DYNAMIC PROGRAMMING CONDITIONS FOR

PARTIALLY OBSERVABLE STOCHASTIC SYSTEMS

by

M. H. A. Davis and P. P. Varaiya

Memorandum No. ERL-M304

30 August 1971

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, California
94720

DYNAMIC PROGRAMMING CONDITIONS FOR PARTIALLY OBSERVABLE STOCHASTIC SYSTEMS

M. H. A. Davis and P. P. Varaiya

Department of Electrical Engineering and Computer Sciences
and the Electronics Research Laboratory
University of California, Berkeley, California 94720

ABSTRACT

In this paper necessary and sufficient conditions for optimality are derived for systems described by stochastic differential equations with control based on partial observations. The solution of the system is defined in a way which permits a very wide class of admissible controls, and then Hamilton-Jacobi type criteria for optimality are derived from a version of Bellman's "principle of optimality".

The method of solution is based on a result of Girsanov: Wiener measure is transformed for each admissible control to the measure appropriate to a solution of the system equation. The optimality criteria are derived for three kinds of information pattern: partial observations (control based on the past of only certain components of the state), complete observations, and "Markov" (observation of the current state). Markov controls are shown to be minimizing in the class of those based on complete observations for system models of a suitable type.

Research sponsored by the National Science Foundation, Grant GK-10656X and the U.S. Army Research Office--Durham, Contract DAHC04-67-C-0046.

Finally, similar methods are applied to two-person, zero-sum stochastic differential games and a version of Isaac's equation is derived.

1. INTRODUCTION.

A. This paper concerns the control of a system represented by a stochastic differential equation of the form

$$dz_t = g(t,z,u)dt + \sigma(t,z)dB_t \tag{1.1}$$

where z_t is the state at time t and the increments $\{dB_t\}$ are "gaussian white noise". The control u is to be chosen so as to minimize the average cost

$$J(u) = E \int_0^T c(t,z,u)dt. \tag{1.2}$$

Here T is either a fixed time or a bounded random time. The solution of (1.1) is defined by the "Girsanov measure transformation" method (see 1.D, 2 below) which permits a wide class of admissible controls. Controls based on three types of information pattern (partial and complete observation of the past, observation of the current state) are considered. In each case a principle of optimality similar to that of Rishel [13] is proved, and criteria for optimality analogous to the Hamilton-Jacobi equation of dynamic programming established by using an Ito process representation of the value function. Controls based on observation of the current state are shown to be minimizing in the class of those based on complete observation for system models of a suitable type. Finally similar methods are applied to 2-person zero sum differential games and a version of Isaac's equation derived.

The results presented here are closely related to those of Fleming on optimal control of diffusion processes. A brief outline of the latter is given in 1.B below in order to give the flavor of the former and for purposes of comparison. Some other possible approaches to stochastic control are mentioned; then, in the light of these, a more detailed statement of the contents of this paper will be found in 1.C.

B. Control of diffusion processes.

The results outlined here will be found in Fleming [9] and in the references there. Let $F \subset R^n$ be open and define the cylinder $Q \subset R^{n+1}$ by

$$Q = [0,1] \times F .$$

The system equation, to be solved in Q, is

$$(1.3) \qquad d\xi_t = g(t,\xi_t,u_t)dt + \sigma(t,\xi_t)dB_t$$
$$\xi_0 = x \in F.$$

$\{B_t\}$ is a separable n-vector Brownian motion process defined on some probability space $(\Omega,\mathcal{A},P)$. σ is an nxn-matrix-valued function on $[o,1]\times R^n$ and $g:[0,1]\times R^n\times R^\ell \to R^n$; both are of class C^2, with g,σ,g_x, σ_x bounded on $[0,1]\times R^n\times K$, for K compact in R^ℓ. Also there exists $c>0$ such that

$$(1.4) \qquad \sum_{i,j} a_{ij}(t,x,y)\mu_i\mu_j \geq c|\mu|^2$$

for each $\mu \in R^n$, where $a = \sigma\sigma'$ ($'$= transpose).

The control $u_t = Y(t,\xi_t)$ where Y is Lipschitz and takes values in $K \subset R^\ell$, compact. Thus the information pattern consists of complete observations of the current state. Under the above conditions (1.3) determines uniquely a diffusion process ξ on $[0,1]$ with $E|\xi_t|^2 < \infty$ for each t. The objective is to choose Y so as to minimize $$J(Y) = E \int_0^\tau c(t,\xi_t,Y[t,\xi_t])dt$$

where τ is the first exit time from Q. Let $g^Y(t,x) = g(t,x,Y[t,x])$ and similarly for c^Y. Define the differential operator

$$\Lambda\phi = \phi_t + 1/2\sum_{i,j} a_{ij}\phi_{x_i x_j} \tag{1.5}$$

and consider the boundary problem

$$\begin{aligned} &\Lambda\psi^Y + \psi_x^Y g^Y + c^Y = 0 \qquad (t,x) \in Q \\ &\psi^Y = 0, \quad (t,x) \in \partial'Q = \partial Q - \{0\}\times F \end{aligned} \tag{1.6}$$

Under the stated conditions this has a unique solution with the required differentiability properties. Applying Ito's differential formula to the function $\psi^Y(t,\xi_t)$, where ξ_t is the solution of (1.3) with $u_t = Y(t,\xi_t)$, gives

$$\psi^Y(0,x) = E\int_0^\tau c^Y dt = J(Y).$$

One can now drop the probabilistic interpretation and regard the problem as that of choosing the coefficients of the partial differential equation (1.6) so as to minimize the initial value $\psi^Y(0,x)$.

Let $U(s,x)$ be the minimum cost, over the class of admissible controls, starting at $(s,x) \in Q$. Formal application of Bellman's principle of optimality leads to the Hamilton-Jacobi equation

$$\begin{aligned} U(t,x) + \min_Y \{U_x(t,x)g^Y(t,x) + c^Y(t,x)\} &= 0 \quad \text{in } Q \\ U(t,x) &= 0 \quad \text{on } \partial'Q \end{aligned} \tag{1.7}$$

Fleming's "verification theorem"[9,Thm.6.1] says that if

(i) $\phi(t,x)$ is a suitably smooth solution of (1.7), and

(ii) $u^o = Y^o(t,x)$ is characterized by the property that $[\phi_x(t,x)g(t,x,v) + c(t,x,v)]$ is minimized in Q by $v = Y^o(t,x)$, then $$\phi(t,x) = U(t,x) = \psi^{Y^o}(t,x)$$

A unique solution of (1.7) satisfying the conditions of the verification theorem exists if a is bounded and Lipschitz and satisfies the uniform ellipticity condition (1.4), K is compact and convex, and the boundary ∂F has certain smoothness properties. (It is possible to relax these conditions somewhat.)

The above theory can be generalized in various ways. Let $C = C_n[0,1]$ be the space of continuous functions on $[0,1]$ with values in R^n. Let $\mathfrak{F}_t$ be the σ-field in C generated by the cylinder sets $\{z \varepsilon C: z_s \varepsilon \Gamma\}$ where Γ is a Borel set in R^n and $s \leq t$. For $z \varepsilon C$, $t \varepsilon [0,1]$, define $\pi_t z \varepsilon C$ by

$$\begin{aligned} \pi_t z(s) &= z_s \qquad s \leq t \\ &= z_t \qquad s > t \end{aligned}$$

A function of the form $u_t = u(t,\pi_t z)$ is "non-anticipative" in that it is adapted to $\mathfrak{F}_t$. A unique solution to (1.3) using a non-anticipative control u is obtained if u is Lipschitz, i.e.

$$|u(t,\pi_t \eta) - u(t,\pi_t \xi)| \leq \kappa ||\eta - \xi||$$

where $|| \cdot ||$ is the uniform norm in C. See [16]. Here the information pattern is complete observation of the past. It turns out that Markov controls are minimizing in the class of non-anticipative controls, so the Markov theory is the natural one in the case of complete observations. The partially observable Markov case where $u_t = Y(t,\xi_t)$ only depends on certain components of ξ_t can be considered though this is a somewhat artificial problem since the controller generally does better by using all the past observations, not just the current value.

D. Methods of the type outlined above suffer from two main drawbacks:

(i) The dependence of the admissible controls on the observations has to be "smooth"(e.g. Lipschitz) to insure the existence of a solution; for optimal control it is undesirable to be limited in this way.

(ii) The observation σ-fields $\{\mathcal{Y}_t^{(u)}\}$ depend on what control is being used. This tends to vitiate variational methods since varying the control at a certain time affects the admissibility of controls applied at subsequent times. There are two cases where this does not apply: (a) complete observations, as above, since then the observation σ-fields are those generated by the Brownian process.
(b) Linear systems of the form

$$(1.8)\qquad \begin{aligned} dx_t &= A_t x_t dt + u_t dt + dB_1(t) \\ dy_t &= F_t x_t dt + dB_2(t). \end{aligned}$$

In this case $\mathcal{Y}_t^{(u)} = \mathcal{Y}_t^o$, where $\mathcal{Y}_t^o$ are the σ-fields generated by $\{y_t^o\}$,$\{x_t^o, y_t^o\}$ being the solution of (1.8) with $u=0$. This is the basic fact behind the separation theorem [18],[19] which says that an optimal control is of the form $u_t = u(t,\hat{x}_t)$, where $\hat{x}_t = E[x_t|\mathcal{Y}_t^o]$. Of course, one can define other problems where the observation σ-fields do not depend on the control (this amounts to observing some function of the noise).Then variational methods can be used; see for example [20]. But then the problem loses its feedback aspects.

The system equation (1.1) treated in this report is more general than (1.3) in that the dependence of the matrices g and σ on the state is non-anticipative rather than "Markov". The method of solution - given in Section 2 - is designed to avoid (i) and (ii) above. In §C above one took a measurable space $(\Omega, \mathcal{A})$ and random variables $\{B_t\}$ constituting a Brownian motion under a measure P, and defined a transformation (1.3) $B \to \xi$ of the random variables. Here a transformation $P \to P_u$ of the measure is defined such that the original random variables generate under P_u the measure in their sample space which is appropriate for the solution of (1.1). This transformation is well-defined with only minimal restrictions on the class of admissible controls, and the observation σ-fields do not depend on the control since they are always generated by the same random variables. On the other hand the most that is claimed for the "solution" is that it has the right distributions.

Skorokhod remarks in the introduction to his book [5] that the methods of probability theory fall into two distinct groups, analytic and probabilistic, the former having to do only with the distributions of random variables, the latter based on operations with the random variables themselves. The method used here is something of a half-way house in that, while it is the distributions one is concerned with, the techniques used to derive them are definitely probabilistic.

This method has previously been used in [7],[8] (on the existence of optimal controls in the case of complete observations) and [19] (on the "separation theorem" of (b) above).

With a solution defined for each admissible control, the objective is to derive Hamilton-Jacobi-type conditions for optimality for the system (1.1) analogous to (1.7).

In [13] Rishel developed dynamic programming conditions for a very general class of stochastic systems. In Section 3 below, Rishel's "principle of optimality" is proved in the present context and in Section 4 conditions for optimality close to Rishel's are established. Using the special structure here these can be recast (Thm. 4.3) in a form close to that of (1.7) above.

In Section 5 the same methods are applied to the (simpler) completely observable case.

Section 6 deals with Markov control of systems similar to (1.3) (but without the technical conditions). The results here are direct extensions of those of §1B above, coinciding with the latter when the relevant conditions are satisfied.

Differential games are susceptible to attack by the same methods. The paper concludes with a brief section (section 7) outlining some of the possibilities in this direction.

2. PRELIMINARIES.

In differential form, the system equation (1.1) is

(2.1) $$dz_t = g(t,z,u_t)dt + \sigma(t,z)dw_t$$

with initial condition $z(0) = z_o \in R^n$, a fixed value. Here $t \in [0,1]$ and for each t, $z_t \in R^n$, $w_t \in R^n$, $u_t \in R^\ell$. When necessary g and σ will be written $g' = (g_1', g_2')$ and $\sigma' = (\sigma_1', \sigma_2')$ (' = transpose) corresponding to $z_t' = (x_t', y_t')$, the "state" and "observation" processes with dimensions (n-m),m, respectively.

Let $\{B_t, t \in [0,1]\}$ be an n-dimensional separable Brownian motion process on some probability space $(\Omega, \mathcal{A}, \mu)$. For each t let $\mathcal{A}_t \subset \mathcal{A}$ be the σ-field generated by the random variables $\{B_s, 0 \leq s \leq t\}$. Consider the stochastic differential equation

(2.2)
$$dz_t = \sigma(t,z)dB_t$$
$$z(0) = z_o .$$

The following properties are assumed for the n×n matrix-valued function $\sigma = [\sigma_{ij}]$:

(2.3)

(i) The elements $\sigma_{ij}:[0,1]\times C \to R$ are jointly measurable functions; $\sigma_{ij}(t,\cdot)$ is $\mathcal{F}_t$-measurable for each t.†

(ii) There exists a process z_t, $t \in [0,1]$, adapted to $\mathcal{A}_t$, satisfying (2.2) and

$$\sum_{i,j} \int_0^1 \sigma_{ij}^2(t,z)dt < \infty$$

† C and $\{\mathcal{F}_t\}$ were defined in section 1.B.

(iii) $\sigma(t,z)$ is non-singular, and $\sigma_2(t,z)$ has rank m, for almost all (t,z).

In (ii), the process $\{z_t\}$ is assumed to be unique in the following sense: all solutions of (2.2) , which must necessarily have continuous sample paths, generate the same measure in the sample space $(C,\mathcal{F})$. This is the definition used by Girsanov in [11]. Conditions (ii) and (iii) are given in the form in which they are required rather than in such a form as to be easily verified; any sufficient conditions insuring their satisfaction could be imposed.

Under the condtions (2.3), (2.2) defines a measure P on $(C,\mathcal{F})$ by

$$PF = \mu[z^{-1}(F)] \qquad \text{for } F \in \mathcal{F} .$$

Observe that

$$(2.4) \qquad \mathcal{A}_t = z^{-1}(\mathcal{F}_t)$$

for each t, since

$$w_t = \int_0^t \sigma^{-1} dz .$$

The function g satisfies the following conditions:

(2.5)

(i) $g:[0,1]\times C\times\Xi \to R^n$ is jointly measurable. (Here Ξ, the control set, is a Borel set of R^{ℓ}).

(ii) For fixed (t,u), $g(t,\cdot,u)$ is adapted to $\mathcal{F}_t$.

(iii) For all (t,z,u),

$$|\sigma^{-1}(t,z)g(t,z,u)| \le g^{\circ}(\|z\|)$$

where $||\cdot||$ is the uniform norm in C and g^o is an increasing real-valued function. Thus

$$\int_0^1 |\sigma^{-1}g|^2 dt \leq \left(g^o(||z||)\right)^2 < \infty \qquad \text{a.s.}(P).$$

Admissible controls. The class of admissible controls is denoted by $\mathcal{U}$ and defined as follows. For s,t ε [0,1], s<t, let $\mathcal{U}_s^t$ be the class of functions satisfying (2.6) below.

(i) $u:[s,t]\times C \to \Xi \subset R^\ell$ is jointly measurable in (t,z).

(2.6) (ii) For each t, u(t,·) is adapted to $\mathcal{Y}_t$.

(iii) $E[\rho_s^t(u)|\mathcal{F}_s] = 1$ a.s.(P).

Here $\rho_s^t(u) = \exp[\zeta_s^t(g^{(u)})]$, where $g^{(u)}(t,z) = g(t,z,u[t,z])$ and $\zeta_s^t(g^{(u)})$ is defined by

$$\zeta_s^t(g^{(u)}) = \int_s^t \left(\sigma^{-1}(\tau,z)g(\tau,z,u[\tau,z])\right)' dB_\tau - \frac{1}{2}\int_s^t |\sigma^{-1}(\tau,z)g(\tau,z,u[\tau,z])|^2 d\tau$$

(2.7) $$= \int_s^t \left(\sigma^{-1}g\right)'\sigma^{-1}dz - \frac{1}{2}\int_s^t |\sigma^{-1}g|^2 d\tau .$$

Now define $\mathcal{U} = \mathcal{U}_0^1$.

From(2.7), $\zeta_s^t(u)$ can be computed directly from $\{z_\tau, 0 \leq \tau \leq t\}$. Thus $\rho_s^t(u)$ can be regarded as a random variable on the probability space $(C,\mathcal{F},P)$; in fact this is taken as the basic space from now on, the symbol E referring, as in (2.6)(iii), to integration with respect to the measure P. It is shown in [11] that (2.5) and (2.6)(i,ii) imply

$$E[\rho_s^t(u)|\mathcal{F}_s] \leq 1 \qquad \text{a.s.}$$

There is no known criterion for equality, though various sufficient conditions have been derived; see [8].

Remarks. 1. If $u' \in \mathcal{U}_r^s$ and $u'' \in \mathcal{U}_s^t$, where $r \leq s \leq t$, then $u \in \mathcal{U}_r^t$, where

$$\begin{aligned} u(\tau,z) &= u'(\tau,z) \qquad \tau \in [r,s) \\ &= u''(\tau,z) \qquad \tau \in [s,t]. \end{aligned}$$

Indeed, u clearly satisfies (2.6)(i),(ii), and

$$\begin{aligned} E[\rho_{t'}^{t''}(u)|\mathcal{F}_{t'}] &= E[\rho_{t'}^{s}(u')\ E\{\rho_s^{t''}(u'')|\mathcal{F}_s\}|\mathcal{F}_{t'}] \\ &= E[\rho_{t'}^{s}(u')|\mathcal{F}_{t'}] = 1 \quad \text{a.s.} \end{aligned}$$

for $r \leq t' \leq s \leq t'' \leq t$. The other cases work similarly.

2. If $u \in \mathcal{U}$, then u restricted to [s,t] belongs to $\mathcal{U}_s^t$. This follows from Lemma 2 of [11].

Theorem 2.1 (Girsanov) For $u \in \mathcal{U}$ let the measure P_u on $(C,\mathcal{F})$ be defined by

$$P_u F = \int_F \rho_0^1(u) dP \quad , \qquad F \in \mathcal{F}$$

Then (a) $dw = dB - \sigma^{-1}g\ dt$ is a Brownian motion process under the measure μ_u defined by

$$\mu_u[z^{-1}(F)] = P_u F \ .$$

(This defines μ_u for each $A \in \mathcal{A}$ in view of (2.4)).

(b) The process $\{z_t\}$ satisfies

$$dz_t = g(t,z,u[t,z])dt + \sigma(t,z)dw_t \tag{2.8}$$

$$z(0) = z_o \ .$$

This result is immediate from Girsanov's Theorem 1 [11]. Lemma 6 of [11] states that if $\{\theta_t\}$ is adapted to $\mathcal{A}_t$ and $\int|\theta_t|^2 dt < \infty$ a.s. then

$$\int_0^t \theta_s dB_s = \int_0^t \theta_s dW_s + \int_0^t \theta_s \sigma^{-1} g \, ds .$$

Putting $\theta_t = \sigma(t,z)$, (2.8) follows from (i) above and (2.2).

Theorem 2.1 shows that the process $\{z_t\}$ is, with measure P_u, a solution of (2.1) in the sense that

$$dz = g \, dt + \sigma \, d(\text{Brownian motion}).$$

Remark: All measures arising in this paper are, by definition, mutually absolutely continuous with respect to the measure P; so when some property is stated to hold "almost surely" (a.s.), it is irrelevant which measure is referred to.

Let $c:[0,1]\times C\times\Xi \to R^+$ be a non-negative real valued function satisfying (2.5)(i),(ii) and

(2.9) $$c(t,z,u) \leq k \quad \text{for all } (t,z,u) \in [0,1]\times C\times\Xi$$

where k is a real constant. The cost ascribed to an admissible control u is

(2.10) $$J(u) = E_u\left[\int_0^1 c(s,z,u(s,z))ds\right] = E\left[\rho_0^1(u)\int_0^1 c_s^{(u)} ds\right] .$$

Note that this allows for a random stopping time τ as long as $\tau \leq 1$ a.s. For $c' = c.I_{[\tau \geq t]}$ is an admissible cost rate function and

$$E_u \int_0^\tau c \, ds = E_u \int_0^1 c' \, ds .$$

The following results will be required in subsequent sections.

A. Compactness of the set of densities.

Let $\mathcal{G}$ be the set of measurable functions $\gamma:[0,1]\times C \to R^n$ adapted to $\mathcal{F}_t$ and satisfying:

$$(2.11)\quad \begin{aligned} &\text{(i)} \quad |\sigma^{-1}(t,z)\gamma(t,z)| \le g^{\circ}(||z||) \\ &\text{(ii)} \quad E[\exp\{\zeta_0^1(\gamma)\}] = 1. \end{aligned}$$

Let $\mathcal{D} = \{\exp[\zeta_0^1(\gamma) : \gamma \in \mathcal{G}\}$.

Theorem 2.2 $\mathcal{D}$ is a weakly compact subset of $L_1(C,\mathcal{F},P)$.

This result is contained in Theorem 2 of [8] for the case $\sigma = I$ (the identity matrix). Only minor modifications are required to establish the result as stated.

Note that $\rho_0^1(u) \in \mathcal{D}$ for each $u \in \mathcal{U}$. Thus for any sequence $u_n \in \mathcal{U}$ there is a subsequence $\{u_{n_k}\}$ and an element $h \in \mathcal{G}$ such that

$$\rho_0^1(u_{n_k}) \to \exp[\zeta_0^1(h)]$$

weakly in L_1 as $k \to \infty$.

B. Innovations process and representation of Martingales.

The main result here is Theorem 2.3, which says that any martingale adapted to $\mathcal{Y}_t$ has a representation as a stochastic integral with respect to the "innovations process" of $\{y_t\}$, defined below. This definition was given in [15]. The result is proved in [10] for the case $\sigma = I$; the following is a similar method of proof using also ideas from [15].

Let $\gamma \in \mathcal{G}$, $\gamma = (\bar{h}, h)$ with dimensions m-n, n, and define the measure P^* on $(C,\mathcal{F})$ by the Radon-Nikodym derivative

$$dP^* = \exp[\zeta_0^1(\gamma)]\, dP .$$

P^* is a probability measure in view of (2.11) and from Girsanov's theorem the process $\{z_t\}$ satisfies

(2.12) $$dz_t = \gamma_t dt + \sigma_t dw_t$$

where $(w_t, \mathcal{F}_t, P^*)$ is a Brownian motion. From (2.12), the observation process $\{y_t\}$ satisfies

(2.13) $$dy_t = h_t dt + \sigma_2(t) dw_t$$

and

(2.14) $$\int_0^1 |h_t|^2\, dt < \infty \quad \text{a.s.}$$

Choose any vector $\theta \in R^m$ and let $\xi_t = \theta' y_t$. Applying Ito's differential formula to the function $F(\xi) = \xi^2$ gives

$$\int_0^t \theta'\sigma_2\sigma_2'\theta\, ds = \xi_t^2 - \xi_0^2 - 2\int_0^t \xi_s d\xi_s .$$

This shows that the symmetric positive definite matrix $\sigma_2(t)\sigma_2'(t)$ is $\mathcal{Y}_t$-measurable for each t. Thus there exists a unitary matrix Q_t and a diagonal matrix L_t ,both $\mathcal{Y}_t$-measurable, such that

(2.15) $$\sigma_2(t)\sigma_2'(t) = Q_t L_t Q_t' .$$

Now define
$$T_t = (L_t)^{-1/2} Q_t'$$
$$\hat{h}_t = E^*[h_t | \mathcal{Y}_t]$$
$$\tilde{h}_t = h_t - \hat{h}_t$$

($E^*[\cdot|\mathcal{Y}_t]$ denotes conditional expectation with respect to P^*.)

The <u>innovations process</u> $\{\nu_t\}$ is defined by

(2.17) $$d\nu_t = T_t(dy_t - \hat{h}_t dt)$$
$$= T_t(\sigma_2(t) dw_t + \tilde{h}_t dt)$$

Lemma 2.1 $(\nu_t, \mathcal{Y}_t, P^*)$ is a Brownian motion process.

Proof: It is evident from the definition that $\{\nu_t\}$ is adapted to $\mathcal{Y}_t$ and has almost all sample paths continuous. Again pick $\theta \in R^m$. In view of (2.15) and (2.16), for each t,

$$T_t\sigma_2(t)\sigma_2'(t)T_t' = I \tag{2.18}$$

so that applying Ito's differential formula to the function $f(\nu) = e^{i\theta'\nu}$ gives (using (2.17)),

$$e^{i\theta'\nu(t)} - e^{i\theta'\nu(s)} = \int_s^t i\theta' e^{i\theta'\nu(\tau)} T_\tau \tilde{h}_\tau d\tau + \int_s^t (-\tfrac{1}{2}|\theta|^2) e^{i\theta'\nu(\tau)} d\tau$$
$$+ \int_s^t i\theta' e^{i\theta'\nu(\tau)} T_\tau \sigma_2(\tau) dw_\tau$$

Now,

$$E^*[i\theta' e^{i\theta'\nu(\tau)} T_\tau \tilde{h}_\tau | \mathcal{Y}_s] = E^*\{i\theta' e^{i\theta'\nu(\tau)} T_\tau \; E^*[h_\tau - \hat{h}_\tau | \mathcal{Y}_\tau] | \mathcal{Y}_s\}$$
$$= 0 \quad \text{a.s.},$$

and,

$$E^*[\int_s^t i\theta' e^{i\theta'\nu(\tau)} T_\tau \sigma_2(\tau) dw_\tau | \mathcal{Y}_s] = 0 \quad \text{a.s.}$$

Thus,

$$E^*[e^{i\theta'\nu(t)} - e^{i\theta'\nu(s)} | \mathcal{Y}_s] = E^*[\int_s^t (-\tfrac{1}{2}|\theta|^2) e^{i\theta'\nu(\tau)} d_\tau | \mathcal{Y}_s] ,$$

or, alternatively,

$$E^*[e^{i\theta'\{\nu(t) - \nu(s)\}} - 1 | \mathcal{Y}_s] = -\tfrac{1}{2}|\theta|^2 E^*[\int_s^t e^{i\theta'\{\nu(\tau) - \nu(s)\}} d_\tau | \mathcal{Y}_s] \tag{2.19}$$

Pick $A \in \mathcal{Y}_s$ and define

$$\eta_t = \int_A e^{i\theta'\{\nu(t) - \nu(s)\}} dP^* .$$

Then from (2.19),

$$\eta_t = P^*A - \frac{1}{2}|\theta|^2 \int_A \int_s^t e^{i\theta'\{\nu(\tau)-\nu(s)\}} \, d\tau \, dP^*$$

$$= P^*A - \frac{1}{2}|\theta|^2 \int_s^t \eta_\tau \, d\tau \, .$$

This integral equation has the unique solution

$$\eta_t = P^*A \, e^{-(1/2)|\theta|^2(t-s)}$$

from which it is immediate that

$$E^*[e^{i\theta'\{\nu(t)-\nu(s)\}} | \mathcal{Y}_s] = e^{-(1/2)|\theta|^2(t-s)} \quad .$$

The statement of the lemma follows from this.

Theorem 2.3 Suppose $(M_t, \mathcal{Y}_t, P^*)$ is a martingale. Then there exists a process $\{\psi_t\}$ adapted to $\mathcal{Y}_t$ such that

$$\int_0^1 |\psi_t|^2 dt < \infty \qquad \text{a.s.}$$

and

(2.20) $$M_t = M_0 + \int_0^t \psi_s d\nu_s \quad .$$

Proof: M_0 is a constant a.s. since $\mathcal{Y}_0 = \{C, \emptyset\}$ .For convenience assume that $E^*M_t = M_0 = 0$. For n=1,2... define

$$\tau_n = \min\left(1, \ \inf\{t: \int_0^t |T_s \hat{h}_s|^2 ds \geq n\} \right)$$

This is a stopping time of $\mathcal{Y}_t$, and $\tau_n \uparrow 1$ a.s. from (2.11)(i). Now define

$$\pi_t = \exp\left(\int_0^t (-T_s \hat{h}_s) d\nu_s - \frac{1}{2} \int_0^t |T_s \hat{h}_s|^2 ds \right)$$

and define the measure $\tilde{P}_n$ by

$$d\tilde{P}_n = \pi_{t \wedge \tau_n} \, dP^* \, .$$

From Girsanov's theorem, $\tilde{P}_n$ is a probability measure for each n, and the process

$$(2.21) \qquad Y_t^n = \nu_t + \int_0^{t\wedge\tau_n} T_s \hat{h}_s \, ds$$

is a Brownian motion under $\tilde{P}_n$. Let

$$\mathcal{Y}_t^n = \sigma\{Y_s^n \, , \, 0 \le s \le t\} \ .$$

By Theorem 3 of [6], if $(\tilde{M}_t, \mathcal{Y}_t^n, \tilde{P}_n)$ is a separable martingale then it has continuous sample paths and has the representation

$$\tilde{M}_t = \int_0^t \phi_s^n \, dY_s^n \ .$$

Observe from (2.17) and (2.21) that $Y_t^n = \int_0^t T_s dy_s$ for $t < \tau_n$ and hence that

$$\mathcal{Y}_t^n = \mathcal{Y}_{t\wedge\tau_n} \ .$$

Thus if $(\tilde{M}_t, \mathcal{Y}_{t\wedge\tau_n}, \tilde{P}_n)$ is a martingale,

$$(2.22) \qquad \tilde{M}_{t\wedge\tau_n} = \int_0^{t\wedge\tau_n} \phi_s^n T_s \, dy_s \ .$$

Now suppose $(M_t, \mathcal{Y}_t, P^*)$ is a separable martingale. Then $(\tilde{M}_t, \mathcal{Y}_{t\wedge\tau_n}, \tilde{P}_n)$ is a martingale, where

$$(2.23) \qquad \tilde{M}_t = M_{t\wedge\tau_n}(\pi_{t\wedge\tau_n})^{-1} \ .$$

Indeed, $(M_{t\wedge\tau_n}, \mathcal{Y}_{t\wedge\tau_n}, P^*)$ is a martingale by the optional sampling theorem, and, denoting integration with respect to $\tilde{P}_n$ by $\tilde{E}_n$,

$$\tilde{E}_n[\tilde{M}_t | \mathcal{Y}_{s\wedge\tau_n}] = \tilde{E}_n[M_{t\wedge\tau_n}(\pi_{t\wedge\tau_n})^{-1} | \mathcal{Y}_{s\wedge\tau_n}]$$

$$= \frac{E^*[M_{t\wedge\tau_n}(\pi_{t\wedge\tau_n})^{-1}\pi_{\tau_n} | \mathcal{Y}_{s\wedge\tau_n}]}{E^*[\pi_{\tau_n} | \mathcal{Y}_{s\wedge\tau_n}]}$$

$$= \frac{E^*[M_{t\wedge\tau_n} | \mathcal{Y}_{s\wedge\tau_n}]}{\pi_{s\wedge\tau_n}}$$

$$= M_{s\wedge\tau_n}(\pi_{s\wedge\tau_n})^{-1} = \tilde{M}_s \ .$$

In this case $\tilde{M}_t = \tilde{M}_{t\wedge\tau_n}$ so that from (2.22)

$$\tilde{M}_t = \int_0^{t\wedge\tau_n} \phi_s^n T_s \, dy_s$$

$$= \int_0^{t\wedge\tau_n} \phi_s^n \, d\nu_s + \int_0^{t\wedge\tau_n} \phi_s^n \, T_s \hat{h}_s \, ds \ . \tag{2.24}$$

Now $\pi_{t\wedge\tau_n}$ satisfies the Ito equation

$$\pi_{t\wedge\tau_n} = 1 - \int_0^{t\wedge\tau_n} \pi_s T_s \hat{h}_s \, d\nu_s \ . \tag{2.25}$$

Applying the Ito differential formula to the product in (2.23), using (2.24),(2.25) , gives

$$M_{t\wedge\tau_n} = \int_0^{t\wedge\tau_n} \psi_s^n \, d\nu_s$$

where
$$\psi_s^n = \pi_s(\phi_s^n - \tilde{M}_s T_s \hat{h}_s) \ .$$

Such a representation is clearly unique, so that

$$\psi_s^n = \psi_s^{n'} \qquad \text{for } n' \geq n, \ s \leq \tau_n.$$

If ψ is the function which, for each n, agrees with ψ^n on $[s < \tau_n]$, then

$$M_{t\wedge\tau_n} = \int_0^{t\wedge\tau_n} \psi_s \, d\nu_s$$

i.e.

$$M_t = \int_0^t \psi_s d\nu_s \tag{2.26}$$

on $[t<\tau_n]$ for each n. Thus (2.26) holds a.s. for each t since $\tau_n \uparrow 1$ a.s.

3. VALUE FUNCTION AND PRINCIPLE OF OPTIMALITY.

The results here are similar to those of Rishel [13]. The value function W_u is defined by (3.3) below and shown in Theorem 3.1 to satisfy a version of Bellman's principle of optimality. In Rishel's paper this depended on the class of controls satisfying a condition called "relative completeness". Here it turns out (Lemma 3.1) that this condition is always satisfied.

Suppose control $u \varepsilon \mathcal{U}_0^t$ is used on $[0,t]$ and $v \varepsilon \mathcal{U}_t^1$ on $(t,1]$. Then the expected remaining cost at time t, given the observations up to that time, is

$$(3.1) \qquad \psi_{uv}(t) = E_{uv}[\int_t^1 c_s ds \,|\, \mathcal{Y}_t] = \frac{E[\rho_0^t(u)\rho_t^1(v)\int_t^1 c_s ds \,|\, \mathcal{Y}_t]}{E[\rho_0^1 \,|\, \mathcal{Y}_t]} \quad \text{a.s.}$$

See [3,§24.2]. Define

$$f_{uv}(t) = E[\rho_0^t(u)\rho_t^1(v)\int_t^1 c_s ds \,|\, \mathcal{Y}_t] .$$

The notations $\psi_u = \psi_{uu}$ and $f_u = f_{uu}$ for $u \varepsilon \mathcal{U}_0^1$ will also be used. Now $f_{uv}(t) \varepsilon L_1(C,\mathcal{F},P)$ since $f_{uv} \geq 0$ a.s. and

$$E\, f_{uv}(t) = E[\rho_0^t(u)\rho_t^1(v)\int_t^1 c_s ds] \leq k(1-t)$$

from (2.9). L_1 is a complete lattice [2,p.302] under the partial ordering $f_1 \prec f_2 \iff f_1(z) \leq f_2(z)$ a.s. The set $\{f_{uv}(t): v \varepsilon \mathcal{U}_t^1\}$ is bounded from below by the zero function, so the following infimum exists in L_1 for each t.

$$(3.2)\qquad V(u,t) = \bigwedge_{v\,\varepsilon\,\mathcal{U}_t^1} f_{uv}(t)$$

Notice that the "normalizing factor" $E[\rho_0^1|\mathcal{Y}_t] = E[\rho_0^t|\mathcal{Y}_t]$ does not depend on v. The value function $W_u(t)$ is thus defined as

$$(3.3)\qquad W_u(t) = \bigwedge_{v\,\varepsilon\,\mathcal{U}_t^1} E_{uv}[\int_t^1 c_s ds|\mathcal{Y}_t] = \frac{1}{E[\rho_0^t(u)|\mathcal{Y}_t]}\, V(u,t).$$

Thus $V(u,t)$ is an unnormalized version of the g.l.b. of the expected additional cost at time t. Suppose $\Theta\ \varepsilon\ L_1$, $V(u,t) \prec \Theta$; i.e. $V(u,t)(z) < \Theta(z)$ for $z\ \varepsilon\ M$, $PM>0$. Then there exists $v\ \varepsilon\ \mathcal{U}_t^1$ and a set M_v with $PM_v > 0$ such that $f_{uv}(t,z) < \Theta(z)$ for $z\ \varepsilon\ M_v$. The class $\mathcal{U}$ is said to be relatively complete [13] if for any $t\ \varepsilon\ [0,1]$ and $\varepsilon>0$ there exists $v\ \varepsilon\ \mathcal{U}_t^1$ such that

$$f_{uv}(t) < V(u,t) + \varepsilon \qquad \text{a.s.}$$

This amounts to saying, in the above, that for $\Theta = V + \varepsilon$ there is a v with $PM_v = 1$. The fact (Lemma 3.1) that this is true is used in the proof of Theorem 3.1.

Lemma 3.1. $\mathcal{U}$ is relatively complete.

Proof: Fix $\varepsilon>0, t\ \varepsilon\ [0,1]$ and $u\ \varepsilon\ \mathcal{U}_0^t$. Let $V(z) = V(t,u)(z)$ and for $v \in \mathcal{U}_t^1$ let

$$M_v = \{z: f_{uv}(z) < V(z) + \varepsilon\} \subset \mathcal{Y}_t .$$

A partial ordering is defined on the set $X = \{(v,M_v): v\ \varepsilon\ \mathcal{U}_t^1\}$. Then Zorn's Lemma is used to establish the existence of a maximal element (v^*,M_{v^*}) which has the property that $PM_{v^*} = 1$, proving the lemma.

The partial ordering $\succ$ on X is as follows:

$$(u,M_u) \succ (v,M_v) \quad \text{if and only if}$$

(3.4)
$$\text{(i)}\ M_u \supset M_v$$
$$\text{(ii)}\ PM_u > PM_v$$
$$\text{(iii)}\ u \text{ and } v \text{ agree on } M_v .$$

Each chain in $(X,\succ)$ has an upper bound. Indeed, let $\{(v_\alpha,M_\alpha):\alpha \in A\}$ be a chain in $(X,\succ)$.

1. If for some $\alpha_1 \in A$, $PM_{\alpha_1} = \sup_{\alpha \in A} \{PM_\alpha\}$, then $(u_{\alpha_1},M_{\alpha_1})$ is the upper bound.

2. If $M = PM_{\alpha_1} < \sup_{\alpha \in A} \{PM_\alpha\}$ for each $\alpha_1 \in A_1$ then $PM_\alpha \uparrow m \leq 1$. For each $n=1,2\ldots$ pick α_n such that

$$PM_{\alpha n} > m - \frac{1}{n}$$

Let $M = \bigcup_{\alpha\varepsilon A} M$. Then $M = \bigcup_{n=1}^{\infty} M_{\alpha_n}$; clearly $M \supset \bigcup_{n=1}^{\infty} M_{\alpha_n}$ and conversely, given $\alpha \varepsilon A$, $PM_\alpha < m - \frac{1}{n'}$ for some integer n' and hence $M_\alpha \subset M_{\alpha n'} \subset \bigcup_{n=1}^{\infty} M_{\alpha_n}$. Thus M is $\mathcal{Y}_t$-measurable and $PM = m$.

3.Define the control v on $t \leq \tau \leq 1$ as follows:

$$v(\tau,z) = v_{\alpha_n}(\tau,z) \qquad z \varepsilon M_{\alpha_n} .$$

This specifies v on M ; on M^c let $v(\tau,z) = v_{\alpha_1}(\tau,z)$. v is clearly measurable; and

$$\{z: v(\tau,z) \varepsilon \Gamma\} = \bigcup_{i=1}^{\infty} M_{\alpha_i} \cap \{z:v_{\alpha_i}(\tau,z) \varepsilon \Gamma\} \quad \varepsilon\ \mathcal{Y}_\tau$$

so that v is adapted to $\mathcal{Y}_t$. Let $M_1 = M_{\alpha_1} \cup M^c$, $M_i = M_{\alpha_i} - \bigcup_{j=1}^{i-1} M_j$ for $i=1,2,3\ldots$ Then $\{M_i\}$ is a partition of C into $\mathcal{Y}_t$-measurable sets. Hence

$$E[\rho_t^1(v) | \mathcal{Y}_t] = \sum_{i=1}^{\infty} E[I_{M_i}\ \rho_t^1(v_{\alpha_i}) | \mathcal{Y}_t]$$

$$= \sum_{i=1}^{\infty} I_{M_i} \, E[\rho_t^1(v_{\alpha_i}) \,|\, \mathcal{Y}_t]$$

$$= \sum_{i=1}^{\infty} I_{M_i} \quad = \quad 1 \quad \text{a.s.}$$

Thus $v \in \mathcal{U}_+^1$.

4. (v,M) is an upper bound for $\{(v_\alpha, M_\alpha) : \alpha \in A\}$.

(a) $(v,M) \in X$; i.e. $M = M_v = \{z : f_{uv}(z) < V(z) + \varepsilon\}$, since $z \in M$ $\Rightarrow z \in M_{\alpha_i}$ for some i $\Rightarrow v = v_{\alpha_i}$ $\Rightarrow f_{uv}(z) < V(z) + \varepsilon$, while $z \in M^c \Rightarrow v(z) = v_{\alpha_1}(z) \Rightarrow z \in M_v^c$ since $M_{\alpha_1} \subset M$.

(b) For any $\alpha \in A$, $(v,M) \succ (v_\alpha, M_\alpha)$. This is immediate from (3.4).

Since each chain in X has an upper bound, X has a maximal element, i.e. anelement (v^*, M^*) with the property that for each comparable $(v_\alpha, M_\alpha) \in X$,

$$(v^*, M^*) \succ (v_\alpha, M_\alpha).$$

It remains to show that $PM^* = 1$. Suppose $PM^* < 1$. Then $P(M^*)^c > 0$ so there exists $v' \in \mathcal{U}_t^1$ and a set $\Psi \subset (M^*)^c$ with $P\Psi > 0$ such that

$$f_{uv'}(z) \quad < \quad V(z) + \varepsilon, \quad z \in \Psi$$

Recall that Ψ, M are $\mathcal{Y}_t$ measurable. Define

$$\begin{aligned} v^o(t,z) &= v^*(t,z) \qquad && z \in M^* \\ &= v'(t,z) && z \in (M^*)^c \end{aligned}$$

This is admissible and

$$M^o = \{z : v^o(t,z) < V(z) + \varepsilon\} \supset M^* \cup \Psi$$

Thus $PM^o > PM$ and hence $(v^o, M^o) \succ (v^*, M^*)$, contradicting the maximality of (v^*, M^*). So $PM^* = 1$, as required.

<u>Theorem 3.1</u> For each $t \in [0,1]$ and $u \in \mathcal{U}_0^t$, the value function $W_u(t)$ satisfies the "principle of optimality":

$$(3.5) \qquad W_u(t) \leq E_u[\int_t^{t+h} c_s^{(u)} ds \mid \mathcal{Y}_t] + E_u[W_u(t+h) \mid \mathcal{Y}_t] \quad \text{a.s.}$$

for each $h > 0$.

<u>Proof</u>:

$$V(u,t) = \bigwedge_{v \in \mathcal{U}_t^1} E[\rho_0^t(u)\rho_t^1(v) \int_t^{t+h} c_s^{(v)} ds + \rho_0^t(u)\rho_t^1(v) \int_{t+h}^1 c_s^{(v)} ds \mid \mathcal{Y}_t]$$

$$(3.6) \qquad \leq E[\rho_0^{t+h}(u) \int_t^{t+h} c_s^{(u)} ds \mid \mathcal{Y}_t] + \bigwedge_{v \in \mathcal{U}_{t+h}^1} E[\rho_0^{t+h}(u)\rho_{t+h}^1(v) \int_{t+h}^1 c_s^{(v)} ds \mid \mathcal{Y}_t]$$

(Otherwise there would be a $v \in \mathcal{U}_{t+h}^1$ and $M \in \mathcal{Y}_t$ with $PM > 0$ such that

$$V(u,t) - E[\rho_0^{t+h}(u) \int_t^{t+h} c_s ds \mid \mathcal{Y}_t] > E[\rho_0^{t+h}(u)\rho_{t+h}^1(v) \int_{t+h}^1 c_s ds \mid \mathcal{Y} t]$$

for $z \in M$; i.e.

$$V(u,t) > E[\rho_0^{t+h}(u)\rho_{t+h}^1(v) \int_t^1 c_s ds \mid \mathcal{Y}_t] \quad \text{for } z \in M, \text{ a contradiction.})$$

The next stage is to show that

$$(3.7) \qquad E[V(u,t+h) \mid \mathcal{Y}_t] = \bigwedge_{v \in \mathcal{U}_{t+h}^1} E[\rho_0^{t+h}(u)\rho_{t+h}^1(v) \int_{t+h}^1 c_s^{(v)} ds \mid \mathcal{Y}_t] .$$

For any $v' \in \mathcal{U}_{t+h}^1$,

$$V(u,t+h) \leq E[\rho_0^{t+h}(u)\rho_{t+h}^1(v') \int_{t+h}^1 c_s^{(v')} ds \mid \mathcal{Y}_{t+h}] \text{ a.s.}$$

Hence,

$$E[V(u,t+h) \mid \mathcal{Y}_t] \leq E[\rho_0^{t+h}(u)\rho_{t+h}^1(v') \int_{t+h}^1 c_s^{(v')} ds \mid \mathcal{Y}_t] \quad \text{a.s.}$$

Therefore

$$E[V(u,t+h) \mid \mathcal{Y}_t] \leq \bigwedge_{v' \in \mathcal{U}_{t+h}^1} E[\rho_0^{t+h}(u)\rho_{t+h}^1(v') \int_{t+h}^1 c_s^{(v')} ds \mid \mathcal{Y}_t] \text{ a.s.}$$

Since the class $\mathcal{U}_t^1$ is relatively complete, given t+h, $u \in \mathcal{U}_0^{t+h}$, and $\varepsilon > 0$, there exists $u' \in \mathcal{U}_{t+h}^1$ such that

$$E[\rho_0^{t+h}(u)\rho_{t+h}^1(u')\int_{t+h}^1 c_s^{(u')}ds\,|\,\mathcal{Y}_{t+h}] \le V(u,t+h) + \varepsilon \quad \text{a.s.}$$

Then,

$$E[\rho_0^{t+h}(u)\rho_{t+h}^1(u')\int_{t+h}^1 c_s^{(u')}ds\,|\,\mathcal{Y}t] \le E[V(u,t+h)\,|\,\mathcal{Y}_t] + \varepsilon \quad \text{a.s.}$$

and thus

$$\bigwedge_{u' \in \mathcal{U}_{t+h}^1} E[\rho_0^{t+h}(u)\rho_{t+h}^1(u')\int_{t+h}^1 c_s^{(u')}ds\,|\,\mathcal{Y}_t] \le E[V(u,t+h)\,|\,\mathcal{Y}_t] \text{ a.s.}$$

This establishes (3.7). From (3.6) and (3.7),

$$V(u,t) \le E[\rho_0^{t+h}(u)\int_t^{t+h} c_s^{(u)}ds\,|\,\mathcal{Y}_t] + E[V(u,t+h)\,|\,\mathcal{Y}_t] \quad \text{a.s.}$$

Dividing this through by $E[\rho_0^t(u)\,|\,\mathcal{Y}_t]$ gives (3.5) after noting that

$$\frac{E[V(u,t+h)\,|\,\mathcal{Y}_t]}{E[\rho_0^t(u)\,|\,\mathcal{Y}_t]} = \frac{E[W_u(t+h)\;E\{\rho_0^{t+h}\,|\,\mathcal{Y}_{t+h}\}\,|\,\mathcal{Y}_t]}{E[\rho_0^t\,|\,\mathcal{Y}_t]}$$

$$= \frac{E[W_u(t+h)\rho_0^{t+h}\,|\,\mathcal{Y}_t]}{E[\rho_0^t\,|\,\mathcal{Y}_t]}$$

$$= E_u[W_u(t+h)\,|\,\mathcal{Y}_t]\ .$$

4. CONDITIONS FOR OPTIMALITY.

Two sets of criteria - Theorems 4.2 and 4.3 - are presented in this section. Theorem 4.2. is included for two reasons: it is used later in establishing other criteria, and it is the equivalent in the present context of Rishel's results (Theorems 8 and 9 of [13]). The process "$W_u(t)$" of Rishel's Theorem 9 corresponds to the process $Z_u(t)$ defined by (4.5) below.

The objective of this and the following sections is to get Hamilton-Jacobi-like criteria for optimality; i.e. a local characterization of the optimal policy in terms of the value function. The results are bound to be less than satisfactory in the case of partial observations as there is a different value function for each control: the expected remaining cost from a certain time on depends on what control was applied prior to that time. By restricting attention, as Rishel does, to "value decreasing" controls, in which class the optimal control, if it exists, must lie, one can get some way towards the above characterization. This is Theorem 4.3.

The case of "complete observations" is a great deal simpler as here there is only one value function. This case is treated in section 5.

The potential generated by an integrable increasing process $\{a_t\}$ is

$$b_t = E[a_1 | \mathcal{Y}_t] - a_t$$

It is easy to check that the process $\psi_u(t)$ of (3.1) is a potential under measure P_u.

Lemma 4.1 Under measure P_u, $\psi_u(t)$ is the potential generated by the integrable increasing process

$$a_t = \int_0^t E_u[c_s^{(u)} | \mathcal{Y}_s]\, ds. \tag{4.4}$$

Proof: $\{a_t\}$ is clearly increasing, positive, and adapted to $\mathcal{Y}_t$. Also $\sup_t E_u a_t \leq k$. It remains to show that

$$\psi_u(t) = E_u[a_1 | \mathcal{Y}_t] - a_t .$$

In the following, $c_s \equiv c_s^{(u)}$.

$$E_u[a_1 | \mathcal{Y}_t] - a_t = E_u[\int_0^t E_u(c_s | \mathcal{Y}_s) ds | \mathcal{Y}_t] + E_u[\int_t^1 E_u(c_s | \mathcal{Y}_s) ds | \mathcal{Y}_t] - a_t$$

$$= E_u[\int_t^1 E_u(c_s | \mathcal{Y}_s) ds | \mathcal{Y}_t]$$

$$= \int_t^1 E_u[c_s | \mathcal{Y}_t]\, ds$$

$$= E_u[\int_t^1 c_s\, ds | \mathcal{Y}_t] = \psi_u(t) .$$

The legitimacy of the interchanges of integration and conditional expectation in the above is easily seen in view of the boundedness of c .

__Theorem 4.2__ $u^* \in \mathcal{U}$ is optimal if and only if there exists a constant J^* and for each $u \in \mathcal{U}$ an integrable process $\{\alpha_u(t)\}$ adapted to $\mathcal{Y}_t$ and satisfying

(i) $$E_u \int_0^1 \alpha_u(s)ds = J^*$$

(ii) $$E_{u^*}[c_t^{(u^*)} | \mathcal{Y}_t] - \alpha_{u^*}(t) = 0$$

$$E_u[c_t^{(u)} | \mathcal{Y}_t] - \alpha_u(t) \geq 0 \qquad \text{for almost all } (t,z)$$

Then $J^* = J(u^*)$, the cost of the optimal policy.

__Proof:__ Suppose u^* is optimal. Let $J^* = J(u^*) = \psi_{u^*}(0)$. Define $\kappa_u = J^* (\psi_u(0))^{-1}$; thus $\kappa_u \leq 1$ and $\kappa_u = 1$ if u is optimal. Then the process

$$\alpha_u(t) = \kappa_u E_u[c_t^{(u)} | \mathcal{Y}_t]$$

is clearly integrable and in fact satisfies (i) and (ii). Indeed,

$$E_u \int_0^1 \alpha_u(s)ds = \kappa_u E_u[\int_0^1 c_s ds] = \kappa_u \psi_u(0) = J^*$$

and

$$E_u[c_t | \mathcal{Y}_t] - \alpha_u(t) = (1 - \kappa_u) E_u[c_t | \mathcal{Y}_t] \geq 0.$$

Conversely, suppose there exists an integrable process $\{\alpha_u(t)\}$ satisfying (i) and (ii). Let $Z_u(t)$ be defined by

(4.5) $$Z_u(t) = E_u[a_u(1) | \mathcal{Y}_t] - a_u(t)$$

where $a_u(t) = \int_0^t \alpha_u(s)ds$. Recall that $\psi_u(t)$ is the potential generated by $\int_0^t E_u[c_s | \mathcal{Y}_s]ds$. Thus

$$\psi_u(t) - Z_u(t) = E_u[\int_0^1 E_u(c_s^{(u)} | \mathcal{Y}_s)ds - \int_0^1 \alpha_u(s)ds | \mathcal{Y}_t]$$

$$- \int_0^t E_u[c_s^{(u)} | \mathcal{Y}_s]ds + \int_0^t \alpha_u(s)ds$$

$$= E_u[\int_t^1 \{E_u(c_s^{(u)} | \mathcal{Y}_s) - \alpha_u(s)\}ds | \mathcal{Y}_t]$$

$$\geq 0 \quad \text{a.s.} \quad \text{from (ii)}$$

It follows that

(4.6)(a) $\quad \psi_u \geq Z_u \quad$ a.e. $(d\lambda \times dP)$.

Similar steps using the equality in (ii) lead to

(4.6)(b) $\quad \psi_{u*} = Z_{u*} \quad$ a.e. $(d\lambda \times dP)$.

Thus

(4.7)(a) $\quad E_u\psi_u(0) \geq E_u Z_u(0)$

(b) $\quad E_{u*}\psi_{u*}(0) = E_{u*}Z_{u*}(0)$.

But $\psi_u(0) = J(u)$ and $E_u Z_u(0) = E_{u*}Z_{u*}(0) = J^*$ from(i). So (4.7) says

$$J(u) \geq J^* = J(u^*)$$

for all $u \in \mathcal{U}$. This completes the proof.

Following Rishel [13], a control $u \in \mathcal{U}$ is called <u>value decreasing</u> if

$$W_u(t) \geq E_u[W_u(t+h) | \mathcal{Y}_t] \quad \text{a.s. for each } t,$$

i.e. if $(W_u(t), \mathcal{Y}_t, P_u)$ is a supermartingale. Any optimal control is value decreasing: from Theorem 4.1, $W_u(t) = \psi_u(t)$ if u is optimal, giving equality in (3.5) and hence that

$$W_u(t) - E_u[W_u(t+h) | \mathcal{Y}_t] = E_u[\int_t^{t+h} c_s^{(u)}ds | \mathcal{Y}_t] \geq 0 \quad \text{a.s.}$$

On the other hand, optimal controls could conceivably be the only value decreasing ones, though normally one would expect this class to be a good deal larger.

In the case of value decreasing controls the value function can be represented as an Ito process and the conditions for optimality restated in a more intuitively appealing way.

Lemma 4.2 Let $u \in \mathcal{U}$ be value decreasing. Then there exist processes $\{\Lambda W_u\}, \{\nabla W_u\}^\dagger$ taking values in R, R^n respectively and adapted to $\mathcal{Y}_t$, such that

(i) $$\int_0^1 |\Lambda W_u|^2 dt < \infty \quad \text{a.s.}$$

(ii) $$E\int_0^1 |\nabla W_u| ds < \infty$$

(ii) $$W_u(t) = J^* + \int_0^t \Lambda W_u(s) ds + \int_0^t \nabla W_u(s) dy_s \quad \text{a.s.}$$

under measure P.

Proof: By assumption $(W_u(t), \mathcal{Y}, P_u)$ is a supermartingale, so that from (3.5)

(4.10) $$\left|E_u\left(W_u(t+h) - W_u(t)\right)\right| \leq E_u\left[\int_t^{t+h} c_s^{(u)} ds\right] \leq kh .$$

Thus the function $t \to E_u W_u(t)$ is right-continuous, and therefore (VI T4) (such a reference is to Meyer's book [4]) $\{W_u(t)\}$ admits a right-continuous modification, which is assumed to be the version

† So called because they play a similar role to the functions $\Lambda\phi$ and $\phi_x = \nabla\phi$ in the Markov case (see §1(C)). This will become apparent.

chosen. It is clear from the definition that $W_u(t) \to 0$ as $t\uparrow 1$, a.s. and in $L_1(P_u)$ so that $\{W_u(t)\}$ is a potential. From (VII T29) there exists a unique integrable natural increasing process $\{A_t\}$ which generates $W_u(t)$; i.e. such that

$$(4.11) \qquad W_u(t) = E_u[A_1 | \mathcal{Y}_t] - A_t .$$

Define, for $h>0$,

$$\beta_t^h = \frac{1}{h}\left(W_u(t) - E_u[W_u(t+h) | \mathcal{Y}_t]\right)$$

Then (VII T29) also states that

$$(4.12) \qquad \int_0^t \beta_s^h \, ds \to A_t$$

weakly in $L_1(P_u)$ as $h\downarrow 0$ for each fixed t. Now from (3.5)

$$\beta_t^h \leq \frac{1}{h} E_u\left[\int_t^{t+h} c_s^{(u)} ds \,|\, \mathcal{Y}_t\right] \leq k \quad \text{a.s.}$$

Thus the subset $\mathcal{H} = \{\beta_t^h : h>0\}$ is uniformly integrable and hence, from (II T23), weakly compact in $L_1(P_u)$. There therefore exists a sequence $h_n\downarrow 0$ and an element α_t of L_1 such that

$$\beta_t^{h_n} \xrightarrow{w} \alpha_t \qquad \text{as } n \to \infty$$

It is then immediate that there is a sequence $h_n\downarrow 0$ and a subset $\{\alpha_t : t \in S\} \subset L_1$, where S is a countable dense subset of [0,1], such that

$$\beta_t^{h_n} \xrightarrow{w} \alpha_t \qquad \text{as } n\to\infty \quad \text{, for each } t \in S.$$

For $t \notin S$ define α_t by

$$(4.13) \qquad \alpha_t = \underset{\substack{s\downarrow t \\ s \in S}}{\text{w-lim}} \; \alpha_s$$

To see that this limit exists, note that β_s^h is right-continuous in s for each fixed h. Let $\theta \in L_\infty$. For $t,t' \in S$, $t'>t$,

$$(4.14)\quad |E_u\theta(\alpha_t - \alpha_{t'})| \leq |E_u\theta(\alpha_t - \beta_t^{h_n})| + |E_u\theta(\alpha_{t'} - \beta_{t'}^{h_n})| + |E_u\theta(\beta_{t'}^{h_n} - \beta_t^{h_n})|$$

Now $\theta(\beta_{t'}^{h_n} - \beta_t^{h_n}) \to 0$ a.s. as $t'\downarrow t$ and hence also in L_1, in view of the uniform integrability. Choosing n such that the sum of the first two terms in (4.14) is $< \frac{1}{2}\varepsilon$ and then t' such that $|E_u\theta(\beta_{t'}^{h_n} - \beta_t^{h_n})| < \frac{1}{2}\varepsilon$ gives

$$|E_u\theta(\alpha_{t'} - \alpha_t)| < \varepsilon \quad .$$

Thus if $t_n\downarrow t, \{\alpha_{t_n}\}$ is a weak Cauchy sequence and the limit in (4.13) exists.

For $\theta \in L_\infty$,

$$(4.15)\quad |E_u\theta(\int_0^t \alpha_s ds - A_t)| \leq |E_u\theta(\int_0^t \alpha_s ds - \int_0^t \beta_s^h ds)| + |E_u\theta(\int_0^t \beta_s^h ds - A_t)| \quad .$$

The last term converges to zero along $\{h_n\}$ from (4.12), and since the expectations $E_u\beta_s^h$ are uniformly bounded for $h>0$, by Lebesgue's bounded convergence theorem

$$\int_0^t E_u\theta(\alpha_s ds - \beta_s^{h_n})ds \to 0, \quad n \to \infty \quad .$$

Thus from (4.15),

$$E_u\theta(\int_0^t \alpha_s ds - A_t) = 0 \qquad \theta \in L_\infty, \ t\varepsilon \ [0,1].$$

It follows that

$$(4.17) \qquad A_t = \int_0^t \alpha_s ds \qquad \text{a.s. for each } t.$$

Recalling (4.12) and in view of (4.17), evidently

$$\alpha_t = \underset{n \to \infty}{\text{w-lim}} \; \beta_t^{h_n}$$

for <u>every</u> subsequence $\{h_n\}$ such that the limit exists. Therefore

$$(4.18) \qquad \alpha_t = \underset{h \downarrow 0}{\text{w-lim}} \; \beta_t^h \; .$$

Now (4.11) says

$$(4.19) \qquad W_u(t) = E_u[A_1 | \mathcal{Y}_t] - \int_0^t \alpha_s ds \; .$$

$Y_t = E_u[A_1 | \mathcal{Y}_t]$ is a right-continuous, hence separable, uniformly integrable martingale on $(C, \mathcal{F}, P_u)$. Applying Theorem 2.3 with $\gamma = g^{(u)}$, $P^* = P_u$, shows that $\{Y_t\}$ has the representation

$$(4.20) \qquad Y_t = Y_o + \int_0^t \psi_s d\nu_s$$

where $d\nu_t = T_t\left(dy_t - \hat{g}_2^{(u)} dt\right)$ is a Wiener process under P_u. Here

$$(4.21) \qquad \hat{g}_2^{(u)} = E_u[g_2(t,z,u_t) | \mathcal{Y}_t] \; .$$

Thus from (4.19) and (4.20),

$$W_u(t) = Y_o - \int_0^t \left(\alpha_s + \psi_s T_s \hat{g}_2^{(u)}(s)\right) ds + \int_0^t \psi_s T_s \, dy_s \; .$$

Now $W_u(0) = J^* = Y_o$; and defining

$$\Lambda W_u(t) = - \alpha_t - \psi_t T_t \hat{g}_2^{(u)}(t)$$

and $$\nabla W_u(t) = \psi_t T_t$$

finally gives

$$W_u(t) = J^* + \int_0^t \Lambda W_u(s)ds + \int_0^t \nabla W_u(s)dy_s$$

as required.

<u>Theorem 4.3</u> $u^* \varepsilon \mathcal{U}$ is optimal if and only if there exists a constant J^* and for each value decreasing control $u \varepsilon \mathcal{U}$ processes $\{\eta_t^{(u)}\},\{\xi_t^{(u)}\}$, taking values in R, R^m respectively and adapted to $\mathcal{Y}_t$, and satisfying the following conditions:

(i) $$\int_0^1 |\xi_t^{(u)}|^2 dt < \infty \quad \text{a.s.,} \quad E\int_0^1 \xi_t^{(u)} dy_t = 0.$$

(ii) $$\chi^{(u)}(1) = 0 \quad \text{a.s., where}$$

(4.22) $$\chi^{(u)}(t) = J^* + \int_0^t \eta_s^{(u)} ds + \int_0^t \xi_s^{(u)} dy_s$$

(iii) $$\eta_t^{(u)} + \xi_t^{(u)} \hat{g}_2^{(u)}(t) + \hat{c}_t^{(u)} \geq 0 = \eta_t^{(u^*)} + \xi_t^{(u^*)} \hat{g}_2^{(u^*)}(t) + \hat{c}_t^{(u^*)}$$

for almost all (t,z), for each $u \varepsilon \mathcal{U}$.

Then $\chi_t^{(u^*)} = W_{u^*}(t)$ a.s. and $J^* = J(u^*)$, the minimal cost.

Here $\hat{g}_2^{(u)}(t)$ is defined by (4.21) above, and $\hat{c}_t^{(u)}$ is defined similarly.

<u>Proof:</u> Suppose $u \varepsilon \mathcal{U}$ is value decreasing. Then from (3.5),

(4.23) $$W_u(t) - E_u[W_u(t+h)|\mathcal{Y}_t] \leq E_u[\int_t^{t+h} c_s^{(u)} ds | \mathcal{Y}_t] .$$

Now from Lemma 4.2,

$$W_u(t) = J^* + \int_0^t \Lambda W_u(s)\, ds + \int_0^t \nabla W_u(s)\, dy_s .$$

Under measure P_u , $\{y_t\}$ has, from Lemma 2.1, the innovations process representation

$$dy_t = T_t^{-1} d\nu_t + \hat{g}_2^{(u)}(t)\, dt$$

and thus

$$W_u(t) = J^* + \int_0^t \left(\Lambda W_u(s) + \nabla W_u(s)\hat{g}_2^{(u)}(s)\right)ds + \int_0^t \nabla W_u(s) T_s^{-1}\, d\nu_s .$$

Therefore,

$$W_u(t) - E_u[W_u(t+h) \mid \mathcal{Y}_t] = - E_u\left[\int_t^{t+h} \left(\Lambda W_u(s) + \nabla W_u(s)\hat{g}_2^{(u)}(s)\right) ds \mid \mathcal{Y}_t\right],$$

so (4.23) becomes

$$(4.24) \qquad E_u\left[\int_t^{t+h} \left(\Lambda W_u(s) + \nabla W_u(s)\hat{g}_2^{(u)}(s) + c_s^{(u)}\right)ds \mid \mathcal{Y}_t\right] \geq 0 \quad \text{a.s.}$$

Denote the integrand in (4.24) by X_s and take $\theta \in L_\infty$. Then

$$\frac{1}{h} E_u[\theta\, E_u\{\int_t^{t+h} X_s ds \mid \mathcal{Y}_t\}] = \frac{1}{h} \int_t^{t+h} E_u\{E_u[\theta \mid \mathcal{Y}_t] X_s\} ds$$

$$\rightarrow E_u\{E_u[\theta \mid \mathcal{Y}_t] X_t\} = E_u\{\theta E_u[X_t \mid \mathcal{Y}_t]\}$$

as $h \downarrow 0$ for almost all t. Hence from (4.24),

$$(4.25) \qquad \Lambda W_u(t) + \nabla W_u(t)\hat{g}_2^{(u)}(t) + \hat{c}_t^{(u)} \geq 0$$

for almost all (t,z). If u is optimal then equality holds in (4.23) and hence in (4.25). Thus, identifying

$$\eta_t^{(u)} = \Lambda W_u(t)$$

$$\xi_t^{(u)} = \nabla W_u(t)$$

$$\chi_t^{(u)} = W_u(t) ,$$

properties (i) - (iii) are seen to hold.

Conversely, suppose J^*, $\{\eta_t^{(u)}\}$, $\{\xi_t^{(u)}\}$ exist and satisfy (i) -(iii), for each value decreasing control. Let $u \in \mathcal{U}$ be value decreasing. Then under measure P_u, $\chi_t^{(u)}$ satisfies

$$(4.26) \quad \chi_t^{(u)} = J^* + \int_0^t \left(\eta_s^{(u)} + \xi_s^{(u)}\hat{g}_2^{(u)}(s)\right) ds + \int_0^t \xi_s^{(u)} T_s^{-1} d\nu_s$$

where $\{d\nu_t\}$ is a Brownian motion. Define

$$\alpha_u(t) = -\eta_t^{(u)} - \xi_t^{(u)}\hat{g}_2^{(u)}(t) .$$

Then from (4.26) and (ii),

$$E_u \int_0^1 \alpha_u(s)\, ds = J^*.$$

$\{\alpha_u(t)\}$ is adapted to $\mathcal{Y}_t$, and from (iii),

$$(4.27) \qquad E_u[c_t^{(u)} | \mathcal{Y}_t] - \alpha_u(t) \geq 0 .$$

In the case $u = u^*$, (4.27) holds with equality. It now follows from Theorem 4.2 that u^* is optimal in the class of value decreasing controls. Since these are, as remarked earlier, the only candidates for the optimum, u^* must be optimal in $\mathcal{U}$.

Since u^* is optimal, $W_{u^*}(t) = \psi_{u^*}(t)$ from Theorem 4.1. Now

$$\psi_{u^*}(t) = E_{u^*}[\int_t^1 c_s^{(u^*)} ds | \mathcal{Y}_t]$$

$$= E_{u^*}[\int_t^1 \left(-\eta_s^{(u^*)} - \xi_s^{(u^*)}\hat{g}_2^{(u^*)}\right) ds - \int_t^1 \xi_s^{(u^*)} d\nu_s | \mathcal{Y}_t]$$

from (iii),

$$= E_{u^*}[-\int_t^1 \eta_s^{(u^*)} ds - \int_t^1 \xi_s^{(u^*)} dy_s | \mathcal{Y}_t]$$

$$= E_{u^*}[\int_0^t \eta_s^{(u^*)} ds + \int_0^t \xi_s^{(u^*)} dy_s | \mathcal{Y}_t] + J^* \qquad \text{from (ii)}$$

$$= \int_0^t \eta_s^{(u^*)} ds + \int_0^t \xi_s^{(u^*)} dy_s + J^*$$

$$= \chi_t^{(u^*)} .$$

Thus $\chi_t^{(u^*)} = W_{u^*}(t)$, as claimed.

5. COMPLETELY OBSERVABLE SYSTEMS.

This section treats the case where the entire past of z is available for control; i.e. (in the definitions of Section 2) $m = n$ and $\mathcal{Y}_t = \mathcal{F}_t$ for each t. Thus the admissible controls (denoted by $\mathcal{N}$) are functionals of the past of z, and are for that reason sometimes referred to as "non-anticipative controls" [9].

The considerable simplification that results in this case is due to the fact that there is now only one value function. In fact

$$\psi_{uv}(t) = E_{uv}[\int_t^1 c_s^{(v)} ds \mid \mathcal{F}_t]$$

$$= \frac{E[\rho_0^t(u)\rho_t^1(v)\int_t^1 c_s^{(v)} ds \mid \mathcal{F}_t]}{\rho_0^t(u)}$$

$$= E[\rho_t^1(v)\int_t^1 c_s^{(v)} ds \mid \mathcal{F}_t] \tag{5.1}$$

does not depend on u; thus $W_u(t) = W(t)$ for all u, where

$$W(t) = \bigwedge_{v \in \mathcal{N}_t^1} E[\rho_t^1(v)\int_t^1 c_s^{(v)} ds \mid \mathcal{F}_t] .$$

The principle of optimality (3.5) becomes

$$W(t) \leq E_u[\int_t^{t+h} c_s^{(u)} ds \mid \mathcal{F}_t] + E_u[W(t+h) \mid \mathcal{F}_t] . \tag{5.2}$$

Using this, a genuine Hamilton-Jacobi-type criterion (Theorem 5.1) for optimality can be obtained. The method is as follows: first one shows (Lemma 5.1) that there is a measure P^* such that $(W_t, \mathcal{F}_t, P^*)$

is a supermartingale. Then an Ito process representation for $W(t)$ and conditions for optimality are obtained as in the previous section.

Recall the definitions of the sets $\mathcal{G}$ and $\mathcal{D}$ from section 2.

<u>Lemma 5.1</u> There exists a process $h \in \mathcal{G}$ such that $(W_t, \mathcal{F}_t, P^*)$ is a supermartingale, where

$$\frac{dP^*}{dP} = \exp[\zeta_0^1(h)] .$$

<u>Proof:</u> Select a sequence $\{u_n\} \subset \mathcal{N}$ such that

$$J(u_n) = \psi_{u_n}(0) + W(0) = J^* .$$

Now $g^{(u_n)} \in \mathcal{G}$ and hence $\rho_0^1(u_n) \in \mathcal{D}$ for each n. From Theorem 2.2 there exists a subsequence, also denoted by $\{\rho(u_n)\}$, and an element $h \in \mathcal{G}$ such that

(5.3) $$\rho_0^1(u_n) \to \rho^* \quad \text{weakly in } L_1(P)$$

where $$\rho^* = \exp[\zeta_o^1(h)] .$$

Evidently, from (5.3), for any $t \in [0,1]$,

(5.4) $$\rho_0^t(u_n) = E[\rho_0^1(u_n) \mid \mathcal{F}_t] \to E[\rho^* \mid \mathcal{F}_t] = \exp[\zeta_0^t(h)] .$$

Define the measure P^* by $dP^* = \rho^* dP$ and let

$$\rho^{*t}_0 = E[\rho^* \mid \mathcal{F}_t] .$$

To show that $(W_t, \mathcal{F}_t, P^*)$ is a supermartingale it suffices to prove that for any t, h, $F \in \mathcal{F}_t$,

$$(5.5) \qquad \int_F (W_{t+h} - W_t)dP^* = \int_F \rho^{*}{}_0^{t+h}(W_{t+h} - W_t)dP \leq 0 .$$

Let $\rho_* = \rho^{*}{}_0^{t+h}$ and $\rho_n = \rho_0^{t+h}(u_n)$. Then

$$(5.6) \qquad \int_F \rho_*(W_{t+h} - W_t) = \int_F (\rho_* - \rho_n)(W_{t+h} - W_t) + \int_F \rho_n(\psi_{u_n}(t) - W_t)$$

$$+ \int_F \rho_n(W_{t+h} - \psi_{u_n}[t+h]) + \int_F \rho_n(\psi_{u_n}(t+h) - \psi_{u_n}(t)).$$

The third and fourth terms of (5.6) are non-positive, the third because $\psi_{u_n}(t+h)$ majorizes W_{t+h} and the fourth because ψ_{u_n} is a supermartingale under P_{u_n}.

Fix $\varepsilon > 0$ and choose n' such that $\psi_{u_n}(0) < W(0) + \varepsilon$ for $n \geq n'$. From (5.2.) (with t=0, h=t),

$$E_{u_n}[\psi_{u_n}(t) - W_t] < \varepsilon$$

for each t. Hence

$$(5.7) \qquad \int_F \rho_n[\psi_{u_n}(t) - W_t] \leq \int_C \rho_n[\psi_{u_n}(t) - W_t]$$

$$= E_{u_n}[\psi_{u_n}(t) - W_t] \leq \varepsilon \quad \text{for } n \geq n'.$$

Now $[W_{t+h} - W_t]I_F \in L_\infty$, so there exists n" such that for $n \geq n''$,

$$\int_F (\rho_* - \rho_n)(W_{t+h} - W_t) < \varepsilon .$$

Thus for $n \geq \max[n', n'']$, in (5.6),

$$\int_F \rho_*(W_{t+h} - W_t) < \varepsilon$$

which is equivalent to (5.5) since ε was arbitrary. This completes the proof.

Lemma 5.2 There exist processes $\{\Lambda W_t\}$, $\{\nabla W_t\}$ taking values in R, R^n, respectively, and adapted to $\mathcal{F}_t$, such that

(i) $$\int_0^1 |\nabla W|^2 ds < \infty \quad \text{a.s.}$$

(ii) $$E\int_0^1 |\Lambda W| ds < \infty$$

(iii) $$W_t = J^* + \int_0^t \Lambda W_s ds + \int_0^t \nabla W_s dz_s$$

almost surely under measure P.

Proof: Choose a sequence $\{u_n\} \subset \mathcal{U}$ satisfying (5.3) and such that

$$J(u_n) \to W(0) \text{ as } n \to \infty.$$

Now

(5.8) $$|E^*(W_{t+h} - W_t)| = |E[\rho^*(W_{t+h} - W_t)]|$$
$$\leq |E[(\rho^* - \rho(u_n))(W_{t+h} - W_t)]|$$
$$+ |E[\rho(u_n)(W_{t+h} - W)]|.$$

The first term on the right goes to zero as $n \to \infty$ since $(W_{t+h} - W_t) \in L_\infty$ and since $\rho(u_n) \to \rho^*$ weakly in L_1 by (5.3). Also

$$E[\rho(u_n)(W_{t+h} - W_t)] = E[\rho(u_n)(W_{t+h} - \psi_{u_n}(t+h)]$$
$$+ E[\rho(u_n)(\psi_{u_n}(t) - W_t)] + E[\rho(u_n)(\psi_{u_n}(t+h) - \psi_{u_n}(t))],$$

and by (5.7) the first two terms on the right go to zero as $n \to \infty$. Finally from (5.1) it is easy to check that

$$E[\rho(u_n)(\psi_{u_n}(t) - \psi_{u_n}(t+h))] = E[\rho_t^{t+h}(u_n) \int_t^{t+h} c_s ds] \leq kh.$$

Thus letting $n \to \infty$ in (5.8) we get

$$(5.9) \qquad |E^*(W_{t+h} - W_t)| \leq kh.$$

This implies, as in Lemma 4.2, the existence of a right-continuous modification of W_t; and since $(W_t, \mathcal{F}_t, P^*)$ is a potential, that

$$W_t = E^*[A_1 | \mathcal{F}_t] - A_t$$

where $$A_t = \underset{h \downarrow 0}{\text{w-lim}} \int_0^t \beta_s^h ds$$

and $$\beta_t^h = 1/h \left(W_t - E^*[W_{t+h} | \mathcal{F}_t]\right).$$

The next stage is to show that $\alpha_t = \frac{d}{dt} A_t = \underset{h \downarrow 0}{\text{w-lim}}\, \beta_t^h$. It suffices to show that $\mathcal{H} = \{\beta_t^h : h>0\}$ is uniformly integrable; then the rest of the proof is exactly as in the proof of Lemma 4.2. From (IIT19) of [4], $\mathcal{H}$ is uniformly integrable if

(i) $E^*\beta_t^h$ are uniformly bounded for $h>0$, and

(ii) $\int_F |\beta_t^h| dP^* \to 0$ as $P^*F \to 0$, uniformly in h.

(i) follows from (5.9). Since β_t^h is $\mathcal{F}_t$-measurable, in proving (ii) we can restrict ourselves to $F \in \mathcal{F}_t$. Now

$$(5.10) \qquad \int_F h\beta_t^h \, dP^* = \int_F [W_t - W_{t+h}]dP^*$$

$$= \int_F [W_t - W_{t+h}](\rho^* - \rho(u_n))dP$$

$$+ \int_F [W_t - W_{t+h}]\rho(u_n)dP$$

Once again since $(W_t - W_{t+h}) \in L_\infty$ and $\rho(u_n) \overset{w}{\to} \rho^*$, the first term on the right goes to zero as $n \to \infty$. Next,

$$(5.11) \qquad \int_F [W_t - W_{t+h}]\rho(u_n)dP = \int_F \rho(u_n)(W_t - \psi_{u_n}(t))dP$$

$$+ \int_F \rho(u_n)(\psi_{u_n}(t+h) - W_{t+h})dP$$

$$+ \int_F \rho(u_n)(\psi_{u_n}(t) - \psi_{u_n}(t+h))dP.$$

From (5.7), the first two terms on the right go to zero as $n \to \infty$. On the other hand from (5.1)

$$\psi_{u_n}(t) = E[\rho_t^{t+h}(u_n) \int_t^{t+h} c_s ds | \mathcal{F}_t]$$

$$+ E[\rho_t^{t+h}(u_n)\psi_{u_n}(t+h) | \mathcal{F}_t],$$

so that

$$\int_F \rho(u_n)\psi_{u_n}(t)dP = \int_F \rho_0^t(u_n)\psi_{u_n}(t)dP$$

$$= \int_F \rho_0^{t+h}(u_n)[\int_t^{t+h} c_s ds]dP$$

$$+ \int_F \rho_0^{t+h}(u_n)\psi_{u_n}(t+h)dP.$$

Also

$$\int_F \rho(u_n)\psi_{u_n}(t+h)dP = \int_F \rho_0^{t+h}(u_n)\psi_{u_n}(t+h)dP$$

so that the last term in (5.11) is equal to

$$\int_F \rho_0^{t+h}(u_n)[\int_t^{t+h} c_s ds]dP \leq kh \int_F \rho_0^{t+h}(u_n)dP$$

and converges to khP^*F as $n \to \infty$. Thus letting $n \to \infty$ in (5.10) we conclude that

$$\int_F h\beta_t^h \, dP^* \leq khP^*F.$$

and (ii) is established. Therefore

$$W_t = E^*[A_1 | \mathcal{F}_t] - \int_0^t \alpha_s ds. \tag{5.12}$$

To represent the separable martingale $E^*[A_1 | \mathcal{F}_t]$, again Theorem 2.3 is used. Recall from Lemma 5.1 that

$$dP^* = \exp[\zeta_0^1(h)]dP .$$

Thus

(5.13) $$dw = \sigma^{-1}(dz - h_t dt)$$

is a Brownian motion under P^* and is in fact the innovations process for $\{z_t\}$ since it is adapted to $\mathcal{F}_t$. From Theorem 2.3 there exists a process $\{\phi_t\}$ such that

(5.14) $$E^*[A_1 | \mathcal{F}_t] = E^*[A_1] + \int_0^t \phi_s dw_s$$

Combining (5.12) - (5.14) gives

$$W_t = J^* + \int_0^t \Lambda W_s ds + \int_0^t \nabla W_s dz_s .$$

where $$\Lambda W_t = -\alpha_t - \phi_t \sigma_t^{-1} h_t$$

$$\nabla W_t = \phi_t \sigma_t^{-1} .$$

This is the desired result.

<u>Theorem 5.1</u> (Non-anticipative controls).

$u^* \in \mathcal{U}$ is optimal if and only if there exist a constant J^* and processes $\{\eta_t\}, \{\xi_t\}$ taking values in R, R^n respectively, adapted to $\mathcal{F}_t$, and satisfying the following conditions:

(i) $\int_0^1 |\xi_t|^2 dt < \infty$ a.s., $E\int_0^1 \xi_t dz_t = 0$

(ii) $\chi(1) = 0$ a.s., where

(5.15) $$\chi(t) = J^* + \int_0^t \eta_s ds + \int_0^t \xi_s dz_s$$

(iii) $\eta_t + \xi_t g_t^{(u)} + c_t^{(u)} \geq 0 = \eta_t + \xi_t g_t^{(u^*)} + c_t^{(u^*)}$

for almost all (t,z), for each $u \in \mathcal{N}$.

Then $\chi(t) = W_t$ a.s. and $J^* = J(u^*)$, the minimal cost.

Proof: Let $u \in \mathcal{N}$. Then from Lemma 5.2 and Girsanov's theorem,

$$W_t = J^* + \int_0^t \left(\Lambda W_s + \nabla W_s g_s^{(u)}\right) ds + \int_0^t \nabla W_s \sigma_s dw_s$$

where $\{w_t\}$ is a Brownian motion under P_u. From the principle of optimality (5.2),

$$(5.16) \quad E_u[W_t - W_{t+h} \mid \mathcal{F}_t] = -E_u\left[\int_t^{t+h} \left(\Lambda W_s + \nabla W_s g_s^{(u)}\right) ds \mid \mathcal{F}_t\right]$$

$$\leq E_u\left[\int_t^{t+h} c_s^{(u)} ds \mid \mathcal{F}_t\right]$$

i.e.,

$$(5.17) \quad E_u\left[\int_t^{t+h} \left(\Lambda W_s + \nabla W_s g_s^{(u)} + c_s^{(u)}\right) ds \mid \mathcal{F}_t\right] \geq 0 \quad \text{a.s.}$$

Denote the integrand in (5.17) by X_s and pick $\theta \in L_\infty$.

$$\frac{1}{h} E_u\left(\theta\, E_u\left[\int_t^{t+h} X_s ds \mid \mathcal{F}_t\right]\right) = \frac{1}{h} \int_t^{t+h} E_u\left(E_u[\theta \mid \mathcal{F}_t] X_s\right) ds$$

$$\to E_u\left(E_u[\theta \mid \mathcal{F}_t] X_t\right) = E_u \theta X_t$$

as $h \downarrow 0$ for almost all t. It follows that

$$(5.18) \qquad X_t = \Lambda W_t + \nabla W_t g_t^{(u)} + c_t^{(u)} \geq 0 \qquad \text{a.e. } (d\lambda \times dP).$$

If $u = u^*$, optimal, then equality holds in (5.16) and hence in (5.18). Thus (iii) is satisfied with

$$\eta_t = \Lambda W_t , \quad \xi_t = \nabla W_t , \quad \chi_t = W_t .$$

(i) and (ii) are easily seen to hold also. Conversely, suppose J^*, $\{\eta_t\},\{\xi_t\}$ exist and satisfy (i) - (iii). Take $u \in \mathcal{U}$. Then from (5.15),

(5.19) $$\chi_t = J^* + \int_0^t \left(\eta_s + \xi_s g_s^{(u)}\right) ds + \int_0^t \xi_s \sigma_s dw_s$$

where $(W_t, \mathcal{F}_t, P_u)$ is a Brownian motion. Define

$$\alpha_u(t) = -\eta_t - \xi_t g_t^{(u)}.$$

Then from (5.19) and (ii),

$$E_u \int_0^1 \alpha_u(s) ds = J^* ,$$

and from (iii),

$$c_t^{(u^*)} - \alpha_{u^*}(t) = 0 \leq c_t^{(u)} - \alpha_u(t) \qquad \text{a.e.}(d\lambda \times dP).$$

It now follows from Theorem 4.2 that u^* is optimal. From Theorem 4.1, $W_t = \psi_{u^*}(t)$ and so

$$\psi_{u^*}(t) = E^*\left[\int_t^1 c_s^{(u^*)} ds \,\middle|\, \mathcal{F}_t\right]$$

$$= E^*\left[\int_t^1 \left(-\eta_s - \xi_s g_s^{(u^*)}\right) ds - \int_t^1 \xi_s \sigma_s dw^* \,\middle|\, \mathcal{F}_t\right] \qquad \text{from(iii)}$$

$$= E^*\left[-\int_t^1 \eta_s ds - \int_t^1 \xi_s dz_s \,\middle|\, \mathcal{F}_t\right]$$

$$= E^*\left[\int_0^t \eta_s ds + \int_0^t \xi_s dz_s \,\middle|\, \mathcal{F}_t\right] + J^* \qquad \text{from (ii)}$$

$$= \chi(t).$$

Thus $\chi(t) = W_t$, as stated.

6. MARKOV CONTROLS.

In this section a more resticted class of models is considered, namely those where the system matrices g and σ depend at a given time on the state only at that time. More precisely, let $\mathcal{B}_t$ be the σ-field generated by the single random variable z_t. The definitions (2.3) and (2.5) are unchanged except for (2.3.)(ii) and (2.5.)(ii) which now read:

(6.1) (ii)′ For each fixed (t,u), $g(t,\cdot,u)$ and $\sigma(t,\cdot)$ are $\mathcal{B}_t$-measurable.

In view of [3, §35.1a] this amounts to saying that g and σ are functions on $[0,1]\times R^n\times R^\ell$ taking on values $g(t,z_t,u)$ and $\sigma(t,z_t)$ at (t,z,u).

The class of Markov controls is denoted by $\mathcal{M}=\mathcal{M}_0^1$, where $\mathcal{M}_s^t$ is the class of functions u satisfying the following conditions:

(6.2)
(i) $u:[0,1]\times R^n \to \Xi \subset R^\ell$ is jointly measurable.
(ii) $E[\rho_s^t(u)|\mathcal{F}_s] = 1$ a.s.

where $\rho_s^t(u)$ is defined by (2.6) with

$$g_t^{(u)} = g(t,z_t,u[t,z_t]) .$$

Let $u\in\mathcal{M}$. Then, from Theorem 2.1, under measure P_u the process $\{z_t\}$ satisfies

$$z_t = z_s + \int_s^t g_\tau^{(u)}d\tau + \int_s^t \sigma_\tau dw_\tau \tag{6.3}$$

where $(w_t, \mathcal{F}_t, P_u)$ is a Browniam motion. From (6.3) it is evident that

$$E_u[z_t|\mathcal{F}_s] = E_u[z_t|\mathcal{B}_s] \qquad \text{a.s.}$$

so that z_t is a Markov process under P_u; hence the term "Markov controls". $\{z_t\}$ is also Markov under the original measure P.

The cost rate function c is also assumed to satisfy a condition similar to (6.1), so that

$$c_t^{(u)}(z) = c(t,z_t,u[t,z_t]) \; .$$

Stopping the process at the first exit time τ from a cylinder Q (as in §1(C)) can be accommodated within this framework. For let $I(s,x) = 1$ for $(s,x) \in Q$ and $=0$ elsewhere. Then new system functions $g^o = Ig$, $\sigma^o = I\sigma$ and $c^o = Ic$ satisfy all the relevant conditions. If $u \in \mathcal{M}$ and E_u^o denotes integration with respect to the measure corresponding to $g^{o(u)}$, σ^o, then

$$E_u[\int_0^\tau c_s^{(u)} ds] = E_u^o[\int_0^1 c_s^{o(u)} ds] \; .$$

The remaining cost function $\psi_u(t)$ is defined as

$$\begin{aligned}\psi_u(t) &= E_u[\int_t^1 c_s^{(u)} ds \,|\, \mathcal{B}_t] \\ &= E_u[\int_t^1 c_s^{(u)} ds \,|\, \mathcal{F}_t] \\ &= E[\rho_t^1(u) \int_t^1 c_s^{(u)} ds \,|\, \mathcal{F}_t]\end{aligned}$$

This does not depend on u_s for $s \in [0,t]$; there is therefore, as in the case of complete observations, a single value function $U(t,z_t)$ defined by

$$U_t = U(t,z_t) = \bigwedge_{u\varepsilon\mathcal{M}_t^1} \psi_u(t) \ .$$

Since $\mathcal{M}_t^1 \subset \mathcal{N}_t^1$ it is clear that $U_t \geq W_t$ a.s. for each t. The main result of this section (Theorem 6.2) is that in fact $U_t = W_t$. This is intuitively clear: since the system's evolution from time t depends only on z_t the controller gains nothing by taking account of previous values z_s, s<t. The proof depends on a priciple of optimality for the Markov case and results exactly analogous to Lemma 3.5 and Theorem 5.1 for the completely observable case. The proofs are almost identical here, the Markov property stepping in whereever the fact $\mathcal{F}_s \subset \mathcal{F}_t$ for s<t was used in section 5. So in the following, complete details are provided only where there is significant deviation from the corresponding previous proofs.

<u>Lemma 6.1</u> (Markov Principle of Optimality.)

Let $u\varepsilon\mathcal{M}$. Then for each t,h,

$$U_t \leq E_u[\int_t^{t+h} c_s^{(u)}ds|\mathcal{B}_t] + E_u[U_{t+h}|\mathcal{B}_t] \qquad \text{a.s.}$$

<u>Lemma 6.2</u> There exist measurable functions ΛU: $[0,1]\times R^n \to R$ and U_x: $[0,1]\times R^n \to R^n$ such that

(i) $$E\int_0^1 |\Lambda U(t,z_t)|dt < \infty$$

(ii) $$\int_0^1 |U_x(t,z_t)|^2 \, dt < \infty \qquad \text{a.s.}$$

(iii) $$U(t,z_t) = J_M + \int_0^t \Lambda U(s,z_s)ds + \int_0^t U_x(s,z_s)dz_s$$

where $J_M = \inf_{u\varepsilon\mathcal{M}} J(u)$, the minimum Markov cost.

Proof: The methods of Lemma 5.2 can be used to show that U_t has the representation

$$(6.5) \qquad U_t = J_M + \int_0^t \eta_s ds + \int_0^t \xi_s dz_s$$

where $\{\eta_\tau\}$ $\{\xi_\tau\}$ are adapted to $\mathcal{F}_t$. It remains to show that η_t, ξ_t are $\mathcal{B}_t$-measurable for each t. For n=1,2... let

$$\tau_n = \min \left(1, \inf\{t: \int_0^t |\xi_s|^2 ds \geq n\}\right)$$

τ_n is a stopping time of $\mathcal{F}_t$ and $\tau_n \uparrow \infty$ a.s. since

$$\int_0^1 |\xi_s|^2 ds < \infty \qquad \text{a.s.}$$

Let
$$\xi_t^{(n)} = \xi_t \qquad \text{for } t \leq \tau_n$$
$$= 0 \qquad \text{for } \tau_n < t \leq 1$$

Let:

$$(6.6) \qquad M_t = \int_0^t \xi_s dz_s$$

$$M_t^{(n)} = M_{t \wedge \tau_n} = \int_0^t \xi_n^{(n)} dz_s$$

Now $E\int_0^1 |\xi_s^{(n)}|^2 ds \leq n$, so that $M_t^{(n)}$ is a second-order (square integrable) martingale for each n; thus M_t is by definition a local second-order martingale. The following results are proved in Kunita and Watanabe[12]. Let

$T = \{(a_1(t) - a_2(t)) : a_i(t \wedge \tau_n)$ is a natural, integrable increasing process adapted to $\mathcal{F}_t$, i=1,2; n=i,2...$\}$

If $(X_t, \mathcal{F}_t), (Y_t, \mathcal{F}_t)$ are local second order martingales there exists a unique process $\langle Y,X\rangle_t \; \varepsilon T$ such that for t>s

$$E[(X_{t\wedge\tau_n} - X_{s\wedge\tau_n})(Y_{t\wedge\tau_n} - Y_{s\wedge\tau_n})|\mathcal{F}_s] = E[\langle Y,X\rangle_{t\wedge\tau_n} - \langle Y,X\rangle_{s\wedge\tau_n}|\mathcal{F}_s]$$

a.s.

In addition,

(6.7) $$\langle Y,X\rangle_t = 1/4(\langle X+Y\rangle_t - \langle X-Y\rangle_t)$$

where $\langle X\rangle_t = \langle X,X\rangle_t$.

$\langle X\rangle_t$ is known as the <u>quadratic variation</u> of X for the following reason: if X has continuous sample paths then [12, Thm.1.3] there exists a sequence of partitions $\{t_k^{(n)}, k=1,2..k_n\}$ of $[0,t]$ such that

(6.8) $$\max_k |t_k^{(n)} - t_{k-1}^{(n)}| \to 0 \qquad n\to\infty$$

(6.9) $$\sum_k (X_{t_k^{(n)}} - X_{t_{k-1}^{(n)}})^2 \to \langle X\rangle_t - \langle X\rangle_0 \quad \text{a.s. as } n\to\infty$$

It is shown in [15] that for local martingales of the form (6.6),

(6.10) $$\langle M\rangle_t = \int_0^t |\xi_s|^2\, ds \qquad \text{a.s.}$$

Also, referring to (6.5) and (6.9),

(6.11) $$\sum_k (U_{t_k^{(n)}} - U_{t_{k-1}^{(n)}})^2 \to \langle M\rangle_t - \langle M\rangle_0 \quad \text{a.s. as } n\to\infty.$$

(The sums corresponding to $\int \eta_s ds$ converge to zero a.s. since this term is of bounded variation.)

Let superscript i denote the i'th component of a vector. Then from (6.6),

$$M_t + z_t^i = \int_0^t \sum_{j\neq i} \xi_s^j dz_s^j + \int_0^t (\xi_s^i + 1) dz_s^i ,$$

so that, using (6.10),

$$\langle M+z^i\rangle_t = \int_0^t \left(|\xi|^2 + 2\xi^i + 1\right)ds$$

$$\langle M-z^i\rangle_t = \int_0^t \left(|\xi|^2 - 2\xi^i + 1\right)ds .$$

Therefore

(6.12) $$\langle M,z^i\rangle_t = \int_0^t \xi_s^i ds ,$$

i.e.,

$$\xi_t^i = \frac{d}{dt}\langle M,z^i\rangle_t .$$

In view of (6.9) and (6.11), for each $h>0$ there is a sequence of partitions $\{t_k^{(n)}\}$ of $[t,t+h]$ satisfying (6.8) and

(6.13) $$\sum_k \left(Y_{t_k^{(n)}} - Y_{t_{k-1}^{(n)}}\right)^2 \to \langle X\rangle_{t+h} - \langle X\rangle_t \quad \text{a.s., } n\to\infty$$

where in this case $X_t = M_t + z_t^i$ or $M_t - z_t^i$ and $Y_t = U_t + z_t^i$ or $U_t - z_t^i$. In either case, for any n the sum on the left of (6.13) is an $\mathcal{F}_t^{t+h}$-measurable random variable, where

$$\mathcal{F}_t^{t+h} = \sigma\{z_s, s\epsilon[t,t+h]\}$$

It follows from (6.12) that

$$\xi_t^{(h)} = \frac{1}{h}\int_t^{t+h} \xi_s^i ds$$

is $\mathcal{F}_t^{t+h}$-measurable. Now $\xi_t^{(h)} \to \xi_t^i$ w-L_1 for almost all t. Hence a subsequence of a sequence of convex combinations converges a.s. and therefore ξ_t^i is $\mathcal{F}_t^{t+h}$-measurable for every h and hence measurable with respect to

$$\bigcap_{h>0} \mathcal{F}_t^{t+h} = \mathcal{B}_t .$$

There is thus a measurable function $U_x:[0,1]\times R^n \to R^n$ such that

(6.14) $$U_x(t,z_t) = \xi_t .$$

Referring back to (6.5) now gives

$$\int_t^{t+h} \eta_s ds = U_{t+h} - U_t - \int_t^{t+h} U_x(s,z_s)\, ds .$$

Thus $\frac{1}{h}\int_t^{t+h} \eta_s ds$ is $\mathcal{F}_t^{t+h}$-measurable, so η_t must be $\mathcal{B}_t$-measurable by the same reasoning as above. Defining

$$\Lambda U(t,z_t) = \eta_t$$

concludes the proof of the lemma.

Corollary. Suppose the value function $U(t,x)$ has continuous first, and continuous first and second, partial derivatives respectively in t and in x; then

(6.15) $$U_x(t,x) = \frac{\partial}{\partial x} U(t,x)$$

(6.16) $$\Lambda U(t,x) = \frac{\partial U(t,x)}{\partial t} + 1/2 \sum_{i,j} \frac{\partial^2 U(t,x)}{\partial x_i \partial x_j} (\sigma\sigma')_{ij}$$

Proof: Denote the right hand sides of (6.15) and (6.16) by U_x', $\Lambda U'$ respectively. Under measure P

$$dz_t = \sigma(t,z_t)dB_t ,$$

so applying Ito's lemma to the function $U(t,z_t)$ gives

$$dU_t = U_x'(t)dz_t + \Lambda U'(t)dt .$$

Thus $\int_0^t (U_x - U_x')dz = \int_0^t (\Lambda U' - \Lambda U)dt$ and the left hand member is a local martingale which must be of bounded variation. It follows that $\int_0^t (U_x - U_x')dz = 0$ a.s. for each t, and hence that $\Lambda U_t' = \Lambda U_t$, $U_x'(t) = U_x(t)$ a.s.

Remark: The corollary shows that the results of this section are precisely equivalent to those of Fleming mentioned in the Introduction, when the relevant conditions are satisfied.

Theorem 6.1 (Markov Controls.) $u^* \varepsilon \mathcal{M}$ is optimal if and only if there exists a constant J_M and measurable functions $\eta:[0,1]\times R^n \to R$ and $\xi:[0,1]\times R^n \to R^n$ satisfying:

(i) $$\int_0^1 |\xi(t,z_t)|^2 dt < \infty \quad \text{a.s.,} \quad E\int_0^1 \xi(t,z_t)dz_t = 0$$

(ii) $$\chi(1) = 0 \quad \text{a.s. , where}$$

$$\chi(t) = J_M + \int_0^t \eta(s,z_s)ds + \int_0^t \xi(s,z_s)dz_s$$

(iii) $$\eta(t,z_t) + \xi(t,z_t)g(t,z_t,u[t,z_t]) + c(t,z_t,u[t,z_t]) \geq 0 \quad \text{a.s.}$$

$$\eta(t,z_t) + \xi(t,z_t)g(t,z_t,u^*[t,z_t]) + c(t,z_t,u^*[t,z_t]) = 0 \quad \text{a.s.}$$

Then $\chi(t) = U_t$ a.s. and $J_M = J(u^*)$, the cost of the optimal Markov policy.

Proof: As for Theorem 5.1, using Lemma 6.2.

Notice that since $u(t,z_t)$ can take any value in Ξ, and the restriction of Wiener measure to $\mathcal{B}_t$ is absolutely continuous with respect to Lebesgue measure, (iii) is equivalent to:

(6.17) $$\eta(t,x) + \min_{v\varepsilon\Xi}\{\xi(t,x)g(t,x,v) + c(t,x,v)\} = 0$$

for all $(t,x) \varepsilon [0,1]\times R^n$, and the optimal policy u^* is characterized by the property that $[U_x(t,x)\ g(t,x,v) + c(t,x,v)]$ is minimized by $v = u^*(t,x)$.

Theorem 6.2 For the system considered in this section (i.e. satisfying (6.1)),

$$\inf_{u\in\mathcal{M}} J(u) = \inf_{u\in\mathcal{N}} J(u) ,$$

where $\mathcal{N}$ is the class of non-anticipative controls.

Proof: From Theorem 6.1 and (6.17),

$$(6.18) \quad \Lambda U(t,x) + U_x(t,x)g(t,x,v) + c(t,x,v) \geq 0$$

for all $(t,x,v) \in [0,1]\times R^n \times \Xi$.

Let $u \in \mathcal{N}$. Then the process $\{w_t\}$ defined by

$$dw_t = \sigma_t^{-1}\left(-g^{(u)}dt + dz_t\right)$$

is a Brownian motion under P_u and

$$U_t = J_M + \int_0^t \left(\Lambda U_s + U_x g^{(u)}\right)ds + \int_0^t U_x \sigma dw .$$

Now $U(1) = 0$ a.s., so taking expectations at $t=1$ gives:

$$J_M = E_u\int_0^1 \left(-\Lambda U - U_x g^{(u)}\right)ds$$

$$\leq E_u \int_0^1 c^{(u)}ds \qquad \text{from (6.18),}$$

$$= J(u).$$

Since u was arbitrary,,

$$J_M \leq \inf_{u\in\mathcal{N}} J(u) .$$

The reverse inequality is immediate from the inclusion $\mathcal{M} \subset \mathcal{N}$.

7. A NOTE ON TWO-PERSON, ZERO-SUM STOCHASTIC DIFFERENTIAL GAMES.

Stochastic differential games - control problems where there are several controllers with conflicting objectives - can also be treated by methods based at least implicitly on dynamic programming. For instance Friedman in [22] has developed a theory using partial differential equations analogous to that of Fleming [9] for the optimal control problem. The "Girsanov" method of this paper can also be applied. The intention here is not to provide an exhaustive account but merely to indicate one or two of the possibilities; in particular, attention is restricted to two-person zero-sum games where complete information is available to both players. The method was first applied to games of this type by Varaiya [8],[21]. See Theorem 7.1 below.

The game (G) is defined as follows. The system dynamics are represented by

$$dz_t = g(t,z,u,v)dt + \sigma(t,z)dw_t$$

where g and σ satisfy (2.5) with the obvious modifications. The control <u>strategies</u> u and v take values in $\Xi_1 \subset R^{\ell_1}$ and $\Xi_2 \subset R^{\ell_2}$ respectively and satisfy (2.6) with $\mathcal{Y}_t = \mathcal{F}_t$ (complete observations) The measure P_{uv} is defined for any admissible strategy (u,v) by

$$\frac{dP_{uv}}{dP} = \rho_0^1(uv) = \exp[\zeta_0^1(g^{(uv)})]$$

where

$$g^{(uv)}(t,z) = g(t,z,u[t,z],v[t,z]).$$

The <u>payoff</u> is

$$J(u,v) = E_{uv}[\int_0^1 c_s^{(uv)} ds] .$$

Here E_{uv} denotes expectation with respect to P_{uv} and $c_s^{(uv)} = c(s,z,u[s,z],v[s,z])$ is a bounded function satisfying similar conditions to those satisfied previously by the cost function. Player I (control u) is attempting to minimize the payoff while player II (control v) wants to maximize it. The game has a <u>saddle point</u> if there is a pair of strategies (the <u>equilibrium strategies</u>) (u^*,v^*) such that for all admissible (u,v),

$$J(u^*,v) \leq J(u^*,v^*) \leq J(u,v^*) .$$

<u>Assumption</u>: There exist equilibrium strategies (u^*,v^*) for the game (G).

In [8],[21] it is shown that a saddle point does in fact exist under certain conditions. To be precise,

<u>Theorem 7.1</u> Suppose

(i) $\sigma = I$ (the identity matrix)

(ii) g has the form $g(t,z,u,v) = \begin{bmatrix} g_1(t,z,u) \\ g_2(t,z,v) \end{bmatrix}$

(iii) For fixed (t,z), $g_1(t,z,\cdot)$ and $g_2(t,z,\cdot)$ are continuous on Ξ_1, Ξ_2 respectively.

(iv) $g_1(t,z,\Xi_1)$ and $g(t,z,\Xi_2)$ are closed and convex for each (t,z).

Then game (G) has a saddle point.

Let (u^*,v^*) be an equilibrium strategy for (G) and let $P^* = P_{u^*v^*}$, $E^* = E_{u^*v^*}$. For any admissible strategy, define the process ψ_t^{uv} by

$$\psi_t^{uv} = E_{uv}[\int_t^1 c_s^{(uv)} ds | \tilde{\mathcal{F}}_t] .$$

Let

$$\phi_t = \psi_t^{u^*v^*} .$$

<u>Lemma 7.1</u> For each $t \in [0,1]$ and $h > 0$,

$$(7.2) \quad E_{u^*v}[\int_t^{t+h} c_s^{u^*v} ds | \mathcal{F}_t] + E_{u^*v}[\phi_{t+h} | \mathcal{F}_t] \leq \phi_t$$

$$\leq E_{uv^*}[\int_t^{t+h} c_s^{uv^*} ds | \mathcal{F}_t] + E_{uv^*}[\phi_{t+h} | \mathcal{F}_t] \quad \text{a.s.}$$

<u>Proof:</u> Suppose there is a strategy v for player II such that for some t,h,

$$\phi_t < E_{u^*v}[\int_t^{t+h} c_s^{u^*v} ds | \mathcal{F}_t] + E_{u^*v}[\phi_{t+h} | \mathcal{F}_t]$$

for $z \in M \subset \mathcal{F}_t$, PM>0. Define the strategy v' for player II by

$$v' = v \qquad t \in [t,t+h],\ z \in M$$
$$= v^* \qquad \text{elsewhere.}$$

Then

$$(7.3) \quad J(u^*,v') - J(u^*,v^*) = E_{u^*v'}(I_M \int_t^{t+h} c_s^{u^*v'} ds)$$

$$+ E_{u^*v'}(I_M \int_{t+h}^1 c_s^{u^*v'} ds) - E^*(I_M \int_t^1 c_s^* ds)$$

where I_M is the indicator function of M. Now,

$$E_{u^*v'}(I_M \int_t^{t+h} c_s^{u^*v'} ds) = E^*(I_M E[\rho_t^{t+h}(u^*v) \int_t^{t+h} c_s^{u^*v} ds | \mathcal{F}_t])$$

$$= E^*\left(I_M\, E_{u^*v}\left[\int_t^{t+h} c_s^{u^*v}\,ds \,|\, \mathcal{F}_t\right]\right)$$

$$> \phi_t I_M - E^*\left(I_M\, E_{u^*v}[\phi_{t+h} \,|\, \mathcal{F}_t]\right)$$

$$= \phi_t I_M - E^*\left(E[\rho_t^{t+h}(u^*v) I_M \phi_{t+h} \,|\, \mathcal{F}_t]\right)$$

$$= \phi_t I_M - E_{u^*v'}\left(I_M \phi_{t+h}\right) \quad . \tag{7.4}$$

From (7.3) and (7.4),

$$J(u^*,v') > J(u^*,v^*).$$

So PM must be zero. The other inequality in (7.2) is proved similarly.

<u>Lemma 7.2</u> $\{\psi_t^{uv}\}$ is the potential generated by the integrable increasing process $\{a_{uv}(t)\}$, where

$$a_{uv}(t) = \int_0^t c_s^{(uv)}\,ds.$$

<u>Proof</u>: As for Lemma 4.1.

<u>Lemma 7.3</u> There exist processes $\Lambda\phi$, $\nabla\phi$ such that

$$\phi_t = J^* + \int_0^t \Lambda\phi_s\,ds + \int_0^t \nabla\phi_s\,dz_s$$

<u>Proof</u>: From Lemma 7.2, ϕ has the representation

$$\phi_t = E^*\left[\int_0^1 c_s^*\,ds \,|\, \mathcal{F}_t\right] - \int_0^t c_s^*\,ds \ .$$

Under measure P^* the innovations process of z is $dw = \sigma^{-1}(dz - g^*dt)$. Hence from Theorem 2.3 there is process $\{\gamma_t\}$ such that

$$E^*\left[\int_0^1 c_s^*\,ds \,|\, \mathcal{F}_t\right] = \int_0^t \gamma_s \sigma_s^{-1}\,(dz_s - g_s^*\,ds) \ .$$

The result follows after defining

$$\nabla\phi = \gamma\, \sigma^{-1}$$

$$\Lambda\phi = c^* - \nabla\phi g^* .$$

Theorem 7.2 (u^*,v^*) is an equilibrium strategy if and only if there exist processes $\{\eta_t\}, \{\xi_t\}$ adapted to $\mathcal{F}_t$, and a constant J^* such that

(i) $\int_0^1 |\xi_t|^2 dt < \infty$ a.s. and $E\int_0^1 \xi_t dz_t = 0.$

(ii) $\chi(1) = 0$ a.s., where

$$\chi(t) = J^* + \int_0^t \eta_s ds + \int_0^t \xi_s dz_s$$

(iii) $$\eta_t + \min_u (g^{(uv^*)}\xi + c^{(uv^*)}) = \eta_t + (g^{(u^*v^*)}\xi + c^{(u^*v^*)})$$

$$= \eta_t + \max_v (g^{(u^*v)}\xi + c^{(u^*v)}) = 0.$$

Then $\chi_t = \phi_t$ a.s. for each t, and J^* is the value of the game.

Proof: Sufficiency is proved as in the proof of Theorem 4.3. Necessity is established by showing that $\eta_t = \Lambda\phi_t$ and $\xi_t = \nabla\phi_t$ satisfy (i) -(iii). Fixing $v = v^*$ and using precisely the methods of Theorem 4.3 together with Lemma 7.1 gives the result with the left-hand side of (iii), while fixing $u = u^*$ similarly gives the right-hand side. This completes the proof.

For $p \in R^n$, define

$$\min_u \max_v H(t,x,u,v,p) = \max_v \min_u H(t,x,u,v,p) \tag{7.5}$$

for all $(t,x,p) \in [0,1]\times C\times R^n$.

The equality (iii) in Theorem 7.2 is a version of Isaacs' equation (the game equivalent of the Hamilton-Jacobi equation). The partial differential equation counterpart of this for the Markov (pure strategies) case was derived by Friedman in [22], and a solution shown to exist under certain conditions; notably, under the assumption that (7.5) is satisfied.

REFERENCES.

[1] J.L.Doob, Stochastic Processes,Wiley, New York, 1953.

[2] N.Dunford and J.T.Schwartz,Linear Operators,Part 1, Interscience, New York, 1958.

[3] M.Loève, Probability Theory, 3rd. ed.,Van Nostrand, Princeton, N.J., 1963.

[4] P.A.Meyer, Probability and Potentials, Blaisdell, Waltham Mass., 1966.

[5] A.V.Skorokhod, Studies in the Theory of Random Processes, Addison Wesley, Reading, Mass., 1965.

[6] J.M.C.Clark, The representation of functionals of Brownian motion by stochastic integrals, Ann. Math. Stat. 41, pp 1282-1295, (1970).

[7] V.E.Beneš, Existence of optimal stochastic control laws, SIAM J. Control 9 (1971).

[8[T.E.Duncan and P.P.Varaiya, On the solutions of a stochastic control system, SIAM J. Control 9 (1971).

[9] W.H.Fleming, Optimal continuous-parameter stochastic control, SIAM Review 11 (1969) pp470-509.

[10] M.Fujisaki, G. Kallianpur & H. Kunita, Stochastic differential equations for the non-linear filtering problem, to appear.

[11] I.V.Girsanov, On transforming a certain class of stochastic processes by absolutely continuous substitution of measures, Theory of Prob. & Appl. 5 (1960) pp285-301.

[12] H.Kunita & S.Watanabe, On square integrable martingales, Nagoya Math. J. 30 (1967) pp209-245.

[13] R.Rishel, Necessary and sufficient dynamic programming conditions for continuous-time stochastic optimal control, SIAM J. Control 8 (1970) pp559-571.

[14] H.S.Witsenhausen, A counterexample in stochastic optimum control, SIAM J. Control 6 (1968) pp131-147.

[15] E.Wong, Representation of martingales, quadratic variation and applications, SIAM J. Control, to appear.

[16] W.H.Fleming & M.Nisio, On the existence of optimal stochastic controls, J.Math. Mech. 15 (1966) pp777-794

[17] W.H.Fleming, Duality and a priori estimates in markovian optimization problems, J. Math. Anal. Appl. 16 (1966) pp254-279.

[18] W.M.Wonham, On the separation theorem of stochastic control, SIAM J. Control 6 (1968), pp312-326.

[19] M.H.A.Davis & P.P.Varaiya, Information states for linear stochastic systems, J. Math. Anal. Appl., to appear.

[20] H.J.Kushner, On the stochastic maximum principle - fixed time of control, J. Math. Anal. Appl. 11 (1965) pp78-92.

[21] P.P.Varaiya, Differential Games, Proc. 6th. Berkeley Symp. on Math. Stat. and Prob., to appear.

[22] A. Friedman, Stochastic Differential Games, to appear.