and new results that evolve around certain central ideas.

KEY WORDS: stochastic optimal control, controlled diffusion processes, existence of optimal controls, Markov controls, dynamic programming principle, control under partial observation, ergodic control.

Research supported by ARO contract no. DAAG-84-K-0005 and AFOSR 85-0227.

TABLE OF CONTENTS

		Page
I.	INTRODUCTION	3
II.	CONTROLLED DIFFUSION PROCESSES	7
III.	COMPACTNESS OF LAWS	14
IV.	EXISTENCE OF OPTIMAL CONTROLS	20
V.	THE DYNAMIC PROGRAMMING PRINCIPLE	26
VI.	THE DEGENERATE CASE	34
VII.	CONTROL UNDER PARTIAL OBSERVATIONS	40
VIII.	THE ERGODIC CONTROL PROBLEM	48
IX.	SOME OPEN PROBLEMS	54
APPENDIX		55
REFERENCES		57

I. INTRODUCTION

This paper attempts to bring into sharp focus a circle of ideas in controlled diffusion processes that has evolved over the last dozen years or so and give a unified exposition thereof. The central characteristic of this circle of ideas is that they view the control problem as an optimization problem on an appropriate set of probability measures and the principal tools are weak convergence and selection theorems. The choice of the title is intended to emphasize the contrast between this and the more common, largely analytic approach as typified by [6], [38] which uses the dynamic programming principle and Hamilton-Jacobi-Bellman equation as the starting point. (This classification is admittedly crude, as there is a lot of grey area in between. Also, the two viewpoints are complementary and not alternative. Neither of them replaces the other.)

The modern probabilistic approach dates back to early seventies when works like [4], [27] introduced the concept of a weak solution of a stochastic differential equation via Girsanov theorem to the control community and formulated the control problem as an optimization problem on a set of probability measures. The initial thrust [4], [13], [27] was to consider control problems where elements of this set were absolutely continuous with respect to a base measure and prove existence of optimal controls by proving the $\sigma(L_1,L_\infty)$ compactness of the corresponding Radon-Nikodym derivatives. Soon after, two probabilistic abstractions of the dynamic programming principle emerged - the martingale approach [25], [57], [59] and the nonlinear semigroup approach [54], [55], [56]. In parallel with this, much work was done on the stochastic maximum principle [7], [12],

[31], [40], and related existence results [23].

Weak convergence techniques were first used in [48] to prove an existence result akin to that of [4], [27]. Note that $\sigma(L_1,L_\infty)$ compactness of Radon-Nikodym derivatives implies the weak compactness of the corresponding probability measures by the Dunford-Pettis compactness criterion ([53], pp. 17). Since this implication goes one way, weak convergence was potentially a more flexible tool, a fact that was borne out by later developments in the degenerate case and control under partial observations. This paper traces these developments up to recent times.

The plan of the paper is as follows.

Section II describes the basic paradigm under scrutiny viz. a controlled stochastic differential equation, and discusses typical classes of controls and costs that will be of interest to us. Here and throughout the rest of the paper, we trade generality for simplicity in the sense that we work with stronger assumptions than what are strictly necessary, in order to simplify the exposition (e.g., the boundedness assumption on the coefficients of (2.1) can be relaxed, the diffusion coefficient σ can be allowed to depend explicitly on the control for many of the results and so on).

Section III establishes the basic compactness results for probability laws under various classes of controls. These are gleaned from [48], [65], though our proofs differ.

Section IV proves that the Markov controls i.e. controls that depend only on the current value of the state is a sufficiently rich class for certain costs under nondegeneracy hypothesis. This section is in the spirit of [18], [20].

Section V uses the foregoing to derive the dynamic programming principle and in the nondegenerate case, the Hamilton-Jacobi-Bellman equation. The approach is essentially probabilistic. Although it is not as economical as the direct analytical approach of [6] for the H.J.B. equation, it offers a different vantage point and establishes a link between the probabilistic and the analytic methods.

Section VI establishes the existence of an optimal Markov control in the degenerate case using the idea of Krylov's Markov selections ([61], Ch. 12). This section is based on [29], [43].

Section VII surveys the problem of control under partial observations. Given the large scope of this section, we are rather brief about each specific topic and work mostly by analogy with the 'complete observation' case studied in Sections III-VI.

Section VIII briefly outlines a probabilistic approach to the ergodic control problem based on a characterization of the a.s. limit points of normalized occupation measures for the joint state and control process. This is based on [21].

Section IX concludes with a short list of some open problems.

A few important disclaimers: This paper does not survey all aspects of controlled diffusions that would legitimately qualify as a part of the 'probabilistic structure'. Some major omissions are: control problems involving optimization over stopping times (optimal stopping, impulse control etc.), stochastic maximum principle, singular control, approximation and robustness issues etc. Also, the bibliography is meant to be only representative and not exhaustive. The surveys [7], [9], [24], [28] can be used to complement the present one in these respects. See also Ch. 16-18 of [32].

Finally a word on notation: For a Polish space (i.e., separable and metrizable with a complete metric) X, P(X) denotes the Polish space of probability measures on X with the topology of weak convergence [11]. C(X;Y) denotes the space of continuous maps $X \to Y$ (Y a complete metric space) with the topology of uniform convergence on compacts. C(X) stands for C(X,R) and $C_b(X)$ for the subset of C(X) consisting of bounded functions. By the natural filtration of a stochastic process Y(t), $t \geq 0$, we shall always mean the filtration $\{F_t\}$ where F_t is the completion w.r.t. the underlying probability measure of the σ-field $\bigcap_{s>t} \sigma(Y(y), y \leq s)$ for $t \geq 0$.

Remark. A familiarity with diffusion theory at the level of the well-known texts by Ikeda-Watanabe or Stroock-Varadhan (references [44] and [61] resp.) is assumed throughout. It should be remarked that much of what follows can be recast in the elegant language of 'martingale problems' introduced in [61], as has been done in [28]-[30]. However, we do not do so for sake of simplicity.

II. CONTROLLED DIFFUSION PROCESSES

The prototypical controlled diffusion we consider is the R^d-valued process $X(\cdot)$ satisfying

$$dX(t) = m(X(t), u(t))dt + \sigma(X(t))dW(t), \; X(0) = X_0, \tag{2.1}$$

where,

(i) $m(\cdot,\cdot) = [m_1(\cdot,\cdot),\ldots,m_d(\cdot,\cdot)]^T : R^d x U \to R^d$ (U being a prescribed compact metric space) is bounded continuous and $m(\cdot,u)$ is Lipschitz uniformly in u,

(ii) $\sigma(\cdot) = [[\sigma_{ij}(\cdot)]] : R^d \to R^{dxd}$ is bounded Lipschitz,

(iii) X_0 is a random variable in R^d with a prescribed law π_0,

(iv) $W(\cdot) = [W_1(\cdot),\ldots,W_d(\cdot)]^T$ is a d-dimensional standard Wiener process independent of X_0, and,

(v) $u(\cdot)$ is a U-valued process (called an 'admissible' control) satisfying: $u(y)$ is independent of $W(t)-W(s)$ for $t \geq s \geq y$.

We distinguish between the nondegenerate case where the least eigenvalue of $\sigma\sigma^T(\cdot)$ is uniformly bounded away from zero and the degenerate case when it is not. In either case, $X(\cdot)$ above can be constructed by Picard iterations as in [53], Section 4.4, given $W(\cdot)$, $u(\cdot)$ on some probability space.

If $u(\cdot)$ is adapted to the natural filtration of $X(\cdot)$, call it a feedback control. Further subclasses of this are Markov controls when $u(\cdot) = v(X(\cdot), \cdot)$ for some measurable $v: R^d x R^+ \to U$ and stationary Markov controls when $u(\cdot) = v(X(\cdot))$ for some measurable $v: R^d \to U$. By abuse of notation, one often refers to the map v itself as the Markov (resp.

stationary Markov) control. For feedback and Markov controls, one cannot obtain a solution of (2.1) by mere Picard iterations unless further strong conditions are imposed on the nature of dependence of u(•) on X(•). These are usually too stringent for control applications and hence one has to seek other proofs of existence or uniqueness and at times, other solution concepts. We shall comment more on this later on in this section.

A control problem is the problem of minimizing the expectation of a prescribed functional of X(•), u(•), called the cost functional, over a prescribed set of admissible controls. Typical cost functionals are:

(C1) $E[F(X(\cdot))],\ F \in C_b(C([0,\infty); R^d))$,

(C2) $E[\int_0^\tau k(X(t), u(t))dt + h(X(\tau))],\ \tau = \inf\{t \geq 0 \mid X(t) \notin \bar{G}\}$
for some bounded connected open $G \subset R^d$ with a C^2 boundary δG and $k \in C_b(\bar{G} x U)$, $h \in C_b(\delta G)$,

(C3) $E[\int_0^T e^{-\alpha t} k(X(t), u(t))dt + e^{-\alpha T} h(X(T))],\ \alpha \in [0,1],\ T \in [0,\infty]$

with $\alpha > 0$ if $T=\infty$, $k \in C_b(R^d x U)$, $h \in C_b(R^d)$.

(A different and rather special cost functional is considered in Section VIII).

A control u(•) for which the minimum of the cost functional is attained will be said to be an optimal control and the corresponding X(•) referred to as the optimal process or the optimal solution to (2.1).

Throughout this paper, we assume the relaxed control framework [34], i.e., we assume that U = **P**(S) for a compact metric space S and there exists

$\bar{m}(\cdot,\cdot) = [\bar{m}_1(\cdot,\cdot),\dots,\bar{m}_d(\cdot,\cdot)]^T : R^d x S \to R^d$ which is bounded continuous with $\bar{m}(\cdot,s)$ Lipschitz uniformly in s, such that

$$m_i(x,u) = \int_S \bar{m}_i(x,s)u(ds), \quad 1 \le i \le d.$$

In addition, if the cost functional is (C2) or (C3), we assume that $k(x,u) = \int \bar{k}(x,s)u(ds)$ for some $k \in C_b(\bar{G} x S)$ or $C_b(R^d x S)$ as the case may be. If $u(\cdot)$ is always concentrated on the subset U_D of U consisting of Dirac measures (itself a compact set), we call it a precise control. (A 'control' will always mean a relaxed control.) If $u(\cdot)$ is a precise control, a straightforward application of the selection theorem in the Appendix (henceforth called 'the selection theorem') shows that

$$m(X(\cdot),\ u(\cdot)) = \bar{m}(X(\cdot),\ s(\cdot)) \qquad (2.2)$$

for an S-valued process $s(\cdot)$ adapted to the natural filtration of $(X(\cdot), u(\cdot))$. The selection theorem can again be employed to show that $s(\cdot)$ will inherit the feedback, Markov or stationary Markov nature of $u(\cdot)$. If $\bar{m}(x,S)$ is convex for all x, (2.2) above is always possible. In any case, the following holds:

Theorem 2.1. Let $X(\cdot)$, $u(\cdot)$, $W(\cdot)$ satisfy (2.1) on a probability space $(\Omega,\mathbf{F},P)$. Then there exists a sequence $\{u_n(\cdot)\}$ of precise controls on $(\Omega,\mathbf{F},P)$ such that if $\{X^n(\cdot)\}$ denote the corresponding solutions of (2.1), then for each $T>0$ and $f \in C_b([0,T] x S)$,

$$\int_0^T\int_S f(t,s)u_n(t)(ds)dt \to \int_0^T\int_S f(t,s)u(t)(ds)dt \text{ on } \Omega \tag{2.3}$$

$$E[\sup_{t\in[0,T]} ||X^n(t) - X(t)||^2] \to 0. \tag{2.4}$$

Proof. (Sketch) W.l.o.g., let T=1. Construct $\{u_n(\cdot)\}$ from $u(\cdot)$ as in [2], pp. 32-33 with the extra proviso that t_i be the left end point of T_i (in the notation of [2]) for each i. A standard argument using the Gronwall inequality shows that

$$E[\sup_{t\in[0,1]} ||X^n(t)-X(t)||^2] \leq K\, E[||\int_0^1 (m(X(t),u(t))-m(X(t),u_n(t)))dt||^2] \tag{2.5}$$

for some K>0. By the results of [2], pp.33, (2.3) holds and the r.h.s. of (2.5) tends to zero as $n \to\infty$. Q.E.D.

Corollary 2.1. For each of the cost functionals considered above, the infima over precise and relaxed controls coincide.

The proof is omitted and will be self-evident by the end of the next section.

There are two ways of posing a control problem. In the strong formulation, $(\Omega,\mathbf{F},P)$, X_0, $W(\cdot)$ are prescribed and one optimizes over a specified class of $u(\cdot)$ on $(\Omega,\mathbf{F},P)$. In the weak formulation, one optimizes over all collections $(\Omega,\mathbf{F},P,W(\cdot)$, $X(\cdot)$, $u(\cdot)$, $X_0)$ as above with $u(\cdot)$ satisfying some prescribed conditions. If $u(\cdot)$ is a feedback control, (2.1) has only a weak solution in general even in the nondegenerate case [4] thus forcing a weak formulation. (It should be mentioned that for the special case of Markov controls, it does have a unique strong solution in the nondegenerate case [63]. In the degenerate case, the existence or uniqueness of even a weak solution is not guaranteed.) We use the weak formulation throughout. The following result shows that not much is lost.

Theorem 2.2. Let (Ω,F,P), X_0, $W(\cdot)$ be given.

(a) Let $u(\cdot)$ be an admissible control on $(\Omega,\mathbf{F},P)$ and $X(\cdot)$ the corresponding solution of (2.1). Then $X(\cdot)$ also satisfies (2.1) with $u(\cdot)$, $W(\cdot)$ replaced by $\tilde{u}(\cdot)$, $\tilde{W}(\cdot)$ where $\tilde{u}(\cdot)$ is a feedback control and $\tilde{W}(\cdot)$ is a d-dimensional standard Wiener process independent of X_0, on a possibly augmented probability space.

(b) Let $\bar{X}(\cdot)$, $\bar{W}(\cdot)$, $\bar{u}(\cdot)$, $\bar{X}_0$ satisfy (2.1) on a probability space $(\bar{\Omega},\bar{\mathbf{F}},\bar{P})$ with $\bar{u}(\cdot)$ a feedback control, such that the laws of $(\bar{W}(\cdot), \bar{X}_0)$ and $(W(\cdot), X_0)$ coincide. Then by augmenting $(\Omega,\mathbf{F},P)$ if necessary, one can construct on it a process $X(\cdot)$ satisfying (2.1) with the prescribed $W(\cdot)$, X_0 and a feedback control $u(\cdot)$ which has the same dependence on $X(\cdot)$ as what $\bar{u}(\cdot)$ had on $\bar{X}(\cdot)$.

Proof. (a) Define $\tilde{u}(\cdot)$ by

$$\int f d\tilde{u}(t) = E[\int f du(t)/X(s),\ s\leq t]\ \text{a.s.},\ t\geq 0,\ f\varepsilon C(S) \tag{2.6}$$

picking a measurable version thereof. Write

$$X(t) = X_0 + \int_0^t m(X(s), \tilde{u}(s))ds + \tilde{M}(t)$$

for the appropriately defined $\tilde{M}(\cdot)$. The results of [64] show that $\tilde{M}(\cdot)$ must be of the form

$$\tilde{M}(t) = \int_0^t \sigma(X(s))d\tilde{W}(s)$$

for some d-dimensional Wiener process $\tilde{W}(\cdot)$, on a possibly augmented probability space.

(b) Let $Q \varepsilon \mathbf{P}(R^d \times C([0,\infty); R^d) \times C([0,\infty); R^d))$ denote the law of $(\bar{X}_0, \bar{W}(\cdot), \bar{X}(\cdot))$. Disintegrate it as

$$Q(dw_1, dw_2, dw_3) = Q_1(dw_1, dw_2)Q_2(w_1,w_2)(dw_3)$$

where Q_1 is the law of $(\bar{X}_0, \bar{W}(\cdot))$ and Q_2 is the regular conditional law of $\bar{X}(\cdot)$ given $(\bar{X}_0, \bar{W}(\cdot))$. Augment Ω to $\Omega' = \Omega x C([0,\infty); R^d)$, F to F' = the product σ-field on Ω' and replace P by P' defined as follows: For $A \varepsilon F$, B Borel in $C([0,\infty); R^d)$,

$$P'(AxB) = E[Q_2(X_0,W(\cdot))(B)I_A]$$

Define $X(\cdot)$ by

$$X(t)((w_1,w_2)) = w_2(t) \text{ for } (w_1,w_2) \varepsilon \Omega'.$$

The rest is routine. Q.E.D.

Relevant references: [4], [6], [9], [24], [28], [32], [34], [38], [47], [64].

III. COMPACTNESS OF LAWS

Let $\mathbf{A}_i \subset \mathbf{P}(C([0,\infty); R^d))$, i=1,2,3, denote resp. the set of laws of $X(\cdot)$ under all admissible/Markov/stationary Markov controls. We prove below that $\mathbf{A}_1$ and in the nondegenerate case, $\mathbf{A}_2$, $\mathbf{A}_3$ are compact. Clearly, it suffices to replace $[0,\infty)$ above by $[0,T]$ for arbitrary $T>0$.

Let $(X^n(\cdot), W^n(\cdot), u^n(\cdot), X_0^n)$, n=1,2,..., satisfy (2.1) on probability spaces $(\Omega^n, \mathbf{F}^n, P^n)$ resp. Let $\{f_j\}$ be countable dense in the unit ball of $C(S)$ and define $\alpha_j^n(t) = \int f_j du^n(t)$, $t \in [0,T]$. Let B denote the closed unit ball of $L_2[0,T]$ with the topology = the weak topology of $L_2[0,T]$ relativized to B. Let D be a countable product of replicas of B with product topology. B,D are compact metrizable. Let $\alpha^n(\cdot) = [\alpha_1^n(\cdot), \alpha_2^n(\cdot),\ldots]$, n=1,2,..., viewed as D-valued random variables.

Using the estimate of Lemma 4.12, pp. 125, [53], one can show that

$$E[||X^n(t_2) - X^n(t_1)||^4] \leq K|t_2-t_1|^2, \; n\geq 1; \; t_1, t_2 \in [0,T],$$

for some T-dependent $K>0$. The criterion of [11], pp. 95, implies that the laws of $\{X^n(\cdot)\}$ are tight in $\mathbf{P}(C([0,T]; R^d))$. Since D is compact, the laws of $(X^n(\cdot), \alpha^n(\cdot))$, $n\geq 1$, are then tight in $\mathbf{P}(C([0,T]; R^d)\times D)$ and hence $(X^n(\cdot), \alpha^n(\cdot))$ converge in law along a subsequence of $\{n\}$ (denoted $\{n\}$ again) to a limit $(X(\cdot), \alpha(\cdot))$. By Skorohod's theorem ([44], pp.9), we may assume that these are defined on a common probability space $(\Omega,\mathbf{F},P)$ and the above convergence holds for all sample points outside a set N with $P(N)=0$. Write $\alpha(\cdot) = [\alpha_1(\cdot), \alpha_2(\cdot),\ldots]$.

<u>Lemma 3.1</u>. There exists a U-valued process $u(\cdot)$ such that

$$\alpha_i(t) = \int f_i du(t), \; i \geq 1.$$

<u>Proof</u>. Fix a sample point $\omega \varepsilon N$. Let $n(1) = 1$ and define $\{n(k)\}$ inductively to satisfy

$$\sum_{j=1}^{\infty} 2^{-j} \max_{1 \leq \ell < k} \left| \int_0^T (\alpha_j^{n(k)}(t) - \alpha_j(t))(\alpha_j^{n(\ell)}(t) - \alpha_j(t))dt \right| < \frac{1}{k},$$

which is possible because $\alpha_j^n(\cdot) \to \alpha_j(\cdot)$ in B. Argue as in the proof of Theorem 1.8.4, pp. 29-30, [2] (Banach-Saks theorem) to conclude that for each j,

$$\frac{1}{m} \sum_{k=1}^{m} \alpha_j^{n(k)}(\cdot) \to \alpha_j(\cdot)$$

strongly in $L_2[0,T]$ and hence a.e. along a subsequence. By a diagonal argument, we may extract a subsequence $\{m_k\}$ of $\{m\}$ such that for t outside a set $M \subset [0,T]$ of zero Lebesgue measure,

$$\frac{1}{m_k} \sum_{\ell=1}^{m_k} \alpha_j^{n(\ell)}(t) \to \alpha_j(t) \quad \forall j. \tag{3.1}$$

Define $v_k(t) \varepsilon U$ by

$$\int f_i dv_k(t) = \frac{1}{m_k} \sum_{\ell=1}^{m_k} \alpha_i^{n(\ell)}(t), \qquad i=1,2,\ldots,$$

for k=1,2,... Fix $t \varepsilon M$. By (3.1), any limit point v(t) of $\{v_k(t)\}$ must satisfy

$$\int f_i dv(t) = \alpha_i(t). \tag{3.2}$$

Now the map

$$\phi : v \varepsilon U \to [\int f_1 dv, \int f_2 dv, \ldots] \varepsilon [-1,1]^{\infty}$$

is a diffeomorphism between U and $\phi(U)$ (See, e.g., Lemma 2.6, pp. 111, [15]). By (3.2), $\alpha(\cdot) \varepsilon \phi(U)$ a.s. where the 'a.s.' may be dropped by an appropriate choice of the version. Define $u(\cdot) = \phi^{-1}(\alpha(\cdot))$. Q.E.D.

Lemma 3.2. For any measurable $f:\Omega x[0,T]xS \to R$ such that $f(w,\cdot,\cdot)$ is continuous for each w,

$$\int_0^t \int_S f(w,s,v)u^n(s)(dv)ds \to \int_0^t \int_S f(w,s,v)u(s)(dv)ds \qquad \text{a.s.,} \quad t \in [0,T].$$

Proof. It clearly suffices to prove this for f of the type $f(w,s,v) = \sum_{\ell=1}^{n} g_\ell(s)f_\ell(v)$, $s \in [0,T]$, $v \in S$, $n \in \mathbf{N}$, $g_\ell \in C[0,T]$, $1 \le \ell \le n$. But then it is immediate from the convergence of $\alpha^n(\cdot)$ to $\alpha(\cdot)$ in D. Q.E.D.

Lemma 3.3. $X(\cdot)$ satisfies (2.1) for some $\tilde{W}(\cdot)$, $\tilde{X}_0$ and $\tilde{u}(\cdot)$.

Proof. Let $f \in C^2(R^d)$ with compact support and write

$$(Lf)(x,u) = \sum_i m_i(x,u) \frac{\partial f}{\partial x_i} + \frac{1}{2} \sum_{i,j,k} \sigma_{ik}(x)\sigma_{jk}(x) \frac{\partial^2 f}{\partial x_i \partial x_j} .$$

Let $0 \le t_1 < t_2 < \ldots < t_m < s < t \le T$ and $g \in C_b((R^d)^m)$. Then for $n \ge 1$,

$$E[(f(X^n(t)) - f(X^n(s)) - \int_s^t (Lf)(X^n(y), u^n(y))dy)g(X^n(t_1),\ldots,X^n(t_m))]=0$$

Letting $n \to \infty$ and using the preceding lemma,

$$E[(f(X(t)) - f(X(s)) - \int_s^t (Lf)(X(y), u(y))dy)g(X(t_1),\ldots,X(t_m))] = 0$$

We can replace $u(\cdot)$ here by $\tilde{u}(\cdot)$ defined as in (2.6). Since g is arbitrary, it follows that $f(X(t)) - \int_0^t (Lf)(X(y),\tilde{u}(y))dy$ is a martingale w.r.t. the natural filtration of $X(\cdot)$. The rest follows from martingale representation theorems of [64] by standard arguments. Q.E.D.

Theorem 3.1. $\mathbf{A}_1$ is compact and in the nondegenerate case, convex.

Proof. Compactness is proved above. Assume nondegeneracy. If $u_1(\cdot)$, $u_2(\cdot)$ are feedback controls, the laws of the corresponding solution to (2.1) on [0,T] are absolutely continuous w.r.t. the law of $Y(\cdot)$ satisfying

$$Y(t) = X_0 + \int_0^t \sigma(Y(s))dW(s)$$

with the Radon-Nikodyn derivatives being $\Lambda_i(T)$, i=1,2, where $\Lambda_i(t)$ is the unique solution to

$$\Lambda_i(t) = 1 + \int_0^t \Lambda_i(s) \langle m(Y(s), u_i(s)), dY(s)\rangle, \ i=1,2. \qquad (3.3)$$

Convexity follows if we show that for $\alpha \in [0,1]$, $\Lambda(t) = \alpha\Lambda_1(t) + (1-\alpha)\Lambda_2(t)$ also satisfies (3.3) with some $u(\cdot)$. This is indeed so for $u(t) = \beta(t)u_1(t) + (1-\beta(t))u_2(t)$ with $\beta(t) = \alpha\Lambda_1(t)/\Lambda(t)$. Q.E.D.

Theorem 3.2. In the nondegenerate case, $\mathbf{A}_2$, $\mathbf{A}_3$ are compact.

Proof. Let $\{u^n(\cdot)\}$ above be Markov and let $T^n_{s,t}$, $t \geq s$, denote the corresponding transition semigroup. In the set-up of Lemma 3.3,

$$E[(f(X^n(t)) - T^n_{s,t}f(X^n(s)))g(X^n(t_1),\ldots,X^n(t_m))] = 0, \; n=1,2,\ldots \tag{3.4}$$

For each n, $T^n_{\cdot,t}f(\cdot)$ satisfies the backward Kolmogorov equation. From standard p.d.e. theory (see [50] or [66], pp. 133-134) it follows that $T^n_{s,t}f(\cdot)$, n=1,2,..., are equicontinuous. Let $T_{s,t}f(\cdot)$ be a limit point of the same in $C(R^d)$. Passing to the limit in (3.4) along an appropriate subsequence, we get

$$E[(f(X(t)) - T_{s,t}f(X(\cdot)))g(X(t_1),\ldots,X(t_m))] = 0,$$

implying that $X(\cdot)$ is Markov. That this implies that $u(\cdot)$ is a Markov control follows by a straightforward application of the selection theorem as in [43], pp. 184-5. Thus $\mathbf{A}_2$ is compact. Compactness of $\mathbf{A}_3$ follows on noting that if $T^n_{s,t}f$ depends on s,t only through t-s, so will $T_{s,t}f$. Q.E.D.

<u>Relevant References:</u> [3], [4], [13], [23], [27], [29], [48], [65].

IV. EXISTENCE OF OPTIMAL CONTROLS

From Theorems 3.1, 3.2 and Lemma 3.2, it follows that for cost functionals C1, C3, the minimum is attained on A_1 and in the nondegenerate case, on A_2, A_3 as well.

Theorem 4.1. In the nondegenerate case, the minimum for C2 on A_i, i=1,2,3 is attained.

Proof. In the set-up of the preceding section, let τ^n, τ be the first exit times from $\bar{G}$ for $X^n(\cdot)$, $X(\cdot)$ resp. and $\sigma = \inf\{t\geq 0 | X(t)\in\delta G\}$. Simple geometric considerations show that for each $\omega\notin N$, any limit point of $\{\tau^n\}$ in $[0,\infty]$ must lie in $[\sigma,\tau]$. Since δG is C^2 and $X(\cdot)$ nondegenerate, $\sigma=\tau$ a.s. Thus $\tau^n \to \tau$ a.s. and the claim will follow from Lemma 3.2 if we prove $\{\tau^n\}$ to be uniformly integrable. The latter can be proved by establishing a bound of the type

$$P(\tau>t) \leq Ke^{-\lambda t}, \quad t\geq 0, \tag{4.1}$$

for some K, $\lambda>0$ uniformly in $u(\cdot)$. (See [19] for details.) Q.E.D.

Remarks: The nondegeneracy can be dropped by insisting that the set $(\delta G)' = \{x\in\delta G | P(\tau>0/X(0)=x) > 0\}$ (= δG in the nondegenerate case) be independent of $u(\cdot)$ and closed in δG, and τ remain uniformly integrable over all $u(\cdot)$.

Theorem 4.2. In all the situations considered above, the subset of the appropriate A_i where the minimum in question is attained is compact and nonempty.

This is obvious in view of the foregoing.

Theorem 4.3. In the nondegenerate case, the minima of C2 over A_1, A_2, A_3 coincide and the minima of C3 over A_1, A_2 coincide.

We shall prove the case of C2 only. W.ℓ.o.g., let $X_0=x \in G$. Define a measure η_x on $\bar{G}xS$ by

$$\int f d\eta_x = E[\int_0^{\tau}\int_S f(X(t), y)u(t)(dy)dt], \quad f \in C(\bar{G}xS),$$

which clearly depends on u(•). Disintegrate η_x as

$$\eta_x(dy,ds) = \nu_x(dy)v(y)(ds)$$

where ν_x is the image of η_x under the projection $\bar{G}xS \to \bar{G}$ (called the Green measure) and $v:\bar{G} \to U$ is the regular conditional law. Let u'(•) denote the stationary Markov control v(X'(•)) where X'(•) denotes the corresponding solution of (2.1) with $X_0=x$.

Remarks: Note that v above is defined only ν_x-a.s. We pick any one representative of this a.s.-equivalence class, it does not matter which. A similar remark applies to other situations in this section and Section VIII where we employ a similar disintegration of measures.

Lemma 4.1. For $f \in C(\bar{G}xS)$, $h \in C(\delta G)$, the quantities

$$E\left[\int_0^\tau\int_S f(X(t),y)u(t)(dy)dt\right],\ E[h(X(\tau))]$$

remain unchanged if $X(\cdot)$, $u(\cdot)$ are replaced by $X'(\cdot)$, $u'(\cdot)$ resp.

Proof. Let

$$\varphi(x) = E\left[\int_0^\tau\int_S f(X'(t),y)u'(t)(dy)dt\right]$$

$$Y(t) = \int_0^t\int_S f(X(s),y)u(s)(dy)ds + \varphi(X(t)).$$

Then $\varphi(\cdot)$ is the unique solution in $W^{2,p}(G)\cap C(\bar{G})$, $p\geq 2$,

$$-(L\varphi)(x,v(x)) = \int_S f(x,y)v(x)(dy) \text{ on } G, \quad \varphi = 0 \text{ on } \delta G. \tag{4.2}$$

(see [6]). By the extended Ito formula of [47], Ch. 2,

$$E[Y(\tau)] - E[Y(0)] = \int f d\eta_x - \varphi(x)$$

$$= E\left[\int_0^\tau \left(\int_S f(X(t),y)u(t)(dy) + (L\varphi)(X(t), v(X(t))\right)dt\right]$$

$$= 0$$

by (4.2) and the definition of v. The claim for f follows. That for h can be reduced to the above by Ito formula if h is the restriction of a C^2 function. The general case then follows by an obvious approximation argument. Q.E.D.

The proof of the first half of Theorem 4.3 is immediate. The second half can be proved by analogous methods. We briefly indicate the line of argument for $T<\infty$, $\alpha=0$. Define a measure γ_x on $R^d x[0,T]xS$ by

$$\int f d\gamma_x = E\left[\int_0^T\int_S f(X(t),t,y)u(t)(dy)dt\right] \text{ for } f \in C_b(R^d x[0,T]xS)$$

and disintegrate it as

$$\gamma_x(ds,dt,dy) = \beta_x(dx,dt)v(x,t)(dy)$$

where β_x is the image of γ_x under the projection $R^d x[0,T]xS \rightarrow R^d x[0,T]$ and $v:R^d x[0,T] \rightarrow U$ is the regular conditional law. Let $u'(\cdot) = v(X'(\cdot),\cdot)$ with $X'(\cdot)$ the corresponding solution of (2.1). Define

$$\varphi(x,t) = E[\int_t^T k(X'(t),u'(s))ds + h(X'(T))/X'(t) = x]$$

$$Y(t) = \int_0^t k(X(s),u(s))ds + \varphi(X(t),t)$$

By arguments analogous to those of Lemma 4.1, one can show that $E[Y(T)] = E[Y(0)]$, implying that $u(\cdot)$, $u'(\cdot)$ yield the same cost. The same proof also shows [18] that the laws of $X(t)$, $X'(t)$ agree for each t. This allows us to prove the existence of optimal Markov controls for a larger class of cost functionals. An example is

$$\int_0^T ||\mu(t)-\mu||dt, \qquad T\geq 0,$$

where $\mu(t)$ = the law of $X(t)$, $\mu \in P(R^d)$ is prescribed and $||\cdot||$ is the total variation norm. Only thing one needs to observe here is that the map

$$\eta \in P(R^d) \to ||\eta-\mu|| \in R$$

is lower semicontinuous.

Such claims are in general false for C1 as the following example shows: Let $d=1$, $S=\{-1,1\}$, $\sigma(\cdot) \equiv 1$, $\bar{m}(x,y) = y$ for all x, and the cost functional is $E[f(X(T))f(X(0))]$ where $f \in C_b(R)$ is given to be monotone increasing and odd. Using the comparison theorem for one dimensional Ito processes [44], one can see that the precise control $s(\cdot) = I\{X(0)<0\} - I\{X(0)\geq 0\}$, which is clearly not Markov, is optimal and does better than any Markov control.

Continuing with the nondegenerate case, one can say more. The set M_x of attainable ν_x, $x \varepsilon G$, is the same whether we consider all admissible controls or just the stationary Markov controls, by virtue of Lemma 4.1. Letting $M(\bar{G})$ = the space of finite measures on $\bar{G}$ with weak* topology, it follows by arguments analogous to those of Theorem 4.1 that M_x is compact in $M(\bar{G})$. By the Krylov inequality ([47], Section 2.2), each ν_x has a density $g(x,\cdot)$ w.r.t. the Lebesgue measure on $\bar{G}$, defined a.e. From (4.2), it follows that $g(x,\cdot)$ satisfies $Ag(x,\cdot) = 0$ on $G\backslash\{x\}$ in the sense of distributions, where A is the formal adjoint of the extended generator of the Markov process under consideration. Standard p.d.e. estimates can then be used [20] to prove that $g(x,\cdot)$ remains equicontinuous bounded on compact subsets of $G\backslash\{x\}$. It is not hard to deduce from this that the attainable $g(x,\cdot)$ form a compact set C_x in $C(G\backslash\{x\})$ and the bijection $\nu_x \longleftrightarrow g(x,\cdot)$ is a homeomorphism between M_x and C_x. Scheffe's theorem [11] then implies that M_x is also compact in the total variation norm topology.

These considerations allow one to prove existence of optimal Markov controls for more general cost functionals. An example is the functional $\int g(x,y)h(dy)$ where h is a finite measure supported on the boundary of a subdomain of G which is bounded away from x, having the interpretation as a 'boundary-crossing cost' [20].

Analogous results seem possible when τ is replaced by a fixed $T>0$.

<u>Relevant references</u>: [18], [20].

V. THE DYNAMIC PROGRAMMING PRINCIPLE

As before, we confine our attention to C2 and assume that either nondegeneracy or the condition in the remarks following Theorem 4.1 holds. Let $\{F_t\}$ denote the natural filtration of $X(\cdot)$.

<u>Lemma 5.1</u>. Let τ be an $\{F_t\}$-stopping time. Then on $\{\tau<\infty\}$, the regular conditional law of $X(\tau+\cdot)$ given F_τ is a.s. the law of some controlled diffusion of the type (2.1).

<u>Proof</u>. W. .o.g., we may assume $u(\cdot)$ to be feedback. Thus there exists a map $f:[0,\infty)xC([0,\infty);\ R^d)\rightarrow U$ which is progressively measurable w.r.t. $\{F_t\}$ such that $u(t) = f(t,X(\cdot))$ a.s. By Lemma 1.3.3, pp. 33, [61], it follows that a version of the regular conditional law of $X(\tau+\cdot)$ given F_τ is given by the law of a controlled diffusion as in (2.1) with initial condition $X(\tau)$ and control $u'(\cdot) = f(\tau+t,X(\cdot))$ with τ and the restriction of $X(\cdot)$ to $[0,\tau]$ being held fixed as parameters. The claim follows. Q.E.D.

For $x\varepsilon\bar{G}$, $t\varepsilon[0,T]$, let $J_x(u(\cdot))$ denote the cost under $u(\cdot)$ when $X_0 = x$. Define $J_x(v)$ analogously for a Markov or stationary Markov control v. For a fixed process $u(\cdot)$, an argument analogous to Theorem 4.1 can be employed to show that $x\rightarrow J_x(u(\cdot))$ is continuous. In the nondegenerate case, $x\rightarrow J_x(v)$ is continuous as a consequence of the Feller property and considerations similar to those of Theorem 4.1. Define the 'value function' $V:\bar{G}\rightarrow R$ by

$V(x) = \min J_x(u(\cdot))$. Then V will be upper semicontinuous and $V = h$ on $(\delta G)'$.

Lemma 5.2. V is continuous.

Proof. Let $x_n \to x_\infty$ in $\bar{G}$ and let $u_n(\cdot)$, $X_n(\cdot)$, $n=1,2,\ldots,$ denote resp. an optimal control and the corresponding process when $X_0 = x_n$. By mimicing the arguments of Theorem 4.1, one has (by dropping to a subsequence if necessary) $X_n(\cdot) \to X_\infty(\cdot)$ in law for some $X_\infty(\cdot)$ satisfying (2.1) with some control $u_\infty(\cdot)$ and with $X_\infty(0) = x_\infty$; and moreover,

$$V(x_n) = J_{x_n}(u_n(\cdot)) \to J_{x_\infty}(u_\infty(\cdot)).$$

For any $u(\cdot)$,

$$J_{x_n}(u(\cdot)) \to J_{x_\infty}(u(\cdot)).$$

Since

$$V(x_n) \leq J_{x_n}(u(\cdot)),$$

we have

$$J_{x_\infty}(u_\infty(\cdot)) \leq J_{x_\infty}(u(\cdot)), \text{ implying } V(x_\infty) = J_{x_\infty}(u(\cdot)).$$ Q.E.D.

<u>Theorem 5.1</u>. (The dynamic programming principle) Let $X_0 = x$. Then for any $\{F_t\}$-stopping time σ,

$$V(x) = \min_{u(\cdot)} E[\int_0^{\sigma\wedge\tau} k(X(t), u(t))dt + V(X(\sigma\wedge\tau))] \tag{5.1}$$

where the minimum is attained if and only if $u(\cdot)$ on $[0,\sigma\wedge\tau]$ is the restriction of an optimal control to $[0,\sigma\wedge\tau]$.

<u>Proof</u>. Under any $u(\cdot)$,

$$V(x) \leq E[\int_0^{\sigma\wedge\tau} k(X(t), u(t))dt] + E[\int_{\sigma\wedge\tau}^{\tau} k(X(t), u(t))dt + h(X(\tau))].$$

Picking $u(\cdot)$ on $[\sigma\wedge\tau,\tau]$ to be an optimal control for the initial condition $X(\sigma\wedge\tau)$ and taking the infimum over all such $u(\cdot)$,

$$V(x) \leq \inf_{u(\cdot)} E[\int_0^{\sigma\wedge\tau} k(X(t), u(t))dt + V(X(\sigma\wedge\tau))].$$

If $u(\cdot)$ is optimal,

$$V(x) = E[\int_0^{\sigma\wedge\tau} k(X(t), u(t))dt] + E[E[\int_{\sigma\wedge\tau}^{\tau} k(X(t), u(t))dt +$$

$$h(X(\tau))/F_{\sigma\wedge\tau}]]$$

$$\geq E[\int_0^{\sigma\wedge\tau} k(X(t), u(t))dt + V(X(\tau\wedge\sigma))]$$

by Lemma 5.1.(5.1) and the 'if' part of the last claim follow. The 'only if' part follows on noting that if $u(\cdot)$ attains the minimum in (5.1), then a control which on $[0,\sigma\wedge\tau]$ coincides with $u(\cdot)$ and thereafter coincides with a control which is optimal for the initial condition $X(\sigma\wedge\tau)$ will have cost $V(x)$ and thus be optimal. Q.E.D.

Remark (5.1) holds with 'min' replaced by 'inf' even if we drop the assumptions of the first paragraph of this section and can be proved by analogous arguments by using near-optimal controls in place of optimal ones. However, V need no longer be continuous.

Corollary 5.1. In the nondegenerate case, if a stationary Markov control v is optimal for some initial x, it is optimal for any initial condition.

Proof. Let σ above be the first exit time from a connected subdomain $A \subset G$ with $x \in A$ and having a C^2 boundary δA. Let μ denote the law of $X(\sigma)$. The above argument then clearly shows that v must also be optimal for initial condition $X(\sigma)$ and hence $V(y) = J_y(v)$ for μ-a.e. y in δA. By the Stroock-

Varadhan support theorem, supp $\mu = \delta A$. Thus by continuity of V and $y \to J_y(v)$, $V(y) = J_y(v)$ on δA. Since δA can be arranged to contain any $y \varepsilon \bar{G}$, we are done. Q.E.D.

Corollary 5.2. In the nondegenerate case, $V \varepsilon W^{2,p}_{loc}(G) \quad C(\bar{G})$ for $p \geq 2$ and satisfies the 'Hamilton-Jacobi-Bellman' equation

$$\min_u [(LV)(x,u) + k(x,u)] = 0 \text{ a.e. in } G, \ V=h \text{ on } \delta G. \tag{5.2}$$

Furthermore, a stationary Markov control v is optimal if and only if for a.e. x, the minimum above is attained at v(x).

Proof. Let v be an optimal stationary Markov control. By standard p.d.e. results [6], there is a unique solution in the specified class to the equation

$$(LV')(x,v(x)) + k(x,v(x)) = 0 \text{ in } G, \ V' = h \text{ on } \delta G.$$

By Corollary 5.1 and the well-known stochastic representation of this solution [6], $V' = V$. Let $X(\cdot)$ be the process starting at some $y \varepsilon G$ and controlled by a constant control $u(\cdot) \equiv u_0$. By Theorem 5.2, the process $V(X(\tau \wedge t)) + \int_0^{\tau \wedge t} k(X(s), u_0)ds$ is an $\{F_t\}$-submartingale. The Ito formula of [47], Ch. 2, gives us its Doob-Meyer decomposition to be $\int_0^{\tau \wedge t}((LV)(X(s), u_0) + k(X(s), u_0))ds$ + a martingale. Thus $(LV)(X(\tau \wedge t), u_0) + k(X(\tau \wedge t), u_0) \geq 0$ a.s. for all $t \geq 0$. But for $t > 0$, the support of the law of $X(\tau \wedge t)$ is all of $\bar{G}$. (This can be proved either from the Stroock-Varadhan support theorem or by p.d.e. methods. We omit the details.) Hence

$(LV)(x,u_0) + k(x,u_0) \geq 0$ a.e. on G.

(5.2) and the 'only if' part of the last claim follow. The 'if' part follows from the stochastic representation mentioned above and the definition of V. Q.E.D.

We shall briefly indicate the corresponding results for C3 with $T<\infty$ and $\alpha=0$. For $t\in[0,T]$, $x\in R^d$, define

$$V(t,x) = \inf_{u(\cdot)} E[\int_t^T k(X(s),u(s))ds + h(X(T))]$$

Then

(i) V is continuous

(ii) V satisfies

$$V(t,x) = \min_{u(\cdot)} E[\int_t^{T\wedge(\sigma\vee t)} k(X(s),u(s))ds + V(X(T\wedge\sigma))] \qquad (5.3)$$

for any $\{F_t\}$-stopping time σ and $u(\cdot)$ attains the minimum in (5.3) if and only if the restriction of $u(\cdot)$ to $[t, T\wedge(\sigma\vee t)]$ coincides with the restriction to the same interval of a control which is optimal for $X(s)$, $s\geq t$, satisfying (2.1) with $X(t) = x$ and with cost functional $E[\int_t^T k(X(s), u(s))ds + h(X(T))]$.

(iii) In the nondegenerate case, V is the unique solution in $W^{1,2}_{p,loc}([o,T]xR^d)\cap C([0,T]xR^d)$, $p\geq 2$ satisfying a suitable growth condition at infinity to the Hamilton-Jacobi-Bellman equation

$$\min_u\left[\frac{\partial V}{\partial t}(t,x) + (LV(t,\cdot))(x,u) + k(x,u)\right] = 0 \text{ on } [0,T]xR^d,\ V(T,\cdot) = h(\cdot) \quad (5.4)$$

Furthermore, a Markov control v is optimal for some initial condition if and only if it is optimal for any initial condition, which is if and only if it attains the minimum in (5.4) for a.e. (t,x) in $[0,T]xR^d$.

In (5.2) and (5.4), the minimum will be clearly obtained by a Dirac measure on S for each fixed x. An appeal to the selection theorem then assures us of an optimal precise control that is stationary Markov (resp. Markov). In the degenerate case, however, an optimal precise control need not exist, as the following example shows: Let d=1, S={-1,1}, $\bar{m}(x,y) = y$, $\sigma \equiv 0$, $X_0=0$ and cost = $E[\int_0^T X^2(t)dt]$ for some T>0. Then the relaxed control $u(\cdot) \equiv (\delta_1+\delta_{-1})/2$ (δ_x being the Dirac measure at x) gives zero cost whereas no precise control does.

Thus it is clearly hopeless in the degenerate case to expect that V would satisfy an H.J.B. equation like (5.2) or (5.4) in the above sense. This has led to the development of a new solution concept for H.J.B. equations called the viscosity solutions which coincide with the conventional ones for the nondegenerate case. V then can be shown to be the unique viscosity solution of the appropriate H.J.B. equation even in the degenerate case. Alternatively, in case of C2, it is characterized as the maximal subsolution of the system

$$-[(LV)(x,u) + k(x,u)] \leq 0 \text{ on } G,\ V \leq h \text{ on } \delta G,\ u \in U \qquad (5.5)$$

with the corresponding analog of (5.4) for C3. See [51], [52] for details.

There is an alternative interesting way of arriving at (5.5) in the nondegenerate case. Theorem 3.1 allows us to consider the control problem for, say, C2 as an optimization problem over a compact convex set of measures. The dual of this problem in the conventional convex analytic sense turns out to be precisely the problem of finding a maximal subsolution of (5.5). See [39] for details.

There are two important abstractions of the dynamic programming principle. We illustrate these for C3 with $\alpha=0$, $T<\infty$. The first is the martingale formulation [25], [57], [59] which states that $V(X(t)) + \int_0^t k(X(s),u(s))ds$, $t\geq 0$, ($u(\cdot)$ feedback) is a submartingale w.r.t. the natural filtration of $X(\cdot)$ and is a martingale if and only if $u(\cdot)$ is optimal. The second is the nonlinear semigroup formulation [54], [55], [56] which states that the map $A_T:C_b(R^d) \to C_b(R^d)$ mapping h into V is a one parameter semigroup in T whose infinitesimal general $\tilde{A}$ is given by $\tilde{A}f = \inf_u((Lf)(\cdot,u) + k(\cdot,u))$ for smooth compactly supported $f\in C(R^d)$. Either formulation is obvious in view of the foregoing. However, their real power lies in the ease with which they generalize to more general semimartingale or Markov process models.

Relevant references: [6], [24], [38], [47], [51], [52], [54], [55], [56], [57], [59].

VI. THE DEGENERATE CASE

In this section, we drop the nondegeneracy assumption and adapt the idea of Markov selections due to Krylov ([61], Ch.12) to establish the existence of an optimal Markov control along the lines of [29], [43]. We consider the simplest case which is C3 with $T=\infty$, $\alpha>0$. See [29] for other cases of C3 and [43] for C2. (For the special case of C3 considered here and C2, one strengthens Markov to stationary Markov.)

Let $\{f_i,\ i\geq 1\} \subset C_b(R^d)$ be dense in $C(R^d)$ and $\{\beta_j,\ j\geq 1\}$ be dense in $(0,\infty)$. Define $F_{ij}:C([0,\infty);\ R^d) \to R$ by

$$F_{ij}(w(\cdot)) = \int_0^\infty e^{-\beta_i t} f_j(w(t))dt,\ i,j\geq 1. \tag{6.1}$$

Enumerate the F_{ij}'s as $F_1, F_2,\ldots$ by a suitable relabelling. Let $\mathbf{A}$ = {all admissible $u(\cdot)$}. (Recall that these need not be defined on the same probability space.) Define $V_i:R^d \to R$ and $\mathbf{A}_i(X_0)$, $i\geq 0$, inductively as follows:

$$V_o=V,\ \mathbf{A}_o(X_0) = \{u(\cdot)\varepsilon\mathbf{A} | u(\cdot) \text{ optimal}\}$$

For $i\geq 1$,

$$\bar{V}_i(X_0) = \min_{u(\cdot)\varepsilon \mathbf{A}_{i-1}(X_0)} E[F_i(X(\cdot))] \tag{6.2}$$

$$\mathbf{A}_i(X_0) = \{u(\cdot)\ \varepsilon\ \mathbf{A}_{i-1}(X_0) \text{ for which the minimum in (6.2) is attained}\}$$

$$V_i(x) = \bar{V}(X_0) \text{ for } X_0 = x.$$

The above is self-explanatory once it is observed that by the same arguments as the ones leading to Theorem 4.2, the set $M_i(X_0)$ of laws of $X(\cdot)$ corresponding to $u(\cdot)\ \varepsilon\ \mathbf{A}_i(X_0)$ is compact nonempty in $\mathbf{P}(C([0,\infty);R^d))$ for each i and thus the minimum in (6.2) is attained. Note that $\mathbf{A}_i(X_0) \supset \mathbf{A}_{i+1}(X_0)$, implying $M_i(X_0) \supset M_{i+1}(X_0)$, $i \geq 0$. Thus $M_\infty(X_0) = \bigcap_{i \geq 0} M_i(X_0)$ is nonempty and therefore $\mathbf{A}_\infty(X_0) = \bigcap_{i \geq 0} \mathbf{A}_i(X_0)$ is nonempty.

A simple conditioning argument shows that $\bar{V}_i(X_0) = E[V_i(X_0)]$. Also, a straightforward adaptation of the argument leading to Lemma 5.2 shows that $V_i\ \varepsilon\ C_b(R^d)$ for all i.

Let $i \geq 0$, $u(\cdot)\ \varepsilon\ \mathbf{A}_i(X_0)$ and $X(\cdot)$ the corresponding solution of (2.1). Let τ be an a.s. finite stopping time w.r.t. the natural filtration $\{F_t\}$ of $X(\cdot)$. Let $\mathbf{L}_\tau$ denote the regular conditional law of $X(\tau+\cdot)$ given F_τ.

<u>Lemma 6.1</u>. $\mathbf{L}_\tau\ \varepsilon\ \bar{M}_i(X(\tau))$ a.s. where $\bar{M}_i(X_0) = M_i(y)|_{y=X_0}$.

<u>Proof</u>. (Sketch) Let i=0. By Lemma 5.1, $\mathbf{L}_\tau$ is a.s. the law of a controlled diffusion of the type (2.1) with initial condition $X(\tau)$. If the claim were false on some $A\varepsilon F_\tau$ with $P(A)>0$, we could modify $u(\cdot)$ on A from τ onwards to obtain a lower cost, a contradiction. Hence the claim holds for i=0.

Suppose it holds for some $i \geq 0$ and take $u(\cdot) \in \mathbf{A}_{i+1}(X_0)$. Then $u(\cdot) \in \mathbf{A}_i(X_0)$ and by the induction hypothesis, $\mathbf{L}_\tau \in \bar{M}_i(X(\tau))$ a.s. Repeat the above argument to obtain a contradiction unless $\mathbf{L}_\tau \in \mathbf{A}_{i+1}(X(\tau))$ a.s. The claim follows by induction. Q.E.D.

Corollary 6.1. If $u(\cdot)$ above is in $\mathbf{A}_\infty(X_0)$, $\mathbf{L}_\tau \in \bar{M}_\infty(X(\tau))$ a.s.

Lemma 6.2. For $f \in L_\infty(R^+)$,

$$\int_0^\infty e^{-\beta_j t} f(t)dt = 0 \text{ for all } j \Rightarrow f(t) = 0 \quad \text{a.e.}$$

Proof. By continuity

$$\int_0^\infty e^{-\beta t} f(t)dt = 0 \quad \text{for } \beta \in (0,1).$$

By successive differentiation w.r.t. β at $\beta=1$,

$$\int_0^\infty t^n e^{-t} f(t)dt = 0 \quad \text{for} \quad n \geq 0$$

The set of measurable $g: R^+ \to R$ for which

$$\int_0^\infty e^{-t} g(t)f(t)dt = 0$$

is a vector space which is closed under uniform convergence and bounded monotone convergence and contains the algebra of polynomials. By the monotone class theorem of [26], pp. 14, it includes $L_\infty(R^+)$ and in particular, f itself. The claim follows. Q.E.D.

Theorem 6.1. For $u(\cdot) \in A_\infty(X_0)$, $X(\cdot)$ is a homogeneous Markov, strong Markov process.

Proof. Fix $i \geq 1$. Then for $w(\cdot) \in C([0,\infty); R^d)$,

$$F_i(w(\cdot)) = \int_0^\infty e^{-\beta_j t} f_\ell(w(t))dt$$

for suitable j,ℓ. By Lemma 6.1,

$$E[F_i(X(\tau+\cdot))/F_\tau] = V_i(X(\tau)) \text{ a.s.}$$

$$= E[F_i(X(\tau+\cdot))/X(\tau)] \quad \text{a.s.} \tag{6.3}$$

By Lemma 6.2,

$$E[f_\ell(X(\tau+t))/F_\tau] = E[f_\ell(X(\tau+t))/X(\tau)] \quad \text{a.e. } t, \text{ a.s.}$$

Using Fubini's theorem to interchange 'a.e.t.' and 'a.s.' and then taking an appropriate version, we conclude that for each $t \geq 0$,

$$E[f\ (X(\tau+t))/F_{\tau}] = E[f\ (X(\tau+t))/X(\tau)] \quad \text{a.s.}$$

The claim follows, the homogeneity being a consequence of the fact that the middle term in (6.3) does not have an explicit τ-dependence. Q.E.D.

Corollary 6.2. If processes $X(\cdot)$, $X'(\cdot)$ have their laws in $M_\infty(x)$ for some $x \varepsilon R^d$, the laws of $X(t)$, $X'(t)$ agree for each t.

Proof. From (6.3) with $\tau=0$,

$$E[F_i(X(\cdot))] = V_i(x) = E[F_i(X'(\cdot))], \ i \geq 1.$$

Use Lemma 6.2 and the density of $\{f_i\}$ in $C(R^d)$ to conclude. Q.E.D.

This allows us to define a collection of transition probabilities $\Gamma = \{p(x,t,\cdot),\ x\varepsilon R^d,\ t \geq 0\}$ $P(R^d)$ by

$$p(x,t,A) = P(X(t)\varepsilon A), \quad A \text{ Borel in } R^d,$$

for any $X(\cdot)$ whose law is in $M_\infty(x)$. The Chapman-Kolmogorov equation for Γ follows from Lemma 6.1. The same fact also shows that Γ is the family of transition probabilities for the Markov process featuring in Theorem 6.1. Since transition probabilities and initial law completely specify the law of a Markov process, this implies that each $M_\infty(X_0)$, in particular $M_\infty(x)$, $x\varepsilon R^d$, is a singleton, strengthening the conclusion of Corollary 6.2.

Using the selection theorem, it is proved in [43], pp. 184-5, that there exists a stationary Markov control v such that for any X_0, the unique element of $M_\infty(X_0)$ is the law of a process satisfying (2.1) with the control v. This assures that there is one optimal solution of (2.1) under the stationary Markov control v. This does not preclude the possibility that there may be others which are not optimal. Nor is it assured that there will be even one solution for any other stationary Markov control.

Relevant references: [29], [43], [61].

VII. CONTROL UNDER PARTIAL OBSERVATIONS

In the problem of control under partial observations, one has in addition to (2.1) an R^m-valued 'observation process' $Y(\cdot)$ satisfying

$$Y(t) = \int_0^t q(X(s))ds + W'(t) \tag{7.1}$$

where $q:R^d \to R^m$ is bounded continuous and $W'(\cdot)$ is an R^m-valued standard Wiener process independent of X_0, $W(\cdot)$. The objective then is to minimize a cost functional over all $u(\cdot)$ adopted to the natural filtration of $Y(\cdot)$. These are called strict sense admissible controls.

The existence of an optimal control in this case is a long standing open problem (except in some simple cases like the well-known 'Linear-Quadratic-Gaussian' case [38]). Therefore one enlarges the class of controls in the following manner: Let (Ω,F,P) denote the underlying probability space and $\{F_t\}$ the natural filtration of $(X(\cdot), Y(\cdot), u(\cdot))$. Define a new probability measure P_0 on (Ω,F) by:

$$\left.\frac{dP}{dP_0}\right|_{F_t} = \Lambda(t) = \exp\left(\int_0^t \langle q(X(s)),dY(s)\rangle - \frac{1}{2} \int_0^t ||q(X(s)||^2 ds\right) \tag{7.2}$$

This $\Lambda(\cdot)$ is seen to be the unique solution to the s.d.e.

$$\Lambda(t) = 1 + \int_0^t \Lambda(s) \langle q(X(s)), dY(s)\rangle \tag{7.3}$$

Under P_0, $Y(\cdot)$ is an m-dimensional Wiener process. Call $u(\cdot)$ a wide sense admissible control if for each t, u(t) is independent of W(•) and of $Y(t_2) - Y(t_1)$ for $t_2 \geq t_1 \geq t$ under P_0. This clearly contains the class of strict sense admissible controls. We seek an optimal control in this larger class.

For simplicity, consider the cost C2 with either nondegeneracy or the condition in the remarks following Theorem 4.1. Letting $E_0[\ ]$ denote the expectation under P_0, one can check that the cost can be rewritten as

$$E_0[\int_0^{\tau} k(X(t),u(t))\Lambda(t)dt + h(X(\tau))\Lambda(\tau)], \tag{7.4}$$

which has the same form as before but with the new (d+1+d+m)-dimensional controlled process (X(•), Λ(•), W(•), Y(•)) whose dynamics is given by (2.1), (7.1) and the trivial s.d.e.s W(•) = W(•), Y(•) = Y(•). One can now repeat the arguments leading to Theorem 4.1 to conclude that an optimal wide sense admissible control exits, the only extra bit needed being the observation that the independence of u(s), $s \leq t$, (identified with the D-valued $\alpha(s)$, $s \leq t$, as in Section III) and W(•) or Y(t+•)-Y(t) for each t is preserved under weak convergence [30].

A similar argument works for C3. For C1, one needs some additional restriction such as that the map F there should depend only on the restriction to [0,T] of its argument for some $T \in (0,\infty)$. This is so because although P, P_0 are mutually absolutely continuous on $\{F_t\}$ for each t, they need not be so on $\bigvee_t F_t$.

One would like to go a step further and have an optimal control that depends in a feedback, or even better, Markovian fashion on an appropriate 'state' process. (Clearly, X(•) no longer qualifies as the latter.) We shall indicate how to achieve this for C3 and a special case of C2, both under the nondegeneracy condition.

Consider C3 with $T<\infty$, $\alpha=0$ (the general case can be handled similarly). The natural candidate for the state process here is the conditional law $\pi(t)$ of $X(t)$ given $Y(s)$, $u(s)$, $s\leq t$, for $t\geq 0$. Introduce the notation $\mu(f) = \int f d\mu$ for $f \in C_c^\infty(R^d)$ = smooth real-valued functions on R^d with compact supports and $\mu \in P(R^d)$. Then the evolution of $\pi(\cdot)$ is described by the well-known Fujisaki-Kallianpur-Kunita equation [53]

$$\pi(t)(f) = \pi_0(f) + \int_0^t \pi(s)((Lf)(\cdot,\ u(s)))ds + \int_0^t \langle \pi(s)(fq)-\pi(s)(f)\pi(s)(q), d\hat{Y}(s)\rangle \quad (7.5)$$

for $f \in C_c^\infty(R^d)$, where fq is the componentwise product of q by f, $\pi(s)(fq)$ and $\pi(s)(q)$ are defined componentwise in an obvious manner, and $\hat{Y}(t) = Y(t) - \int_0^t \pi(s)(q)ds$ is the so-called innovations process which is a Wiener process under P having the same natural filtration as $Y(\cdot)$ [1]. Under P_0, (7.5) becomes

$$\pi(t)(f) = \pi_0(f) + \int_0^t [\pi(s)((Lf)(\cdot,u(s))) + \langle \pi(s)(fq) - \pi(s)(f)\pi(s)(q), \pi(s)(q)\rangle]ds + \int_0^t \langle \pi(s)(fq) - \pi(s)(f)p(s)(q),\ dY(s)\rangle \quad (7.6)$$

for $f \in C_c^\infty(R^d)$ with $Y(\cdot)$ a Wiener process.

We shall assume that given on some probability space (Ω,F,P_0) a Wiener process $Y(\cdot)$ and a U-valued process $u(\cdot)$ satisfying: $u(t)$ is independent of $Y(t+\cdot)-Y(t)$ for each t, the solution to (7.6) is pathwise unique. One situation where such a result is known is the case when q is twice continuously differentiable with suitable growth conditions on its first and second partial derivatives as in [15], [42]. Then $\pi(\cdot)$ in (7.6) is interconvertible to another measure-valued process $\nu(\cdot)$ satisfying a nonstochastic p.d.e. (but with stochastic processes $u(\cdot)$, $Y(\cdot)$ featuring as 'parameters') called the pathwise filtering equation the uniqueness problem for which can be handled by standard methods [15], [42]. We omit the details.

An important consequence of this uniqueness is the fact that given (7.6) on some (Ω,F,P_0), $\pi(\cdot)$ is the process of conditional laws for a partially observed control system of the type described earlier on a possibly augmented probability space, after an appropriate absolutely continuous change of measure. This is achieved as follows: By adjoining a copy of $R^d \times C([0,\infty);R^d)$ to Ω if necessary (cf. the proof of Theorem 2.2 (b)), we can construct an R^d-valued random variable X_0 with law π_0 and a d-dimensional standard Wiener process $W(\cdot)$ which are independent of each other and of $Y(\cdot)$. Construct $X(\cdot)$ by (2.1). Change measure to P by (7.2). Then the conditional law $\pi'(t)$ of $X(t)$ given $Y(s)$, $u(s)$, $s\leq t$, $t\geq 0$, under P satisfies (7.5) under P and hence (7.6) under P_0, thereby coinciding with $\pi(t)$, $t\geq 0$, by the uniqueness hypothesis.

Note that the cost functional can be rewritten as

$$E\left[\int_0^T \pi(t)(k(\cdot,u(t)))dt + \pi(T)(h)\right] \tag{7.7}$$

Thus we can consider $\pi(\cdot)$ itself as a controlled process with (7.7) as cost. This is called the separated control problem. We know that a wide sense admissible optimal $u(\cdot)$ exists for this and by considerations analogous to Theorem 2.2 (a), one may assume that it is in a feedback form, i.e., is adapted to the natural filtration of $\pi(\cdot)$ [15], [30]. In analogy with Section V, we can define a value function $V:[0,T]\times P(R^d) \rightarrow R$ by

$$V(t,\nu) = \min E\left[\int_t^T \pi(s)(k(\cdot,u(s)))ds + \pi(T)(h)/\pi(t) = \nu\right]$$

where the minimum is over all wide sense admissible controls. We can mimic the argument of Lemma 5.2 to claim that V is continuous and that of Theorem 5.1 to claim that it satisfies the dynamic programming principle.

$$V(t,\nu) = \min E\left[\int_t^{T\wedge(\sigma\vee t)} \pi(s)(k(\cdot,u(s)))ds + V(t,\pi(T\wedge\sigma))\right] \tag{7.8}$$

for any σ which is a stopping time w.r.t. the natural filtration of $\pi(\cdot)$, the minimum in (7.8) being once again over all wide sense admissible controls. Furthermore, it follows as before that this minimum is attained for a wide sense admissible $u(\cdot)$ if and only if its restriction to $[t,T\wedge(\sigma\vee t)]$ coincides with that of an optimal wide sense admissible control.

This suggests that one can mimic the arguments of the preceding section to obtain an optimal process $\pi(\cdot)$ which is a Markov process satisfying (7.5) for a $u(\cdot)$ of the form $u(t) = v(\pi(t),t)$, $t\geq 0$, where v is some measurable map $P(R^d)x[0,T] \to U$. This is indeed so. See [30] for details.

We shall now briefly sketch a method of obtaining a separated control problem for C2 when h is the restriction of an C^2 function with bounded first and second partial derivatives. The latter restriction allows us to assume w.l.o.g. that $h\equiv 0$ since we can always replace k by k+Lh to achieve this. Define a $P(\bar{G})$-valued process $\eta(\cdot)$ by

$$\int f d\eta = E[I\{\tau>t\}f(X(t))/Y(s),u(s),s\leq t]/P(\tau>t/Y(s),u(s),s\leq t)$$

for $f\varepsilon C(\bar{G})$, $t\geq 0$, taking a measurable version thereof. The evolution of $\eta(\cdot)$ is described by

$$\eta(t)(f) = \eta(0)(f) + \int_0^t (\eta(s)((Lf)(\cdot,u(s))) + \eta(s)(fg) - \eta(s)(f)\eta(s)(g))ds +$$

$$\int_0^t \langle \eta(s)(fq)-\eta(s)(f)\eta(s)(q),\ d\hat{Y}(s)\rangle \qquad (7.9)$$

where $g:\bar{G} \to R$ is defined as follows: Let $Z(\cdot)$ be the unique solution to

$$dZ(t) = \sigma(Z(t))dW(t),\ Z(0) = x\varepsilon\bar{G},$$

with $W(\cdot)$ a Wiener process. Let $\tau_0 = \inf\{t\geq 0 | Z(t)\varepsilon\bar{G}\}$. Define

$$g(x) = \frac{\partial}{\partial t} P(\tau_0 > t) \Big|_{t=0} .$$

Then g can be shown to be continuous [19]. We do not drive (7.9) here. The evolution of a related 'unnormalized' process taking values in the space of finite nonnegative measures on $\bar{G}$ is derived in [19]. $\eta(\cdot)$ can be obtained from this process by normalizing it to a probability measure. (7.9) is then easily derivable from the evolution equation of the above process by a simple application of the Ito formula. The cost can be rewritten as

$$\int_0^\infty E[\int k(\cdot,u(t))d\eta(t)]dt$$

where the integrand can be shown to be dominated in absolute value by an exponentially decaying function of time uniformly in $u(\cdot)$ [19]. With this as the cost for the state process $\eta(\cdot)$, we can now mimic the foregoing developments for C3, $\pi(\cdot)$.

The use of wide sense admissible controls can be partly justified in case of C2, C3 by the fact that the infima of the costs over strict sense admissible and wide sense admissible controls coincide. A proof of this follows along the following lines: Look at the separated control problem, say for C3. Say that the sequence of controls $\{u^n(\cdot)\}$ approximates a control $u(\cdot)$ if the corresponding D-valued processes $[\int f_1 du^n(\cdot), \int f_2 du^n(\cdot), \ldots]$ $\{f_i\}$ as in Lemma 3.1, converge a.s. in D to $[\int f_1 du(\cdot), \int f_2 du(\cdot) \ldots]$. By familiar arguments, the corresponding costs converge. Now given wide sense admissible control $u(\cdot)$, it is approximated by w.s.

admissible controls $\{u^n(\cdot)\}$ with continuous paths defined by

$$\int f du^n(t) = n \int_{(t-1/n)Vo} \int_S f du(s)ds, \quad n=1,2,\ldots, f \in C(S).$$

In turn, each $u^n(\cdot)$ can be approximated by a piecewise constant w.s. admissible control in an obvious manner. Thus each wide sense admissible control $u(\cdot)$ is approximated by piecewise constant wise sense admissible controls. Consider a specific partition of the time axis into finite intervals and consider optimization of the prescribed cost over only those w.s. admissible controls which are constant on each of these intervals. In this subclass, an optimal strict sense admissible control can be shown to exist by treating the above as a discrete time control problem. The claim follows. See [30] for details.

As for a dynamic programming principle for the control under partial observations, a martingale formulation of the same is given in [25]. A more common approach is the nonlinear semigroup formulation for the separated control problem [8], [17], [35], [36]. Attempts have been made to extend the results available for the H.J.B. equation in the completely observed case to an appropriate generalization of the same to the partially observed case [5], [45], [46], but with limited success.

Relevant references: [5], [8], [9], [14], [15], [17], [19], [25], [30], [35], [36], [37], [41]. [42], [45], [46].

VIII. THE ERGODIC CONTROL PROBLEM

In the ergodic control problem, one seeks to a.s. minimize the cost

$$\limsup_{t \to \infty} \frac{1}{t} \int_0^t k(X(s),u(s))ds, \tag{8.1}$$

where k is as before. Assume nondegeneracy. (8.1) differs strikingly from C1-C3 in that the finite time behaviour of $X(\cdot)$ is now immaterial and only 'average' asymptotic behaviour matters. Call a stationary Markov control v a stable stationary Markov control (SSMC for short) if under v, $X(\cdot)$ is positive recurrent and thus has a unique invariant probability measure η_v [10]. Call it an unstable stationary Markov control (USMC) otherwise. Under an SSMC v, (8.1) a.s. equals $\int k(x,v(x))\eta_v(dx)$. One typically tries to find an optimal SSMC. Thus we assume that at least one SSMC exists. However, an optimal SSMC need not always exist as the following example shows: If $k(x,u) = \exp(-||x||^2)$ and a USMC exists, the USMC will clearly give a strictly lower (i.e., zero) cost than any SSMC. We consider two cases where an optimal SSMC can be shown to exist.

Call k near-monotone if

$$\lim_{||x|| \to \infty} \min_u k(x,u) > \beta = \inf_{\text{all SSMC } v} \int k(x,v(x))\eta_v(dx) \tag{8.2}$$

This nomenclature is motivated by the fact that any k which is of the form $k(x) = f(||x||)$, f monotone increasing on R^+, will satisfy (8.2). This is an important class of costs in practice. Clearly, one expects such k to penalize unstable behaviour and thus an optimal SSMC to exist. This intuition is confirmed by the results to follow.

The second case that we shall consider is a stability condition that ensures that all stationary Markov controls are stable (and more).

The key result in both cases is Theorem 8.1 below. Let $\bar{R}^d = R^d \cup \{\infty\}$ be the one point compactification of R^d. Define a $P(\bar{R}^d xS)$-valued process $\mu(\cdot)$ by

$$\mu(t)(AxB) = \frac{1}{t}\int_0^t I\{X(s)\varepsilon A\}u(s)(B)ds, \ t\geq 0; \ A\subset \bar{R}^d, \ B\subset S \text{ measurable}$$

For any $\mu\varepsilon P(\bar{R}^d xS)$, we can write the decomposition $\mu(A) = \delta_\mu \mu_1(A\cap(R^d xS)) + (1-\delta_\mu)\mu_2(A\cap(\{\infty\}xS))$, $A\subset \bar{R}^d xS$ measurable, where $\delta_\mu\varepsilon[0,1]$, $\mu_1\varepsilon P(R^d xS)$, $\mu_2\varepsilon P(\{\infty\}xS)$. The decomposition is unique for $\delta_\mu\varepsilon(0,1)$ and is rendered unique for $\delta_\mu=0$(resp.1) by fixing an arbitrary choice of μ_1(μ_2 resp.). Disintegrate μ_1 as

$$\mu_1(dx,ds) = \mu^*(dx)v_\mu(x)(ds)$$

where $\mu^*\varepsilon P(R^d)$ is the image of μ_1 under the projection $R^d xS \to R^d$ and $v_\mu : R^d \to U$ the regular conditional law.

<u>Theorem 8.1</u>. For all sample points outside a set N with P(N) = 0, each limit point μ of $\mu(t)$, $t\geq 0$, in $P(\bar{R}^d xS)$ for which $\delta_\mu>0$ satisfies

$$\mu^* = \eta_{v_\mu} \tag{8.3}$$

(In particular, v_μ is an SSSM.)

<u>Proof</u>. Let $f_i: R^d \to R$, $i \geq 1$, be such that (i) $f_i \in C_c(R^d)$, (ii) their continuous extensions to $\bar{R}^d$ (denoted $\{f_i\}$ again by abuse of notation) are dense in $\{g \in C(\bar{R}^d) \mid g(\infty) = 0\}$. Divide the equations

$$f_i(X(t)) - f_i(X(0)) = \int_0^t (Lf_i)(X(s), u(s))ds + \int_0^t \langle \nabla f_i(X(s)), \sigma(X(s))dW(s)\rangle, \ i \geq 1 \tag{8.4}$$

by t throughout and let $t \to \infty$. The last term in (8.4) is a time-changed Brownian motion whose process of time change has a uniformly bounded derivative and hence it is o(t) a.s. [21]. It follows that outside a set N with P(N) = 0, any limit point μ of $\{\mu(t)\}$ for which $\delta_\mu > 0$ must satisfy

$$\int (Lf_i)(x, v_\mu(x))\mu^*(dx) = 0, \ i \geq 1.$$

The claim follows by Theorem 9.19, pp. 252-3, [33]. Q.E.D.

<u>Theorem 8.2</u>. If k is near-monotone, an optimal SSSM exists.

<u>Proof</u> (Sketch) Let $\{v_n\}$ be a sequence of stationary Markov controls such that

$$\int k(x,v_n(x))\eta_{v_n}(dx) \downarrow \beta. \tag{8.5}$$

Define $\nu_n \in P(R^d xS)$ by

$$\int_{R^d xS} fd\nu_n = \int_{R^d}\int_S f(x,s)v_n(x)(ds)\eta_{v_n}(dx), \quad f \in C_b(R^d xS),$$

for $n \geq 1$. (8.2), (8.5) force $\{\nu_n\}$ to be tight [21]. Any limit point μ of $\{\nu_n\}$ can be decomposed as

$$\mu(dx,ds) = \bar{\mu}(dx)v(x)(ds)$$

with $\bar{\mu} \in P(R^d)$, $v:R^d \to U$. Since for each $n \geq 1$,

$$\int (Lf_i)(x,v_n(x))\eta_{v_n}(dx) = 0, \ i \geq 1,$$

letting $n \to \infty$, we get

$$\int (Lf_i)(x,v(x))\bar{\mu}(dx) = 0, \ i \geq 1,$$

implying as above that $\bar{\mu}=\eta_v$. Clearly $\int k(x,v(x))\eta_v(dx) = \beta$. For arbitrary $u(\cdot)$, (8.2) and (8.3) ensure that

$$\liminf_{t \to \infty} \frac{1}{t} \int_0^t k(X(s),u(s))ds \geq \beta \quad \text{a.s.}$$

Thus v above is an optimal SSSM. Q.E.D.

The second case we consider is the following : Let B_1, B_2 with $B_1 \subset B_2$ be concentric spheres with boundaries δB_1, δB_2 resp. Define the stopping times

$$\tau_1 = \inf\{t \geq 0 | X(t) \in \delta B_1\}$$

$$\tau_2 = \inf\{t \geq \tau_1 | X(t) \in \delta B_2\}$$

$$\tau_3 = \inf\{t \geq \tau_3 | X(t) \in \delta B_1\}$$

Assume that

$$\sup E[(\tau_3 - \tau_1)^2] < \infty \tag{8.6}$$

where the supremum is over all initial data and all admissible controls. (This condition is not in an easily verifiable form. A simpler Liapunov-type sufficient condition for the above to hold is given in [21]. This condition requires that there exist a twice continuously differentiable function $w: R^d \to R$ satisfying

(i) $\lim_{||x|| \to \infty} w(x) = +\infty$,

(ii) there exists $\varepsilon > 0$ such that for all x with $||x||$ sufficiently large
$$\sup_{u} (Lw)(x,u) \leq -\varepsilon,$$
plus some additional technical hypotheses.)

Define for each SSSM v, a measure $\nu_v \in P(R^d xS)$ by

$$\int f d\nu_v = \iint f(x,s)v(x)(ds)\eta_v(dx), \quad f \in C_b(R^d xS)$$

Under (8.6), it is proved in [21] that

(i) all stationary Markov controls are stable and the set $\{\nu_v | v \text{ SSSM}\}$ is compact in $P(R^d xS)$

(ii) for a suitable choice of N in Theorem 8.1, δ_ν is always one for ν as in the statement of Theorem 8.1.

Using arguments analogous to those used for the first case, one can show from (i), (ii) above that an optimal SSSM exists.

When d=1, both (i) and (ii) can be derived from the comparison theorem for one dimensional diffusions [44] under the weaker assumption that all stationary Markov controls are stable [21]. The proof is almost obvious.

<u>Relevant references</u>: [10], [16], [21], [22], [49], [58], [62].

IX. SOME OPEN PROBLEMS

We conclude in this section by listing a few open problems related to the foregoing.

(1) Let $M(\cdot)$ be an integrable real-valued process adapted to the natural filtration of $X(\cdot)$ such that the value of $M(t+s)$, $t,s\geq 0$, is completely specified by the value of $M(t)$ and the restriction of $X(\cdot)$ to $[t,t+s]$. Show that for the cost functional $E[f(M(T))]$ for some $T>0$, $f\varepsilon C(R)$, an optimal control $u(\cdot)$ of the form $u(t) = v(X(t), M(t), t)$, $t\geq 0$, exists, where $v:R^{d+1}xR^+ \to U$ is a measurable map. Examples are:

(i) $M(t) = \max_{0\leq s\leq t} ||X(s)||$, $f(x) = x$,

(ii) $M(t) = \int_0^t g(X(s))ds$ for some $g\varepsilon C_b(R)$, $f\varepsilon C_b(R)$.

(2) We proved that for a bounded domain G with a C^2 boundary and under nondegeneracy assumptions, the Green measure and the hitting distribution on the boundary for an admissible control coincide with those under some stationary Markov control (Lemma 4.1). Prove an analog of this for the degenerate case.

(3) Show the existence of an optimal wide sense admissible control for the control problems under partial observations considered in Section VII when q explicitly depends on the control.

(4) Formulate a separated control problem for C2 when h is only continuous.

(5) For the ergodic control problem with $d \geq 2$ show that an optimal stable stationary Markov control exists whenever all stationary Markov controls are stable, thus dispensing with the additional assumption (8.6).

(6) Show the existence of an optimal wide sense admissible control for the ergodic control problem with partial observations.

(7) The ergodic control problem without the nondegeneracy hypothesis needs to be studied.

APPENDIX

The following selection theorem is used frequently in the main text of this paper. See [3], pp. 182-4, for a proof.

Theorem. Let $(M,\mathbf{M})$ be a measure space, A separable metric and U the union of countably many compact metrizable subsets of itself. Let $k:M\times U \to A$ be continuous in its second argument for each value of the first and $\mathbf{M}$-measurable in the first for each value of the second. Let $y:M \to A$ be $\mathbf{M}$-measurable with

$$y(x) \in k(x,U), \quad x \in M.$$

Then there exists an $\mathbf{M}$-measurable $u:M \to U$ such that $y(x) = k(x,u(x))$.

ACKNOWLEDGEMENT

This work was done while the author was visiting the Laboratory for Information and Decision Systems, M.I.T., and was supported by A.R.O. contract DAAG-84-K-0005 and AFOSR 85-0227. The author would like to thank

Prof. S.K. Mitter for his hospitality while at L.I.D.S. He would also like to thank Prof. Fleming of Brown University for introducing him to some of the unpublished material surveyed here while still in preprint form.

REFERENCES

[1] D. Allinger, S.K. Mitter, New results for the innovations problem for nonlinear filtering, Stochastics 4 (1981), 339-348.

[2] A.V. Balakrishnan, Applied functional analysis, 2nd Ed., Springer, 1981.

[3] V.E. Beneš, Existence of optimal strategies based on specified information, for a class of stochastic decision problems, SIAM J. Control and Opt. 8 (1970), 179-188.

[4] V.E. Beneš, Existence of optimal stochastic controls, SIAM J. Control and Opt. 9 (1971), 446-472.

[5] V.E. Beneš,, I. Karatzas, On the relation of Zakai's and Mortensen's equations, SIAM J. Control and Opt. 21 (1983), 472-489.

[6] A. Bensoussan, Stochastic control by functional analysis methods, North Holland, 1982.

[7] A. Bensoussan, Lectures on stochastic control, in 'Nonlinear filtering and stochastic control', S.K. Mitter, A. Moro (Eds.) Lecture Notes in Mathematics No. 972, Springer, 1983, 1-62.

[8] A. Bensoussan, Maximum principle and dynamic programming approaches of the optimal control of partially observed diffusions, Stochastics 9 (1983), 169-222.

[9] A. Bensoussan, Current results and issues in stochastic control, in 'Stochastic differential systems', N. Christopeit, K. Helmes, M. Kohlmann (Eds.), Lecture Notes in Control and Info. Sciences No. 78, 1986, 36-88.

[10] R.N. Bhattacharya, Asymptotic behaviour of several dimensional diffusions, in 'Stochastic nonlinear systems', L. Arnold, R. Lefever (Eds.), Springer, 1981, 86-99.

[11] P. Billingsley, Convergence of probability measures, Wiley, 1968.

[12] J.-M. Bismut, Conjugate convex functions in optimal stochastic control, J. Math. Anal. and Appl. 44 (1973), 384-404.

[13] J.-M. Bismut, Théorie probabiliste du contrôle des diffusions, Memoirs of A.M.S. 4, No. 167 (1976).

[14] J.-M. Bismut, Partially observed diffusions and their control, SIAM J. control and Opt. 20 (1982), 261-285.

[15] V.S. Borkar, Existence of optimal controls for partially observed diffusions, Stochastics 11 (1983), 103-142.

[16] V.S. Borkar, A note on controlled diffusion on line with time-averaged cost, Systems and Control Letters 4 (1984), 1-4.

[17] V.S. Borkar, The Nisio semigroup for controlled diffusions with partial observations, Stochastics 17 (1986), 145-160.

[18] V.S. Borkar, A remark on the attainable distributions of controlled diffusions, Stochastics 18 (1986), 17-23.

[19] V.S. Borkar, Control of a partially observed diffusion up to an exit time, to appear in Systems and Control Letters.

[20] V.S. Borkar, Controlled diffusions with boundary-crossing costs, preprint.

[21] V.S. Borkar, M.K. Ghosh, Ergodic control of multidimensional diffusions I: the existence results, preprint.

[22] R.M. Cox, I. Karatzas, Stationary control of Brownian motion in several dimensions, Adv. Appl. Prob. 17 (1985), 531-561.

[23] M.H.A. Davis, On the existence of optimal policies in stochastic control, SIAM J. Control and Opt. 11 (1973), 587-594.

[24] M.H.A. Davis, Martingale methods in stochastic control, in 'Stochastic control theory and stochastic differential systems', M. Kohlmann, W. Vogel (Eds.), Lecture Notes in Control and Info. Sciences No. 16, Springer, 1979, 85-117.

[25] M.H.A. Davis, P.P. Varaiya, Dynamic programming conditions for partially observable stochastic systems, SIAM J. Control and Opt. 11 (1973), 226-261.

[26] C. Dellacherie, P.-A. Meyer, Probability and potentials, North Holland, 1978.

[27] T. Duncan, P.P. Varaiya, On the solutions of a stochastic control system, SIAM J. Control and Opt. 9 (1971), 354-371.

[28] N. El Karoui, Les aspects probabilistes du contrôle stochastique, in 'Ecole d'eté de probabilités de Saint-Flour IX-1979' P.L. Hennequin (Ed.), Lecture Notes in Math. No. 876, Springer, 1981, 74-238.

[29] N. El Karoui, D. Huu Nguyen, M. Jeanblanc-Pique, Compactification methods in the control of degenerate diffusions: existence of an optimal control, to appear in Stochastics.

[30] N. El Karoui, D. Huu Nguyen, M. Jeanblanc-Pique, Existence of an optimal Markovian filter for the control under partial observations, preprint.

[31] R.J. Elliot, The optimal control of a stochastic system, SIAM J. Control and Opt. 15 (1977), 756-778.

[32] R.J. Elliot, Stochastic calculus and applications, Springer, 1982.

[33] S. Ethier, T. Kurtz, Markov processes: characterization and convergence, Wiley 1986.

[34] W. Fleming, Generalized solutions in optimal stochastic control, in 'Differential games and control theory III', E. Roxin, P.T. Liu, R.L. Sternberg (Eds.), Marcel Dekker, 1977, 147-165.

[35] W. Fleming, Nonlinear semigroup for controlled partially observed diffusions. SIAM J. Control and Opt. 20 (1982), 286-301.

[36] W. Fleming, M. Nisio, On stochastic relaxed control of partially observed diffusions, Nagoya J. Math. 93 (1984), 71-108.

[37] W. Fleming, E. Pardoux, Optimal control of partially observed diffusions, SIAM J. Control and Opt. 20 (1982), 261-285.

[38] W. Fleming, R. Rishel, Deterministic and stochastic optimal control, Springer, 1975.

[40] U.G. Haussmann, On the stochastic maximum principle, SIAM J. Control and Opt. 16 (1978), 236-251.

[41] U.G. Haussmann, On the existence of optimal controls for partially observed diffusions, SIAM J. Control and Opt. 20 (1982), 385-407.

[42] U.G. Haussmann, L'équation de Zakai et le problème du contrôle optimal stochastique, in 'Seminaire de Probabilitiés XIX 1983/84', J. Azema, M. Yor (Eds.), Lecture Notes in Maths. No. 1123, Springer, 1985, 37-62.

[43] U.G. Haussmann, Existence of optimal Markovian controls for degenerate diffusions, in 'Stochastic differential systems', N. Christopeit, K. Helmes, M. Kohlmann (Eds.), Lecture Notes in Control and Info. Sciences No. 78, Springer, 1986, 171-186.

[44] N. Ikeda, S. Watanabe, Stochastic differential equations and diffusion processes, North Holland, Kodansha, 1981.

[45] M. Kohlmann, Concepts for the derivation of optimal partially observed controls, preprint.

[46] M. Kohlmann, A maximum principle and dynamic programming for a partially observed control problem, preprint.

[47] N.V. Krylov, Controlled diffusion processes, Springer, 1980.

[48] H.J. Kushner, Existence results for optimal stochastic control, J. Opt. Theory and Appl. 15 (1975), 347-359.

[49] H.J. Kushner, Optimality conditions for the average cost per unit time problem with a diffusion model, SIAM J. Control and Opt. 16 (1978), 330-346.

[50] O.A. Ladyženskaja, V.A. Solonnikov, N.N. Ural'ceva, Linear and quasilinear equations of parabolic type, Translations of Mathematical Monographs Vol. 23, A.M.S., 1968.

[51] P.L. Lions, Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equation, part 1: The dynamic programming principle and applications, Comm. in P.D.E. 8 (1983), 1101-1174.

[52] P.L. Lions, Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equation, Part II: Viscosity solutions and uniqueness, Comm. in P.D.E. 8 (1983), 1229-1276.

[53] R.S. Liptser, A.N. Shiryayev, Statistics of random processes. I General theory, Springer, 1977.

[54] M. Nisio, On a nonlinear semigroup attached to stochastic optimal control, Publ. of R.I.M.S., Kyoto U., 1976.

[55] M. Nisio, On stochastic optimal controls and envelope of Markovian semigroups in 'Proc. of the Int. Symp. on Stoch. Diff. Eq.', K. Ito (Ed.), Wiley, 1978.

[56] M. Nisio, Lectures on stochastic control theory, I.S.I. Lecture Notes No. 9, McMillan (India), 1983.

[57] R. Rishel, Necessary and sufficient dynamic programming conditions for continuous time stochastic optimal control, SIAM J. Control and Opt. 8 (1970), 559-571.

[58] M. Robin, Long-term average cost control problems for continuous time Markov processes - a survey, Acta Appl. Math. 1 (1983), 281-300.

[59] C. Striebel, Martingale methods for the optimal control of continuous time stochastic systems, Stoch. Proc. and Appl. 18 (1984), 329-347.