January, 1979 LIDS-P-874

MARTINGALE METHODS IN STOCHASTIC CONTROL

M.H.A. Davis

Abstract

The martingale treatment of stochastic control problems is based on the idea that the correct formulation of Bellman's "principle of optimality" for stochastic minimization problems is in terms of a submartingale inequality: the "value function" of dynamic programming is always a submartingale and is a martingale under a particular control strategy if and only if that strategy is optimal. Local conditions for optimality in the form of a minimum principle can be obtained by applying Meyer's submartingale decomposition along with martingale representation theorems; conditions for existence of an optimal strategy can also be stated.

This paper gives an introduction to these methods and a survey of the results that have been obtained so far, as well as an indication of some shortcomings in the theory and open problems. By way of introduction we treat systems of controlled stochastic differential equations, the case for which the most definitive results have been obtained so far. We then outline a general semimartingale formulation of controlled processes, state some optimality conditions and indicate their application to other specific cases such as that of controlled jump processes. The martingale approach to some related problems - optimal stopping, impulse control and stochastic differential games - will also be outlined.

Paper presented at the Workshop on Stochastic Control Theory and Stochastic Differential Systems, University of Bonn, January, 1979. Proceedings to be published in the Springer-Verlag Lecture Notes in Control and Systems Sciences Series, edited by M. Kohlmann.

MARTINGALE METHODS IN STOCHASTIC CONTROL*

M.H.A. Davis
Laboratory for Information and Decision Systems
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

LIDS-P-874

CONTENTS

1. INTRODUCTION

The status of continuous-time stochastic control theory ten years ago is admirably summarized in Fleming's 1969 survey paper [40]. The main results, of which a very brief outline will be found in §2 below and a complete account in the book [41], concern control of completely-observable diffusion processes, i.e. solutions of stochastic differential equations. Formal application of Bellman's "dynamic programming" idea quickly leads to the "Bellman equation" (2.3), a quasi-linear parabolic equation whose solution, if it exists, is easily shown to be the value function for the control problem. At this point the probabilistic aspects of the problem are finished and all the remaining work goes into finding conditions under which the Bellman equation has a solution. The reason why dynamic programming is a fruitful approach in stochastic control is precisely that these conditions are so much weaker than those required in the deterministic case. As regards problems with *partial observation* the best result was Wonham's formulation of the "separation theorem" [78] which he proved by reformulating the problem as one of complete observations, with the "state" being the conditional mean estimate produced by the Kalman filter; see §6 below.

* Work supported by the U.S. Air Force Office of Sponsored Research under Grant AFOSR 77-3281 and by the Department of Energy under Contract EX-76-A-01-2295.

The dynamic programming approach, while successful in many applications, suffers from many limitations. An immediate one is that the controls have to be smooth functions of the state in order that the resulting stochastic differential equation (2.1) have a solution in the Ito sense. This rules out, for example, "bang-bang" controls which arise naturally in some applications (e.g. [3]). Thus a weaker formulation of the solution concept seems essential for stochastic control; this was provided by Stroock and Varadhan [71] for Markov processes and by various forms of measure transformations, beginning with the Girsanov Theorem [43], for more general stochastic systems; these are outlined in §3. But even with the availability of weak solution concepts it seems that the Bellman equation approach is essentially limited to Markovian systems and that no general formulation of problems with partial observations is possible (A Bellman equation for partially observed diffusions was formally derived by Mortensen [65], but just looking at it convinces one that some other approach must be tried).

Since 1969 a variety of different approaches to stochastic control have been investigated, among them the following (a very partial list). Krylov [51] has studied generalized solutions of the Bellman equation; methods based on potential theory [5] and on convex analysis [7] have been introduced by Bismut; necessary conditions for optimality using general extremal theory have been obtained [44] by Haussmann; a reformulation of dynamic programming in terms of nonlinear semigroups has been given by Nisio [66]; variational inequality techniques have been introduced by Bensoussan and Lions [4], and computational methods systematically developed by Kushner [54].

This survey outlines the so-called "martingale approach" to stochastic control. It is based on the idea of formulating Bellman's "principle of optimality" as a *submartingale inequality* and then using Meyer's submartingale decomposition [63] to obtain local conditions for optimality. This is probably the most general form of dynamic programming and applies to a very general class of controlled processes, as outlined in §5 below. However, more specific results can be obtained when more structure is introduced, and for this reason we treat in some detail in §§4,6 the case of stochastic differential equations, for which the best results so far are available. Other specific cases are outlined in §7.

I have attempted to compile, in §9, a fairly complete list of references on this topic and related subjects. Undoubtedly this list will suffer from important omissions, but readers have my assurance that none of these is intentional. It should also be mentioned that no systematic coverage of martingale representation theorems has been attempted, although they are obviously germane to the subject.

2. CONTROL OF DIFFUSION PROCESSES

To introduce the connection between dynamic programming and submartingales, let us consider a control problem where the n-dimensional *state process* x_t satisfies the Ito stochastic differential equation

$$(2.1)\qquad dx_t = f(t, x_t, u_t)dt + \sigma(t, x_t)dw_t$$

$$x_0 = \xi \in R^n$$

Here w_t is an n-dimensional Brownian motion and the components of f and σ are C^1 functions of x, u, with bounded derivatives. The *control* u_t is a feedback of the current state, i.e. $u_t = u(t, x_t)$ for some given function $u(t, x)$ taking values in the *control set* U. If u is Lipschitz in x, then (2.1) is a stochastic differential equation satisfying the standard Ito conditions and hence has a unique strong solution x_t. The cost associated with u is then

$$J(u) = E\left[\int_0^T c(t, x_t, u_t)dt + \Phi(x_T)\right]$$

where T is a fixed terminal time and c, Φ are, say, bounded measurable functions. The objective is to choose the function $u(\cdot,\cdot)$ so as to minimize J(u). An extensive treatment of this kind of problem will be found in Fleming and Rishel's book [41].

Introduce the *value function*

$$(2.2)\qquad V(t, x) = \inf_u E_{(t, x)}\left[\int_t^1 c(s, x_s, u_s)ds + \Phi(x_T)\right]$$

Here the subscript (t, x) indicates that the process x_s starts at $x_t = x$, and the infimum is over all control functions restricted to the interval [t, T]. Formal application of Bellman's "principle of optimality" together with the differential formula suggests that V should satisfy the *Bellman equation*:

$$(2.3)\qquad V_t + 1/2 \sum_{i,j} (\sigma\sigma')_{ij} V_{x_i x_j} + \min_{u\in U} [V'_x f(t, x, u) + c(t,x,u)] = 0$$

$$(t, x) \in [0, T[\times R^n$$

$$(2.4)\qquad V(T, x) = \Phi(x),\ x\in R^n$$

($V_t = \partial V/\partial t$ etc., and V_t, V_x etc. are evaluated at (t, x) in (2.3)). There is a "verification theorem" [41 ,§ VI 4] which states that if V is a solution of (2.3), (2.4) and u^o is an admissible control with the property that

$$V'_x(t,x)\ f(t,x,u^\circ(t,x)) + c(t,x,u^\circ(t,x)) = \min_{u\in U} [V'_x(t,x)\ f(t,x,u) + c(t,x,u)]$$

then u° is optimal. Conditions under which a solution of (2.3), (2.4) is guaranteed will be found in [41 ,§ VI 6]. Notable among them is the *uniform ellipticity* condition: there exists $\kappa>0$ such that

$$(2.5)\qquad \sum_{ij} (\sigma\sigma')_{ij}\ \xi_i \xi_j \geq \kappa|\xi|^2$$

for all $\xi\in R^n$. This essentially says that noise enters every component of equation (2.1), whatever the coordinate system.

Let us reformulate these results in martingale terms, supposing the conditions are such that (2.3), (2.4) has a solution with suitable growth properties (see below). For any admissible control function u and corresponding trajectory x_t define a process M_t^u as follows:

(2.6) $$M_t^u = \int_0^t c(s, x_s, u_s)ds + V(t, x_t)$$

Note that M_t^u is the minimum expected *total* cost given the evolution of the process up to time t. Expanding the function $V(t, x_t)$ by the Ito rule gives

(2.7) $$M_t^u = V(0,\xi) + \int_0^t [V_t + 1/2 \sum_{ij} (\sigma\sigma')_{ij} V x_i x_j + V_x' f^u + c]ds + \int_0^t V_x \sigma dw$$

where $f^u(t,x) = f(t, x, u(t, x))$. But note from (2.3) that the integrand in the second term of (2.7) is always non-negative. Thus this term is an *increasing process*. If u is optimal then the integrand is identically zero. Assuming that the function V is such that the last term is a martingale, we thus have the following result:

(2.8) *For any admissible* u, M_t^u *is a submartingale and* u *is optimal if and only if* M_t^u *is a martingale.*

The intuitive meaning of the submartingale inequality is clear: the difference

$$E[M_t^u | x_r, r \leq s] - M_s^u$$

is simply the expected cost occasioned by persisting in using the non-optimal control over the time interval [s, t] rather than switching to an optimal control at time s. The other noteworthy feature of this formulation is that an optimal control is constructed by minimizing the *Hamiltonian*

$$H(t,x,V_x,u) = V_x' f(t,x,u) + c(t,x,u)$$

and, conveniently, the "adjoint variable" V_x is precisely the function that appears in the integrand of the stochastic integral term in (2.7).

Abstracted from the above problem, the "martingale approach" to stochastic control of systems with complete observations (i.e. where the controller has exact knowledge of the past evolution of the controlled process) consists of the following steps:

1. Define the value function V_t and conditional minimal cost processes M_t^u as in (2.2), (2.6)
2. Show that the "principle of optimality" holds in the form (2.8)
3. Construct as optimal policy by minimizing a Hamiltonian, where the adjoint variable is obtained from the integrand in a stochastic integral representation of the martingale component in the decomposition of the submartingale M_t^u.

In evaluating the cost corresponding to a control policy u in the above problem, all that is required is the sample space measure induced by the x_t process with control u. It is also convenient to note that the cost can always be regarded as a terminal cost by introducing an extra state variable x_t° defined by

(2.9) $$dx_t^\circ = c(t, x_t, u_t)dt + dw_t^\circ$$

where w_t° is an additional Brownian motion, independent of w_t. Then since $E\, w_T^\circ = 0$ we have

(2.10) $$J(u) = E\,[x_T^\circ + \Phi(x_T)] = E\,[\tilde{\Phi}(x_T^\circ, x_T)]$$

Let C denote the space of R^{n+1}- valued continuous functions on [0, T] and (F_t) the increasing family of σ-fields generated by the coordinate functions $\{\chi_t\}$ in C. Since (2.1), (2.9) define a process (x_t°, x_t) with a.s. continuous sample functions, this induces a measure, say μ_u, on (C, F_T) and the cost can be expressed as

$$J(u) = \int_C \Phi(\chi_T^\circ, \chi_T)\ \mu_u(d\chi)$$

It turns out that each μ_u is absolutely continuous with respect to the measure μ induced by (x_t°, x_t) with $f \equiv c \equiv 0$. Thus in its abstract form the control problem has the following ingredients:

(i) A probability space (Ω, F_T, μ)

(ii) A family of measures $(\mu_u, u \in \mathcal{U})$ absolutely continuous with respect to (or, equivalently, a family of positive random variables (ℓ_u) such that $E\,\ell_u = 1$ for each $u \in \mathcal{U}$)

(iii) An F_Y-measurable random variable $\tilde{\Phi}$

The problem is then to choose $u \in U$ so as to minimize $E_u \Phi = E[\ell_u \tilde{\Phi}]$. In many cases it is possible to specify the Radon-Nikodym derivative ℓ_u directly in order to achieve the appropriate sample-space measure. We outline this idea in the next section before returning to control problems in section 4.

3. ABSOLUTELY CONTINUOUS TRANSFORMATION OF MEASURES

Let (Ω, F, P) be a probability space and $(F_t)_{0 \le t \le 1}$ be an increasing family of sub-σ-fields of F such that

(3.1)

(i) Each F_t is completed with all null sets of F

(ii) (F_t) is right-continuous: $F_t = \bigcap_{s>t} F_s$

(iii) F_o is the completion of the trivial σ-field $\{\emptyset, \Omega\}$.

(iv) $F_1 = F$

Suppose P_u is a probability measure such that $P_u \ll P$. Define

(3.2) $$L_1 = dP_u/dP$$

and

(3.3) $$L_t = E\,[L_1 | F_t]$$

Then L_t is a positive martingale, $EL_t = 1$, and $L_o = 1$ a.s. in view of (3.1) (iii). According to [63 , VI T4] there is a modification of (L_t) whose paths are right-continuous with left hand limits (we denote $L_{t-} = \lim_{s\uparrow t} L_s$). Define

$$T = 1 \wedge \inf \{t: L_t \wedge L_{t-} = 0\}$$

$$T_n = 1 \wedge \inf\{t: L_t < 1/n \}$$

Then $T_n \uparrow$, $T_n \leq T$ and Meyer shows in [64 , VI] that $L_t(w) = 0$ for all $t > T(w)$, a.s.

Suppose (X_t) is a given non-negative *local* martingale of (F_t) with $X_o=1$ a.s. Then X_t is always a *supermartingale*, since, if s_n is an increasing sequence of localizing times and $s<t$, using Fatou's lemma we have:

$$x_s = \lim_n X_{s\wedge s_n} = \lim_n E[X_{t\wedge s_n} | F_s] \geq E[\lim_n \inf X_{t\wedge s_n} | F_s] = E[X_{t\wedge s_n} | F_s]$$

It follows that $EX_t \leq 1$ for all t and X_t is a martingale if and only if $EX_1 = 1$. This is relevant below because we will want to use (3.2), (3.3) to *define* a measure P_u from a given process L_t which, however, is *a priori* only known to be a local martingale.

Let (M_t) be a local martingale of (F_t) and consider the equation

(3.4) $$L_t = 1 + \int_0^t L_{s-} dM_s$$

It was shown by Doléans-Dade [28] (see also [64 , IV 25], that there is a unique local martingale (L_t) satisfying this, and that L_t is given explicitly by

$$L_t = \exp\ (M_t - 1/2 \langle M^c, M^c\rangle_t) \prod_{s\leq t} (1 + \Delta M_s)\ e^{-\Delta M_s}$$

Here M_t^c is the "continuous part" of the local martingale M_t (see [64, IV 9] and the countable product is a.s. absolutely convergent. We denote $L_t = \mathcal{E}(M)_t$ (the "Doléans-Dade exponential").

Suppose $\Delta M_s \geq -1$ for all (s,ω). Then L_t is a non-negative local martingale, and hence according to the remarks above is a martingale if and only if $EL_1 = 1$. Its utility in connection with measure transformation lies in the following result, due to van Schuppen and Wong [69].

(3.5) *Suppose* $EL_1 = 1$ *and define a measure* P_u *on* (Ω, F_1) *by* (3.2). *Let* x *be a local martingale such that the cross-variation process* $\langle X, M\rangle$ *exists. Then* $\tilde{X}_t := X_t - \langle X, M\rangle_t$ *is a* P_u *local martingale.*

Note that from the general formula connecting Radon-Nikodym derivatives and conditional expectations we have

(3.6) $$E_u(\hat{X}_t | F_s) = \frac{E[L_t \hat{X}_t | F_s]}{L_s}$$

and consequently $\hat{X}_t$ is a P_u-local martingale if and only if $\hat{X}_t L_t$ is a P-local martingale. One readily verifies that this is so with X_t defined as above, using the general change of variables formula for semimartingales [64 , IV 21].

Conditions for the existence of $\langle X, M\rangle$ are given by Yoeurp [79]. Recall that the "square brackets" process [x, M] is defined for any pair of local martingales x, M by

$$[X, M] = \langle X^c, M^c\rangle_t + \sum_{s\leq t} \Delta X_s \Delta M_s$$

Yoeurp defines <X, M> as the dual predictable projection (in the sense of Dellacherie [27]) of [X, M], when this exists and gives conditions for this [79, Thm. 1.12]. (This definition coincides with the usual one [52] when x and M are locally square integrable.) In fact a predictable process A such that X-A is a P_u-local martingale exists *only* when these conditions are satisfied (see also [64, VI 22]).

An exhaustive study of conditions under which $E\mathcal{E}(M)_1 = 1$ is given by Lepingle and Memin in [57]. A typical condition is that $\Delta M > -1$ and

$$(3.7)\qquad E\,[\exp\,(1/2\ \langle M^c, M^c\rangle_1 \prod_{t\leq 1} (1 + \Delta M_t)\ \exp(\frac{-\Delta M_t}{1+\Delta M_t})\ ] < \infty$$

This generalizes an earlier condition for the continuous case given by Novikov [67]. We will mention more specific results for special cases below; see also references [2], [3], [12], [13], [30], [36], [43], [56], [60], [77].

Let us now specialize the case where x_t is a Brownian motion with respect to the σ-fields F_t, and M_t is a stochastic integral

$$M_t = \int_0^t \phi_s dX_s$$

where ϕ_s is an adapted process satisfying

$$(3.8)\qquad \int_0^t \phi_s^2 ds < \infty \quad \text{a.s. for each } t$$

Then $\langle M^c, M^c\rangle_t = \langle M, M\rangle_t = \int_0^t \phi_s^2\, ds$ and $\langle M, X\rangle_t = \int_0^t \phi_s ds$ so that

$$(3.9)\qquad L_t = \exp\,(\int_0^t \phi_s dX_s - 1/2\int_0^t \phi_s^2\, ds)$$

and

$$(3.10)\qquad B_t := X_t - \int_0^t \phi_s ds$$

in a P_u-local martingale (assuming $EL_1 = 1$). Since X_t has continuous paths, $\langle X, X\rangle_t$ is the sample path quadratic variation of X_t [52] and this is invariant under absolutely continuous change of measure. It follows from (3.10), since the last term is a continuous process of bounded variation, that

$$\langle B, B\rangle_t^{(P_u)} = \langle X, X\rangle_t^{(P)} = t$$

and hence that B_t *is a* P_u *-Brownian notion,* in view of the Kunita-Watanabe characterization [64, III 102]. This is the original "Girsanov theorem" [43]. A full account of it will be found in Chapter 6 of Liptser and Shiryaev's book [60]. In particular, theorem 6.1 of [60] gives Novikov's condition: $EL_1 = 1$ if ϕ satisfies (3.7) and

$$(3.11)\qquad E\,\exp(1/2\int_0^1 \phi_s^2 ds) < \infty$$

The Girsanov theorem is used to define "weak solutions" in stochastic differential equations. Suppose $f : [0, 1] \times C \to R$ is a bounded non-anticipative functional on the space of continuous functions and define

$$\phi(t, \omega) = f(t, x(\cdot,\omega))$$

where x_t is a P-Brownian motion as above. Then (3.11) certainly holds and from (3.10) we see that under measure P_u the process x_t satisfies

$$dx_t = f(t, x)dt + dB_t \tag{3.12}$$

where B_t is a P_u-Brownian motion, i.e. (x_t, F_t, P_u) is a "weak solution" of the stochastic differential equation (3.12). (It is not a "strong" or "Ito" solution since B does not necessarily generate x; a well-known example of Tsyrelson [72], [60, §4.4.8] shows that this is possible). The reader is referred to [60] for a comprehensive discussion of weak and strong solutions, etc. Suffice it to say that the main advantage of the weak solution concept for control theory is that there is no requirement that the dependence of f on x in (3.12) be smooth (e.g., Lipshitz as the standard Ito conditions require), so that such things as "bang-bang" controls [3], [21] fit naturally into this framework.

4. CONTROLLED STOCHASTIC DIFFERENTIAL EQUATIONS - COMPLETE OBSERVATIONS CASE

This problem, a generalization of that considered in §2, is the one for which the martingale approach has reached its most definitive form, and it seems worth giving a self-contained outline immediately rather than attempting to deduce the results as special cases of the general framework considered in §5. The results below were obtained in a series of papers: Rishel [68], Beneš [2], Duncan and Varaiya [30], Davis and Varaiya [25], Davis [16], and Elliott [34].

Let Ω be the space of continuous functions on [0, 1] to R^n, (w_t) the family of coordinate functions and $F_t^\circ = \sigma\{w_s, s \le t\}$. Let P be Wiener measure on (Ω, F_1°) and F_t be the completion of F°_t with null sets of F_1°. Suppose $\sigma : [0, 1] \times \Omega \to R^{n \times n}$ is a matrix-valued function such that

(4.1)
(i) $\sigma_{ij}(\cdot,\cdot)$ is F_t- predictable

(ii) $|\sigma_{ij}(t, x) - \sigma_{ij}(t, y)| \le K \sup_{0\le s\le t} | x_s - y_s|$

(iii) $\sigma(t, x)$ is non-singular for each (t, x) and $|(\sigma^{-1}(t, x))_{ij}| \le K$

(Here K is a fixed constant, independent of t, i, j). Then there exists a unique strong solution to the stochastic differential equation

$$dx_t = \sigma(t, x)dw_t, \quad x_o \in R^n \text{ given.}$$

Now let U be a compact metric space, and $f: [0, 1] \times C \times U \to R^n$ a given function which is continuous in $u\in U$ for fixed $(t, x) \in [0, 1] \times C$, an F_t-predictable process as a function of (t, x) for fixed $u\in U$, and satisfies

(4.2) $$|f(t,x,u)| \le K(1+\sup_{s\le t}|x_s|)$$

Now let U be the family of F^t-predictable U-valued processes and for $u\in U$ define

$$L_t(u) = \exp\left(\int_0^t (\sigma^{-1}(s,x)\ f(s,x,u_s))'dw_s - 1/2\int_0^t |\sigma^{-1}f|^2 ds\right)$$

The Girsanov theorem as given in §3 above generalizes easily to the vector case, and condition (4.2) implies the vector version of Novikov's condition (3.10) (see [60, p. 221]). Thus $EL_1(u) = 1$ and defining a measure P_u by

$$\frac{dP_u}{dP} = L_1(u)$$

we see that under P_u the process x_t satisfies

(4.3) $$dx_t = f(t,x,u_t)dt + \sigma(t,x)dw_t^u$$

where w_t^u is a P_u-vector Brownian motion. The cost associated with $u\in U$ is now

(4.4) $$J(u) = E_u\left[\int_0^1 c(t,x,u_t)dt + \Phi(x_1)\right]$$

where c, Φ are bounded measurable functions and c satisfies also the same condition as f.

It is clear that σ must be non-singular if weak solutions are to be defined as above (cf. the uniform ellipticity conditions (2.5)), but an important class of "degenerate" systems is catered for, namely those of the form

(4.5) $$dx_t^1 = f^1(t,x_t^1, x_t^2)dt$$

(4.6) $$dx_t^2 = f^2(t,x_t^1,x_t^2,u_t)dt + \bar{\sigma}(t,x_t^1,x_t^2)dw_t$$

where $\bar{\sigma}$ is nonsingular and f^1 is Lipschitz in x_t^1 uniformly in (t,x_t^2). Then (4.5) has a unique solution $x_t^1 = \chi_t(x^2)$ for each given trajectory x^2, and (4.6) can be rewritten as

$$dx_t^2 = f^2(t,\chi_t(x^2),x_t^2,u_t)dt + \bar{\sigma}(t,\chi_t(x^2),x_t^2)dw_t$$

which is in the form (4.3). This situation arises when a scalar n'th-order differential equation is put into 1st-order vector form.

Fix $t\in[0,1]$ and define the conditional remaining cost at time t as

$$\psi_t^u = E_u\left[\int_t^1 c^u(x,s)ds + \Phi(x_1)\,|F_t\right]$$

(Here and below we will write $c(x,s,u_s)$ as $c^u(x,s)$ or c_s^u, and similarly for f). It is seen from the formula (3.6) that ψ_t^u only depends on u restricted to the interval $[t,1]$ and since all measures P_u are equivalent the null sets up to which ψ_t^u is defined are also control-independent; in fact ψ_t^u is a well-defined element of $L_1(\Omega,F_t,P)$ for each $u\in U$. Since L_1 is a complete lattice we can define the lattice infimum

$$W_t = \bigwedge_{u \in U} \psi_t^u$$

as an F_t-measurable random variable. This is the *value function* (or *value process*). It satisfies the following *principle of optimality*, originally due to Rishel [68]: for each fixed $u \in U$ and $0 \le t \le \tau \le 1$,

$$W_t \le E_u\left[\int_t^\tau c_s^u ds \middle| F_t\right] + E_u[W_\tau | F_t] \tag{4.7}$$

The proof of this depends on the fact that the family $[\psi_t^u : u \in U]$ has the "ε-lattice property": see §5 below. Now define

$$M_t^u = \int_0^t c_s^u \, ds + W_t$$

This has the same interpretation as in (2.6) above. Note that since x_0 is assumed to be a fixed constant,

$$\begin{aligned} M_0^u &= W_0 = \inf_{v \in U} J(v) \\ M_1^u &= \int_0^1 c_s^u ds + \Phi(x_1) = \text{"sample cost"} \end{aligned} \tag{4.8}$$

The statement of the principle of optimatlity is now exactly as in (2.8). Firstly (4.7) implies that M_t^u is a P_u-submartingale for each u. Now if M_t^u is a P_u-martingale then $E_u M_0^u = E_u M_1^u$ which implies u is optimal in view of (4.8), while if u is optimal then for any t,

$$W_0 = E_u\left[\int_0^t c_s^u ds + \psi_t^u\right]$$

Now for any control we have from (4.7)

$$W_0 \le E_u\left[\int_0^t c_s^u ds + W_t\right]$$

and hence

$$E_u[W_t - \psi_t^u] \ge 0 .$$

But by definition $W_t \le \psi_t^u$ a.s.; thus $W_t = \psi_t^u$ a.s. and therefore $M_t^u = E_u(M_1^u | F_t]$. So M_t^u is a martingale if and only if u is optimal.

Fix $u \in U$. A direct argument shows that the function $t \to EM_t^u$ is right continuous, and it follows from [63, VI T4] that M_t^u has a right-continuous modification. The conditions for the Meyer decomposition [63, VII T31] are thus met, so there exists a unique predictable increasing process A_t^u with $A_0^u = 0$ and a martingale N_t^u such that

$$M_t^u = W_0 + A_t^u + N_t^u$$

We now want to represent the martingale N_t^u as a stochastic integral. If the σ-fields F_t were generated by a Brownian motion then this representation would be a standard result [15], [52], [60], but here (4.3) is only a weak solution, so (w_t^u) does not necessarily generate (F_t). Nevertheless it was proved by Fujisaki, Kallianpur and Kunita [42] (see also [25], [60]) that all F_t-martingales are in fact stochastic integrals of w_t^u, i.e. there exists an adapted process g_t such that

$$\int_0^t |g_s|^2 ds < \infty \quad \text{a.s.}$$

and

(4.9) $$N_t^u = \int_0^t g_s \sigma_s dw_s^u$$

From the definition of M_t^u we now have

(4.10) $$W_t = W_0 + \int_0^t g_s \sigma_s dw_s^u + A_t^u - \int_0^t c_s^u ds$$

Now take another control $u \in \mathcal{U}$. By definition

$$M_t^v = \int_0^t c_s^v ds + W_t$$

and hence, using (4.3) and (4.10) we get

(4.11) $$M_t^v = W_0 + \int_0^t g_s \sigma_s dw_s^v + A_t^u + \int_0^t (H_s(v_s) - H_s(u_s)) ds$$

where

(4.12) $$H_s(u_s) = g_s f(s,x,u_s) + c(s,x,u_s)$$

Now (4.11) gives a representation of M_t^v as a "special semimartingale" (= local martingale + predictable bounded variation process) under measure P_u and it is known that such a decomposition is unique [64,IV32]. But we know that M^v is a submartingale with decomposition

(4.13) $$M^v = W_0 + N_t^v + A_t^v$$

so the terms in (4.11), (4.13) must correspond. In particular this shows that *the integral* g *in* (4.9) *does not depend on the control* u. We can now state some conditions for optimality.

(4.14) A necessary condition. If $u^* \in \mathcal{U}$ is optimal then it minimizes (a.s. dP x dt) *the Hamiltonian* H_s *of* (4.12)

Indeed, if u^* is optimal then $A_t^{u^*} = 0$. Referring to (4.11) with $u = u^*$ we see that (4.14) is just the statement that the last term in (4.11) is an increasing process.

(4.15) <u>A sufficient condition for optimality</u>. *For a given control* u^*, *defined the* P^{u^*} *-martingale*

$$p_t^* = E_{u^*}[M_1^{u^*} | F_t]$$

Then u^* *is optimal if for any other* $u \in \mathcal{U}$ *the process*

$$I_t^u = p_t^* + \int_0^t (c_s^u - c_s^{u^*}) ds$$

is a P_u *-submartingale.*

This is evident since then

$$J(u^*) = I_0^u = E_u I_0^u \leq E_u I_1^u = J(u).$$

We can recast (4.15) as a local condition: since it is a martingale, p_t^* has a representation

$$p_t^* = J(u^*) + \int_0^t \tilde{g}_s \sigma_s dw_s^{u^*}$$

Now suppose that

(4.16) $\quad \tilde{H}_t(u_t) \leq \tilde{H}_t(v) \quad$ a.e. for all $v \in U$

where $\tilde{H}$ is as in (4.12) but with $\tilde{g}$ replacing g. Then a calculation similar to (4.11) shows that I_t^u is a local P_u-submartingale for any $u \in \mathcal{U}$; since $I_0^u = J(u^*)$, this implies that if T_n is a sequence of localizing times then

$$E_u[I^u_{1 \wedge T_n}] \geq J(u^*)$$

But the process I_t^u is uniformly bounded and $I^u_{1\wedge T_n} \to I_1^u$ as $n \to \infty$, so that

$$E_u[I^u_{1\wedge T_n}] \to J(u).$$

Thus (4.16) is a sufficient condition for optimality and it is easily seen that if it is satisfied then $p_t^* = M_t^{u^*}$ and $\tilde{g}_t = g_t$, a.e. See [21] for an application.

Since the process g_t is defined independently of the existence of any optimal control it seems clear from the above that an optimal control should be constructed by minimizing the Hamiltonian (4.12). Under the conditions we have stated, an implicit function lemma of Beneš [1] implies the existence of a predictable process u_t^0 such that

$$H_t(u_t^0) = \min_{v \in U} H_t(v) \text{ a.e.}$$

Using (4.11) with $u = u^0$ gives

$$M_t^v \geq W_0 + \int_0^t g_s \sigma_s dw_s^v + A_t^{u^0}$$

and hence, taking expectations at t=1,

(4.17) $\quad E_v[A_1^{u^0}] \leq J(v) - W_0$

To show u^0 is optimal it suffices, according to the criterion (2.8), to show that $A_1^{u^0} = 0$ a.s. Here we need some results on compactness of the sets of Girsanov exponentials, due to Beneš [2] and Duncan and Varaiya [30]. Let $\mathcal{A}$ be the set of R^n-valued F_t-predictable processes ϕ satisfying

$$|\phi(t,x)| \leq \kappa(1 + \sup_{s \leq t} |x_s|), \quad (t,x) \in [0,1] \times \Omega$$

(thus $f^u \in \mathcal{A}$ for $u \in \mathcal{U}$, see (4.2)) and let

$$\mathcal{D} = \{\delta(\phi) : \phi \in \mathcal{A}\}$$

where

$$\delta(\phi) = \exp\left(\int_0^1 (\sigma^{-1}\phi)'dw - 1/2 \int_0^1 |\sigma^{-1}\phi|^2 dt \right)$$

then Beneš' result is

(4.18) *D is a weakly compact subset of $L_1(\Omega,F,P)$ and $\ell>0$ a.s. for all $\ell\in D$.*

Returning to (4.17) we can, in view of (4.8), choose a sequence $u_n \in U$ such that $J(u_n) \downarrow W_0$ and hence such that for any positive integer N,

(4.19) $$E_{u_n}[A_1^{u^0} \wedge N] = E[\delta(f^{u^n})(A_1^{u^0} \wedge N)] \to 0, \; n \to \infty.$$

In view of (4.18) there is a subsequence of $\delta(f^{u^n})$ converging weakly to some $\rho\in D$; hence from (4.19)

$$E[\rho(A_1^{u^0} \wedge N)] = 0$$

and it follows that $A_1^{u^0} = 0$ a.s. We thus have:

(4.20) *Under the stated conditions, an optimal policy* u^0 *exists, constructed by minimizing the Hamiltonian* (4.12).

Two comments on this result: firstly, it is possible to recast the problem so as to have a purely terminal cost by introducing an extra state x^0 as in (2.9), (2.10). However it is important *not to do this* here, since an extra Brownian motion w_t^0 is introduced as well, and there is then no way of showing that the optimal policy u^0 does not depend on w^0 - i.e. one gets a possibly "randomized" optimal policy this way. Secondly, the existence result (4.20) was originally proved in [2] and [30] just by using the compactness properties of the density sets. However they were obliged to assume convexity of the "velocity set" $f(t,x,U)$ in order that the set $D(U) = \{\delta(f^u) : u\in U\}$ be convex (and can then be shown to be weakly closed). Finally it should be remarked that (4.20) is a much stronger result than anything available in *deterministic* control theory, the reason being of course that the noise "smooths out" the process.

A comparison of (2.3) and (4.12) shows that the process g_t plays the role of the gradient $V_x(t,x_t)$ in the Markov case, so that in a sense the submartingale decomposition theorems are providing us with a weak form of differentiation. The drawback with the martingale approach is of course that while the function V_x can (in principle) be calculated by solving the Bellman equation, the process g_t is only defined implicitly by (4.9), so that the optimality conditions (4.14) (4.15) do not provide a *constructive* procedure for calculating the optimal u^0, or for verifying whether a candidate control satisfies the necessary condition (4.14). Some progress on this has been made by Haussmann [44], but it depends on $u^0(t,x)$ being a smooth function of $x\in\Omega$, which is very restrictive.

Suppose u^0 is optimal and that the random variable

$$M_1^{u^0} = \int_0^1 c(s,x,u^0(s,x))ds + \Phi(x_1)$$

is Frechet differentiable as a function of $x\in\Omega$; then by the Riesz representation theorem there is, for each $x\in\Omega$ an R^n-valued Radon measure μ_x such that for $y\in\Omega$

$$M_1^{u^0}(x+y) = M_1^{u^0}(x) + \int_{[0,1]} y(s)\,\mu_x(ds) + o(||y||)$$

Since u^0 is optimal $M_t^{u^0}$ satisfies

$$M_t^{u^0} = J(u^0) + \int_0^t g_s \sigma_s dw_s^{u^0}$$

and Haussmann [45] [46] (see also [19]) shows that, under some additional smoothness assumptions, g_t is given by

$$g_t = E_{u^0}[\int_{]t,1]} \mu_x'(ds)\Psi(s,t)|F_t]$$

where $\Psi(s,t)$ is the (random) fundamental matrix solution of the linearized equation corresponding to (4.3) with $u = u^0$. Ths representation gives, in some cases, an "adjoint equation" satisfied by g_t, along the lines originally shown by Kushner [ ].

Finally let us remark that in all of the above the state space of x_t is R^n. Some problems - for example, control of the orientation of a rigid body - are more naturally formulated with a differentiable manifold as state space. Such problems have been treated by Duncan [29] using versions of the Girsanov theorem etc. due to Duncan and Varaiya [31].

5. GENERAL FORMULATION OF STOCHASTIC CONTROL PROBLEMS

The first abstract formulation of dynamic programming for continuous-time stochastic control problems was given by Rishel [68] who isolated the "principle of optimality" in a form similar to (4.7). The submartingale formulation was given by Striebel [70] who also introduced the important "ε-lattice property." Other papers formulating stochastic control problems in some generality are those of Boel and Varaiya [11], Memin [61], Elliott [37] [38], Boel and Kohlmann [9] [10], Davis and Kohlmann [23] and Brémaud and Pietri [14].

We shall sketch briefly a formulation, somewhat similar to that of (2.7), which is less general than that of Striebel [70] but sufficiently general to cover all of the applications considered in this paper.

The basic ingredients of the control problem are

(i) A probability space (Ω, F, P)

(ii) Two families (F_t), (Y_t) $(0\le t\le 1)$ of increasing, right-continuous, completed sub-σ-fields of F, such that $Y_t \subset F_t$ for each t.

(iii) A non-negative F_1-measurable random variable Φ.

(iv) A measurable space (U, Ξ)

(v) A family of control processes $\{\mathcal{U}_s^t,\ 0\le s\le t\le 1\}$

Each control process $u\in\mathcal{U}_s^t$ is a Y_t-predictable U-valued function on $]s,t]\times\Omega$. The family $\{U_s^t\}$ is assumed to be closed under

restriction: $u \in U_s^t \to u|_{[s,\tau]} \in U_s^\tau$ for $s \le \tau \le t$

concatenation: $u \in U_s^\tau,\ v \in U_\tau^t \Rightarrow w \in U_s^t$ where

$$w(\sigma,\omega) = \begin{cases} u(\sigma,\omega) & \sigma \in]s,\tau] \\ v(\sigma,\delta) & \sigma \in]\tau,t] \end{cases}$$

(5.1)

finite mixing: $u,v \in U_s^t,\ A \in Y_s \to w \in U_s^t$ where

$$w(\sigma,\omega) = \begin{cases} u(\sigma,\omega), & \omega \in A \\ v(\sigma,\omega), & \omega \in A^c \end{cases}$$

We denote $U = U_0^1$ (In most cases U will consist of *all* predictable U-valued processes, but (5.1) is the set of conditions actually required for the principle of optimality below). A control $u \in U_0^t$ is assumed to determine a measure P_u on (Ω, F_t) which is absolutely continuous with respect to $P|_{F_t}$ such that $P_u|_{F_o} = P|_{F_0}$ and such that the assignment is compatible in the sense that if $u \in U_0^t$, $s < t$ and $v = u|_{[0,s]}$ (so that $v \in U_0^s$) then $P_v = P_u|_{F_s}$. If $u \in U_s^t$ and X is an F_t-measurable random variable, then $E_u X$ denotes expectation with respect to measure P_u. We finally assume that $E_u \Phi < \infty$ for all $u \in U$ and the problem is then to choose $u \in U$ so as to minimize $J(u) = E_u \Phi$.

The value process corresponding to $u \in U_0^t$ is

(5.2) $$W_t^u = \bigwedge_{v,t} E_v[\Phi|Y_t]$$

where "$\bigwedge_{v,t}$" denotes the lattice infimum in $L_1(\Omega, Y_t, P)$, taken over all $v \in U$ such that $v|_{[0,t]} = u$. Note that, in contrast to the situation in §4, W_t^u is in general *not* control-independent. We nevertheless have a result analogous to (2.8), namely

(5.3) W_t^u *is a submartingale for each* $u \in U$ *and is a martingale if and only if* u *is optimal.*

Note that by inclusion and using the compatability condition, for any $\tau > t$

$$W_t^u \le \bigwedge_{v,\tau} E_v[\Phi|Y_t] = \bigwedge_{v,\tau} E_u[E_v[\Phi|Y_\tau]|Y_t]$$

so that the first statement of (5.3) is equivalent to the assertion that $\bigwedge_{v,\tau}$ and $E_u[\cdot|Y_t]$ may be interchanged, and according to Striebel [70] (see also [26] for a summary) this is possible if the random variables $E_v[\Phi|Y_t)$ have the ε-*lattice property:* if $v_1, v_2 \in U_t^1$ then there exists $v_3 \in U_t^1$ such that, with $\bar{v}_i$ denoting the concatentation of u and v_i,

(5.4) $$E_{\bar{v}_3}[\Phi|Y_t] \le E_{\bar{v}_1}[\Phi|Y_t] \wedge E_{\bar{v}_2}[\Phi|Y_t] + \varepsilon \quad \text{a.s.}$$

Now it is evident that under assumptions (5.1) the set $\{E_v[\Phi|Y_t]\}$has the 0-lattice property, because given v_1, v_2 as above one only has to define

$$A = \{w : E_{\bar{v}_1}[\Phi|Y_t] \le E_{\bar{v}_2}[\Phi|Y_t]$$

and, for $\tau \in]t,1]$,

$$v_3(\tau,\omega) = \begin{cases} v_1(\tau,\omega), & \omega \in A \\ v_2(\tau,\omega), & \omega \in A^c \end{cases}$$

Then (5.4) holds with $\Sigma=0$.

It is clear from the definition (5.2) that u is optimal if W_t^h is a P_u-martingale while conversely if u is optimal then for any $t\in[0,1]$

(5.5) $$E_u[W_0^u] = \inf_{v\in U} J(v) = J(u) = E_u[E_u[\Phi|Y_t]]\ .$$

But by the submartingale property $E_u[W_0^u] \le E_u[W_t^u]$ and this together with (5.2) and (5.5) implies that $W_t^u = E_u[\Phi|Y_t]$, i.e. W_t^u is a P_u -martingale.

Statement (5.3) is a general form of optimality principle but its connection with conventional dynamic programming is tenuous as there is a different value function for each control, reflecting the fact that past controls can affect the expectation of future performance. This is suggestive of Feldbaum's "dual control" idea, namely that an optimal controller will act so as to "acquire information" as well as to achieve direct control action.

The postulates of the general model above are not, as they stand, sufficient to endure that there is a single value function if $Y_t = F_t$ (complete information). Let

(5.6) $$L_t(u) = E[\frac{dP_u}{dP}|F_t]$$

Now fix $s\in[0,1]$ and for $s\le t\le 1$ define

$$L_t(u,v) = \begin{cases} L_t(u)/L_s(v) & \text{if } L_s(v) > 0 \\ 1 & \text{if } L_s(v) = 0 \end{cases}$$

then $L_t(u,v)$ is a positive martingale and $L_s(u,v) = 1$. Then the following hypothesis ensures that there is a process W_t such that $W_t^u = W_t$ in case $Y_t = F_t$:

(5.7) *For any* $v\in U$, *and* $u_1, u_2 \in U$ *such that* $u_1|_{]s,1]} = u_2|_{]s,1]}$ *we have*

$$L_t(u_1,v) = L_t(u_2,v) \quad \text{for all } t\in\ ]s,1]$$

See [61, Lemma 3.2]. Clearly the densities $L_t(u)$ of §4 above satisfy (5.7)

A minimum principle - complete observations case

If we are to use the principle of optimality (5.3) to obtain *local* conditions for optimality in the form of a minimum principle it is necessary to be more specific about how the densities $L_t(u)$ are related to the controls $u\in U$. This is generally through a transformation of measures as described in §3 above. A general formulation will be found in Elliott's paper [38] in this volume, but to introduce the idea let us consider the following rather special set-up.

Suppose $Y_t = F_t$ for each t, and let M_t be a given F_t-martingale with almost all paths continuous. Now take a function $\phi : [0,1]\times\Omega\times U \to R$ such that ϕ is a predictable process for each $u \in U$ and continuous in u for each fixed (t,ω), and for $u \in \mathcal{U}$ let ϕ^u denote the predictable process $\phi^u(t,\omega) = \phi(t,\omega,u(t,u))$. We suppose that for each $u \in \mathcal{U}$

(5.8) $$E \exp(1/2 \int_0^1 (\phi_s^u)^2 d\langle M\rangle_s) < \infty$$

and that the measure P_u is defined by

$$\frac{dP_u}{dP} = \mathcal{E}(\int \phi^u dM)_1$$

(see 3). From (3.7), condition (5.8) ensures that P_u is a probability measure and that $P_u \approx P$. Now $L_t(u)$ (defined by (5.6)) satisfies the equation

$$L_t(u) = \int_0^t L_s(u)\phi_s^u \, dM_s$$

The uniqueness of the solution to this equation shows that condition (5.7) is satisfied, and hence that there is a single value process W_t, which can be shown to have a right-continuous modification [61], assuming *the cost function is bounded.* Then for any $u \in \mathcal{U}$, W_t has the submartingale decomposition

(5.9) $$W_t = W_0 + N_t^u + A_t^u$$

where N_t^u is a P_u-martingale and A_t^u a predictable increasing process. According to the translation theorem, the process

(5.10) $$dM^u_t = dM_t - \phi^u d\langle M\rangle_t$$

is a continuous P_u-martingale. Decompose N_t^u into the sum

$$N_t^u = \overline{N}_t^u + \tilde{N}_t$$

where $\overline{N}_t^u$ is in the stable subspace generated by M_t^u (see [64]) and $\tilde{N}_t$ is orthogonal to this stable subspace. There is a predictable process g_t such that

$$\overline{N}_t^u = \int_0^t g_s dM_s^u$$

Now consider another admissible control v. Using (5.9), (5.10), we see, as in (4.11), (4.12) above that W_t can be written

$$W_t = W_0 + \int_0^t g_s dM_s^u + \tilde{N}_t + \int_0^t g_s(\phi_s^v - \phi_s^u) d\langle M\rangle_s + A_t^u$$

Now $\tilde{N}_t$ *is a* P_u*-martingale*, since the Radon-Nikodym derivative $E_u[dP_v/dP_u|F_t]$ is in the stable subspace generated by M^u (see [37], [38]) and hence, by the uniqueness of the semi-martingale decomposition (5.9) we have

$$A_t^u = \int_0^t g_s(\phi_s^u - \phi_s^u) d\langle M\rangle_s + A_t^u$$

Since A_t^u is an increasing process and $A_t^u = 0$ if u is optimal, we have the following minimum principle:

(5.11) *If* $u \in U$ *is optimal and* v *is any admissible control then for almost all* ω

$$g_s \phi(s,\omega,u_s) \leq g_s \phi(s,\omega,v_s) \quad \text{a.e. } (d\langle M\rangle_s)$$

In particular if U *consists of all predictable* U*-valued processes then*

$$g_s \phi(s,\omega,u_s) = \min_{v \in U} g_s \phi(s,\omega,v)$$

The importance of this type of result is that no martingale representation result is required, since the "orthogonal martingale" $\tilde{N}_t$ plays no role in the optimality conditions (things are somewhat more complicated if the basic martingale m_t is not continuous).

Partial observations case

Further progress in the case when $Y_t \neq F_t$ appears to depend on representation theorems for Y_t-martingales, although possibly a development similar to the above could be carried out. For each $u \in U$ the P_u-submartingale W_t^u is decomposed into the sum of a martingale and an increasing process. In Memin's paper it is assumed that all (Y_t,P)-martingales have a representation as a sum of stochastic integrals with respect to a continuous martingale and a random measure. It is shown in [48] that a similar representation then holds for (Y_t,P_u)-martingales since $P_u \ll P$. Using this some somewhat more specific optimality conditions can be stated, but these do not lead to useful results as no genuine minimum principle can be obtained. Rather than describe them we revert to the stochastic differential equation model of §4 for which better results have been obtained.

6. CONTROLLED STOCHASTIC DIFFERENTIAL EQUATIONS WITH PARTIAL INFORMATION

Returning to the problem of §4, let us suppose that the state vector x_t is divided into two sets of components $x_t' = (y_t', z_t')$ of which only the first is observed by the controller. Define $Y_t = \sigma\{y_s, s \leq t\}$. Then the class of admissible controls is the set N of Y_t-adapted processes with values in U. The objective is to choose $u \in N$ so as to minimize $J(u)$ given by (4.4). Following Elliott [34] we will outline a *necessary condition* for optimality. Thus we suppose that $u^* \in N$ is optimal (and write c^*, E_* instead of c^{u^*}, E_{u^*}, etc.). Let

$$\psi_t^* = E_*\left[\int_t^1 c_s^* ds + \Phi(x_1) \,\middle|\, F_t\right]$$

and for any $u \in N$ define

$$N_t^u = \int_0^t c_s^u ds + \psi_t^* .$$

Then N_t^* is an (F_t,P_*)-martingale and it is easily shown that

(6.1)
(i) $E_*[N_t^* | Y_t]$ is a (Y_t,P_*)-martingale

(ii) $E^*[N_t^u | Y_t] \leq E_*\left[E_u[N_{t+h}^u | F_t] \,|\, Y_t\right]$ for any $u \in U$ and $h > 0$

As in §4, we can represent N^*_t as a stochastic integral with respect to the Brownian motion $w^*_t = w^{u^*}_t$, i.e. there exists an F_t-adapted process g^*_t such that

(6.2) $$N^*_t = \psi^*_0 + \int_0^t g^*_s \sigma_s dw^*_s$$

Using an argument similar to that of (4.11)-(4.12) we see that N^u_t can be written

(6.3) $$N^u_t = \psi^*_0 + \int_0^t g^*_s \sigma_s dw^u_s + \int_0^t \Delta H^*_s(u)ds$$

where

$$\Delta H^*_s(u) = [g^*_s f(s,x,u_s) + c(s,x,u_s)] - [g^*_s f(s,x,u^*_s) - c(s,x,u^*_s)]$$

It now follows from (6.1) (ii) and (6.3) that

$$(1/h)E_*[E_u(\int_t^{t+h} \Delta H^*_s(u)ds|F_t)|Y_t] \geq 0$$

A rather delicate argument given in [34] shows that taking the limit as $h \downarrow 0$ gives $E_*[\Delta H^*_t(u)|Y_t] \geq 0$. We thus obtain the following minimum principle:

(6.4) *Suppose* $u^* \in N$ *is optimal and* $u \in N$. *Then there is a set* $T \subset [0,1]$ *of zero Lebesgue measure such that for* $t \notin T$

$$E_*[g^*_t f(t,x,u^*_t) + c(t,x,u^*_t)|Y_t] \leq E_*[g^*_t f(t,x,u_t) + c(t,x,u_t)|Y_t] \quad \text{a.s.}$$

where g^*_t *is the process of* (6.2).

This is a much better result than the original minimum principle (theorem 4.2 of [25]) since the optimal control minimizes the conditional expectation of a Hamiltonian involving a single "adjoint process" g^*. A similar result (including some average value state space constraints) was obtained by Haussmann [44] using the Girsanov formulation together with L.W. Neustadt's "general theory of extremals."

It is shown in [39] that a *sufficient condition* for optimality is that an inequality similar to (6.4) but with E_u replacing E_* should hold for all admissible u.

The disadvantage of the types of result outlined above is that they ignore the general cybernetic principle that in partially observable problems the conditional distribution of the state given the observations constitutes an "information state," on which control action should be based. In other words, the filtering operation is not explicitly brought in. Although there is a well-developed theory of filtering for stochastic differential equations [42], [60], it turns out to be remarkably difficult to incorporate this into the control problem. A look at the "separation theorem" of linear control [18], [78], [41], chapter 7] will show why. The separation theorem concerns a linear stochastic system of the form

(6.5) $$dx_t = Ax_t dt + \beta(u_t)dt + Gdw^{1u}_t$$
$$dy_t = Fx_t dt + R^{1/2}dw^{2u}_t$$

where w^{1u}, w^{2u} are independent vector Brownian motions, the distribution of the initial state x_0 is normal, and the coefficient matrices can be time-varying. It is assumed that GG' and R are symmetric and strictly positive definite, that the controls u_t

take values in a compact set U and that the function β is continuous. The solution of (6.5) for a given Y_t-adapted control policy u_t is then defined by standard application of the Girsanov technique and the (non-quadratic) cost is given by

$$J(u) = E_u\left[\int_0^1 c(t,x_t,u_t)dt + \Phi(x_1)\right]$$

It is shown in [24] that *the conditional distribution of* x_t *given* Y_t *is normal,* with mean $\hat{x}_t$ and covariance Σ_t given by the Kalman filter equations:

(6.6) $$d\hat{x}_t = A\hat{x}_t dt + \beta(u_t)dt + \Sigma_t F'R^{-1/2}d\nu_t$$
$$\hat{x}_0 = Ex_0$$

(6.7) $$\dot{\Sigma} = A\Sigma + \Sigma A' + GG' - \Sigma F'R^{-1}F\Sigma$$
$$\Sigma(0) = \mathrm{cov}(x_0)$$

Here ν_t is the normalized innovations process

$$\nu_t = \int_0^t R^{-1/2}(dy - F\hat{x}_s ds)$$

which is a standard vector Brownian motion. Let us denote $K(t) = \Sigma_t F'R^{-1/2}$, and let $n(\cdot,x,t)$ be the normal density function with mean x and covariance Σ_t. Now define

$$\hat{c}(t,x,u) = \int_{R^n} c(t,\xi,u)n(\xi,x,t)d\xi, \quad \hat{\Phi}(x) = \int_{R^n}\Phi(\xi)n(\xi,x,t)d\xi$$

Then the cost J(u) can be expressed as

(6.8) $$J(u) = E_u\left[\int_0^1 \hat{c}(t,\hat{x}_t,u_t)dt + \hat{\Phi}(x_1)\right]$$

The original problem is thus seen to be equivalent to a "completely observable" problem (6.6), (6.8) with "state" $\hat{x}_t$ (this characterizes the entire conditional distribution since the covariance $\Sigma(t)$ is non-random). This suggests studying "separated controls" of the form $u_t = \psi(t,\hat{x}_t)$ for some given measurable function ψ: $[0,1]\times R^n \to U$. However, such controls are, in general, *not admissible:* admissible controls are *specified functionals* of y, whereas the random variable $\hat{x}_t$ depends on past controls $\{u_s, s\leq t\}$. One way round this difficulty is to consider (6.6)-(6.8) as an independent problem of the type considered in §4, i.e., to define the solution of (6.6) by Girsanov transformation on a new probability space, for separated controls $u(t,\hat{x})$. However we then run into the fresh difficulty that weak solutions of (6.6) are only defined if the matrix $K(t)K'(t)$ is strictly positive definite, which cannot happen unless the dimension of y_t is at least as great as that of x_t - a highly artificial condition. If this condition is met then we can apply (4.17) to conclude that there exists an optimal separated control, and an extra argument as in [18] shows that its cost coincides with $\inf_{u\in\mathcal{N}} J(u)$. If $\dim(y_t) < \dim(x_t)$ then some form of approximation must be resorted to.

With these elementary obstacles standing in the way of a satisfactory martingale treatment of the separation theorem, it is not surprising that a proper formulation of information states for nonlinear problems has not yet been given. It is possible that the Girsanov solution concept is still too strong to give existence of optimal controls for partially-observable systems in any generality.

7. OTHER APPLICATIONS

This section outlines briefly some other types of optimization problems to which martingale methods have been applied. The intention is merely to indicate the martingale formulation and not to give a survey of these problems as a whole: most of them have been extensively studied from other points of view and the associated literature is enormous. Nor is it claimed that the martingale approach is, in all cases, the most fruitful.

7.1 Jump processes

A jump process is a piecewise-constant right-continuous process x_t on a probability space (Ω,F,P) with values in, say, a complete separable metric space X with Borel σ-field S. It can be identified with an increasing sequence of times $\{T_n\}$ and a sequence of X-valued random variables $\{Z_n\}$ such that

$$x_t = \begin{cases} Z_n, & t\in[T_n,T_{n+1}[\\ z_\infty, & t\geq T_\infty \end{cases}$$

where $T_\infty = \lim_n T_n$ and z_∞ is a fixed element of X. (Generally $T_\infty = \infty$ a.s. in application.) Jump processes are useful models in operations research (queueing and inventory systems) and optical communication theory, among other areas. Their structure is analysed in Jacod [47], Boel, Varaiya and Wong [12] and Davis [17]. A jump process can be thought of as an integer valued random measure μ on $E = R^+ \times X$ defined by

$$\mu(\omega,dt,dz) = \sum_n \delta_{(T_n(\omega),\, X_n(\omega))}(dt,dz)$$

where δ_e is the Dirac measure at $e\in E$. Now let

$$F_t = \sigma\{\mu(]0,s] \times A),\ s\leq t,\ A\in S\} = \sigma\{x_s,\ s\leq t\}$$

and let P be the F_t-predictable σ-field on $R^+ \times\Omega$. A random measure μ is *predictable* if the process

(7.1) $$\int_{]0,t]\times X} g(\omega,s,z)\mu(\omega,ds,dz)$$

is predictable for all bounded measurable functions g on $(\Omega\times R^+\times X,\ P_* S)$. The fundamental result of Jacod [47] is that there is a unique predictable random measure ν such that

(7.2) $$E[\int_E g(s,z)\mu(ds,dz)] = E[\int_E g(s,z)\nu(ds,dz)]$$

for all g as above. ν is also characterized by the fact that for each $A\in S$, $\nu(]0,t]\times A)$ is the dual predictable projection (in the sense of Dellacherie [27]) of $\mu(]0,t],A)$, i.e. the process

$$q(t,A) = \mu(]0,t]\times A) - \nu(]0,t]\times A)$$

is an F_t-martingale. An explicit construction for ν in terms of the distributions of the (T_n, Z_n) sequence is given in [2.3]. We will denote by $\int g\,dq$ integrals of the form $(\int g\,d\mu - \int g\,d\nu)$ where $\int g\,d\mu$ and $\int g\,d\nu$ are defined as in (7.1) then the process

$$g \cdot q_t = \int_{]0,t]} {}_X g\,dq$$

is an F_t-martingale for a suitable class of predictable integrands g, and the *martingale representation theorem* [12], [17], [47] states that all F_t-martingales are of this form for some g.

Denote

$$\Lambda_t = \nu(]0,t] \times X)$$

For each ω this is an increasing function of t and evidently the measure it defines on R^+ dominates that defined by $\nu(]0,t] \times A)$ for any $A \in S$. Thus there is a positive function $n(\omega,s,A)$ such that

$$\text{(7.3)} \qquad \nu(]0,t] \times A) = \int_{]0,t]} n(\omega,s,A)\,d\Lambda_s$$

Owing to the existence of regular conditional probabilities it is possible to choose n so that it is measurable and is a probability measure in A for each fixed (s,ω). The pair (n,Λ) is called the *local description* of the process and has the interpretation that Λ_t is the *integrated jump rate*: roughly, $d\Lambda_s \approx P[x_{s+ds} \neq x_s | F_s]$ and $n(\omega,s,\cdot)$ is the *conditional distribution* of x_s given that $x_s \neq x_{s-}$.

Optimization problems arise when the local description of the process can be controlled to meet some objective. This is normally formulated [11], [22] by absolutely continuous change of measure, as in §3: we start with a "base measure" P on (Ω, F_1) with respect to which the jump process has a local description (n,Λ) and define a new measure P_u by

$$\frac{dP_u}{dP} = \mathcal{E}(m^u)_1$$

where m^u is a (P,F_t) martingale. Under P_u the process x_t has a different local description which can be identified by the translation theorem (.). More specifically, it is supposed that the admissible controls $\mathcal{U}$ consist of F_t-predictable, (U,Ξ)-valued processes and that a real-valued measurable function ϕ on $(R^+ \times \Omega \times X \times U,\ P * S * \Xi)$ is given. Denoting $\phi^u(t,\omega,z) = \phi(t,\omega,z,u(t,\omega))$ for $u \in \mathcal{U}$, m^u is defined by

$$m^u_t(\omega) = \int_{]0,t]\times X} \phi^u(s,\omega,z)\,q(\omega,ds,dz)$$

The Doleans-Dade exponential (.) then takes the specific form

$$\mathcal{E}(m^u)_t = \exp\left(-\int_0^t \int_X \phi^u dn\, d\Lambda^c\right) \prod_{T_i \le t} \left(1+\phi^u(T_i,Z_i) - \Delta\Lambda_{T_i} \int_X \phi^u(T_i,z)\,n(T_i,dz)\right)$$

$$\times \prod_{s \le t}{}' \left(1 - \Delta\Lambda_s \int_X \phi^u(s,z)\,n(s,dz)\right)$$

where Λ^c is the continuous part of Λ and the second product is taken over the countable set of s such that $\Delta\Lambda_s > 0$ and $s \notin \{T_1, T_2, \ldots\}$. Assuming that $E\mathcal{E}(M^u)_1 = 1$, x_t is, under measure P_u, a jump process with local description

(7.4)
$$\Lambda_t^u = \int_{]0,t]\times X} ((1 + \phi_s^u - \Delta\Lambda_s \int_X \phi^u dn)\nu(ds,dz)$$

$$n^u(s,A) = \frac{\int_A (1+\phi_s^u - \Delta\Lambda_s \int_X \phi^u dn)\; n(s,dz)}{\int_X (1 + \phi_s^u - \Delta\Lambda_s \int_X \phi^u dn)\; n(s,dz)}$$

See [22], [36] for details of these calculations and conditions under which $E\mathcal{E}(m^u)_1 = 1$. Generally, only weak conditions on ϕ are needed to ensure that P_u is a probability measure on F_{T_n} for each n and hence on F_{T_∞}. If $T_\infty = \infty$ a.s. (P) then extra conditions on ϕ can be imposed to ensure that $T_\infty = \infty$ a.s.(P_u) and then P_u is a probability on F_t for each fixed t; see [77]. Let us suppose that the control problem is to choose $u \in U$ so as to minimize

$$J(u) = E_u \Phi$$

where Φ is a bounded F_1-measurable random variable. Then the problem is in the general framework of §5 and furthermore we have a martingale representation theorem analogous to that of the Brownian case. Thus local conditions for optimality can be obtained by following the steps of §4.

Suppose $u^* \in U$ is optimal. Then by the martingale representation theorem there is an integran g such that

(7.5)
$$E_*[\Phi|F_t] = J(u^*) + \int_{]0,t]\times X} g(s,z)q^*(ds,dz)$$

where $q^* = \mu - \nu^*$, and ν^* is the dual projection of μ under measure P_* (cf. (7.2)). Now let $u \in U$ be any other control; then we can rewrite (7.3) in the form

(7.6)
$$E_*[\Phi|F_t] = J(u^*) + \int_{]0,t]\times X} g\, dq^u + \int_{]0,t]\times X} g(d\nu^u - d\nu^*)$$

According to the criterion (5.3), $E_*[\Phi|F_t]$ is a P_u-submartingale, and hence the last term in (7.5) must be an increasing process. Using (7.3) and the specific forms of local description provided by (7.4), this statement translates into the following result:

(7.7) *Suppose* u^* *is optimal, let* g *be as in* (7.5) *and define*

$$h(t,z,\omega) = g(t,z,\omega) - \Delta\Lambda(t,\omega) \int_X g(t,\xi,\omega)n(t,d\xi,\omega)$$

Then for almost all ω

$$\int_X h(t,z)\phi(t,z,u_t^*)n(t,dz) = \min_{u\in U} \int_X h(t,z)\phi(t,z,u)n(t,dz) \quad \text{a.e. } (d\Lambda_t)$$

Thus, as in (4.14), the optimal control minimizes a "Hamiltonian." A sufficient condition for optimality similar to (4.15) can also be obtained. In the literature [12], [22], [77] various forms of Hamiltonian appear, depending on the nature of the cost function and the function ϕ. In [77] an existence theorem along the lines of (4.20) is obtained; however this only holds under very restrictive assumptions, related to the absolute continuity of the measures. In the Brownian case all the measures P_u are *mutually* absolutely continuous under very natural conditions, and this is crucial in the proof of the existence result, as is seen in (4.18), (4.19). In the jump process context mutual absolute continuity is very unnatural, but one is apparently obliged to insist on it if an existence result is to be obtained.

Finally, let us mention some other work related to the above. Optimality conditions for jump processes are obtained by Kohlmann [50] using Neustadt's extremal theory in a fashion analogous to Haussmann's treatment of the Brownian case [44]. Systems with both Brownian and jump process disturbances are dealt with in Boel and Kohlmann [9], [10] (based on a martingale representation theorem of Elliott [33]) and Lepeltier and Marchal [58]. The survey [13] by Bremaud and Jacod contains an extensive list of references on martingales and point processes.

7.2 Differential games [32], [35], [73], [74], [75], [76]

The set-up here is the same as that of §4 except that we suppose $U = U_1 x U_2 x \ldots x U_N$ where each U_i is a compact metric space. Then $\mathcal{U} = \mathcal{U}_1 x \ldots x \mathcal{U}_N$ where $\mathcal{U}_i$ is the set of F_t-predictable U_i-valued processes, and we assume that each $u^i \in \mathcal{U}_i$ is to be chosen by a *player* i with the objective of minimizing a *personal cost*

$$J_i(u) = J_i(u^1 \ldots u^N) = E_u[\int_0^1 c_i(s,x,u_s)ds + \Phi_i(x_1)]$$

(c_i and Φ_i satisfy the same conditions as c, Φ of §4). Thus each player is assumed to have perfect observations of the state process x_t.

Various solution concepts are available for this game [76]: $u^* = (u^{1*}, \ldots u^{N*})$ is

- *a Nash equilibrium* if there is no i and $u^i \in \mathcal{U}_i$ such that

  $J_i(u^{*1}, \ldots, u^{*(i-1)}, u^i, u^{*(i+1)}, u^{*N}) < J_i(u^*)$

- *efficient* if there is no $u \in \mathcal{U}$ such that

  $J_i(u) < J_i(u^*)$ for all i

- *in the core* if there is no subset $S \subset \{1,2\ldots,N\}$ and $u \in \mathcal{U}$ such that

  $J_i(v) < J_i(u^*) \quad i \in S$

  where $v^i = u^i$ for $i \in S$ and $v^i = u^{i*}$ for $i \notin S$.

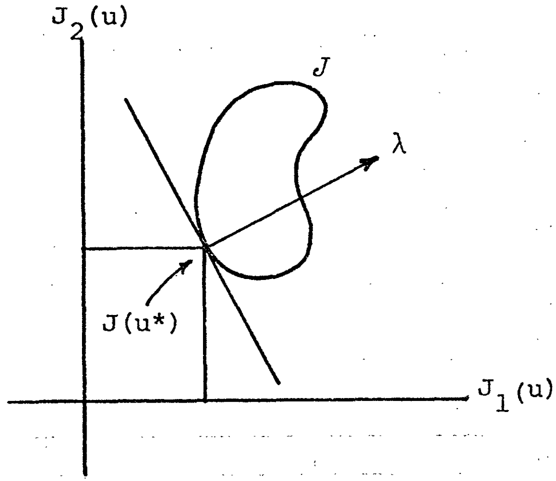
Thus an equilibrium point is one from which it does not pay any player to deviate unilaterally, a strategy is efficient if no strategy is better for everybody and a strategy is in the core if no coalition can act jointly to improve its lot. Evidently a core strategy is both efficient and an equilibrium, but equilibrium solutions are not necessarily efficient or conversely.

For $u \in \mathcal{U}$ denote $J'(u) = (J_1(u),\ldots,J_N(u))$ and let

$$\mathcal{J} = \{J(u) \mid u \in \mathcal{U}\}$$

This is a bounded subset of R^N, and a *sufficient* condition for efficiency of a strategy u^* is the existence of a non-negative vector $\lambda \in R^N$ such that

$$(7.8) \qquad \lambda' J(u^*) \leq \lambda' \xi \quad \text{for all } \xi \in \mathcal{J}$$



(see diagram for N=2). If $\mathcal{J}$ is convex, this condition is also necessary. It follows from results of Benes [2] (see the remarks following (4.20)) that convexity of the set $(f(t,x,U),\ c'(t,x,U_1)\ldots,c^N(t,x,U_N)) \subset R^{n+N}$ implies convexity of $\mathcal{J}$. Now (7.8) says that u^* is optimal for the control problem of minimizing the weighted average cost $J_\lambda(u) = \sum_i \lambda^i J_i(u)$. Fix $u^* \in \mathcal{U}$, and as in §4,, let $g^i, i=1,\ldots,N$, be adapted processes such that

$$E_{u^*}\left[\int_0^1 c_{is}^{u^*} ds + \Phi_i(x_1) \mid F_t\right] = J_i(u^*) + \int_0^t g_s^i \sigma_s dw_s^{u^*}$$

For any other strategy $u \in \mathcal{U}$ the right-hand side can be expressed, as in (4.11), as

$$J^i(u^*) + \int_0^t g_s^i \sigma_s dw^u + \int_0^t (H_s^i(u_s) - H_s^i(u_s^*))ds$$

where

$$H_s^i(u) = g^i f(t,x,u) + c_i(t,x,u)$$

Combining the remarks above with (4.16) shows that u^* is efficient if there exists $\lambda \in R^N$ such that

$$(7.9) \qquad \sum_i \lambda^i H^i(u_t^*) \leq \sum_i \lambda^i H^i(v), \qquad \text{a.e. for all } v \in U$$

under the convexity hypothesis, this condition is also necessary.

u^* is a Nash equilibrium if, for each i, u^{*i} minimizes $J_i(u^{*1},\ldots,u^{*(i-1)},u,u^{*(i+1)},\ldots,U^{*N})$ over $u \in \mathcal{U}_i$. Applying condition (4.16) we see that this will be the case if

$$(7.10) \qquad H^i(u_t^*) \leq H^i(v), \qquad \text{a.e. for all } v \in U_i, \quad i=1,2,\ldots,N$$

Thus u^* is an *efficient equilibrium* if u_t minimizes each "private" Hamiltonian as in (7.10) and also minimizes a "social" Hamiltonian (7.9) formed as a certain weighted average of these. Analogous conditions can be formulated under which u^* lies in the core.

For $(t,x,p_i,u) \in R^+ \times \Omega \times R^n \times U$ define the Hamiltonians

$$\overline{H}^i(t,x,p_i,u) = p_i' f(t,x,u) + c_i(t,x,u)$$

We say that the *Nash condition* holds if there exists for $i=1,\ldots,N$ measurable functions $u_i^0(t,x,p_1,\ldots,p_n)$ such that u_i^0 is a predictable process for each fixed $(p,u)=(p_1\ldots{}_N,u)$

and

$$\overline{H}^i(t,x,p_i,u_1^0(t,x,p),\ldots,u_N^0(t,x,p)) \le \overline{H}^i(t,x,p_i,u_1^0,\ldots,u_{i-1}^0,v,u_{i+1}^0,\ldots,u_N^0)$$

for all $v \in U_i$, for each $(t,x,p) \in R^+ \times \Omega \times R^{Nn}$. Uchida shows in [73] that *the game has a Nash equilibrium point if the Nash condition holds.* The proof is by a contradiction argument using the original formulation of the results of §4 as given in Davis and Varaiya [25]. Conditions under which the Nash condition holds are stated in [74].

Now consider the case N=2, $J_2(u) = -J_1(u)$, so that the game is 2-person, 0-sum. Then the core concept is ugatory, all strategies are efficient and an equilibrium is a *saddle point*, i.e. a strategy u* such that (denoting $J_1 = J$) for all $u \in \mathcal{U}$

$$J(u^{*1}, u^2) \le J(u^{*1},u^{*2}) \le J(u^1,u^{*2})$$

In this case the relevant condition is the *Isaacs' condition:* for each $(t,x,p) \in R^+ \times \Omega \times R^n$,

$$\max_{u_2 \in U_2} \min_{u_1 \in U_1} \overline{H}^1(t,x,p,u_1,u_2) = \min_{u_1 \in U_1} \max_{u_2 \in U_2} \overline{H}^1(t,x,p,u_1,u_2)$$

The main result is analogous to the above, namely that *a saddle strategy* u* *exists if the Isaacs' condition holds.* The argument, given by Elliott in [32], [35], is constructive, along the lines leading to the existence result (4.20) for the control problem. One considers first the situation where the minimizing player I announces his strategy $u_1 \in \mathcal{U}_1$ in advance. It is immediate from (4.20) that the maximizing player II has an optimal reply $u_2^0(u_1)$ to this. Now introduce the *upper value function*

$$W_t^+ = \bigwedge_{u_1 \in \mathcal{U}_1} E_{u_1,u_2^0(u_1)}\left[\int_t^1 c_1(s,x,u_1,u_2^0(u_1))ds + \Phi_1(x_1) \middle| F_t\right]$$

An analysis of this somewhat similar to that of §4 shows that player I has a best strategy, i.e. a strategy $u_1^0 \in \mathcal{U}_1$, such that

$$J(u_1^0,u_2^0(u_1^0)) = \min_{u_1 \in \mathcal{U}_1} J(u_1,u_2^0(u_1))$$

If it is player II who announces his strategy first, then we can define in an analogous manner the *lower value* function W_t^-. In general $W_t^+ \ge W_t^-$, but if the Isaacs' condition holds then $W_t^+ = W_t^-$ and it follows that u* given by $u^{*1} = u_1^0$, $u^{*2} = u_2^0(u_1^0)$ is a saddle strategy.

A somewhat more restricted version of this result was given by Varaiya in [75], using a compactness-of-densities argument similar to that of Benes [1] and Duncan and Varaiya for the control problem. No results are available if the players do not have complete observations. Some analogous results for a differential game including a jump process component are given in [49].

7.3 Optimal stopping and impulse control

In the conventional formulation of optimal stopping one is given a Markov process x_t on a state space S and a bounded continuous function ϕ on S, and asked to find a Markov time τ such that $E_x\phi(x_\tau) \ge E_x\phi(x_\sigma)$ for all $x \in S$ and Markov times σ. Let

$$\psi(x) = \sup_\sigma E_x \phi(x_\sigma)$$

Then under some regularity conditions ψ is the "least excessive majorant" of ϕ (i.e., $\psi(x) \geq \phi(x)$ and $\psi(x_t)$ is a supermartingale) and the first entrance time of x_t into the set $\{x: \phi(x) = \psi(x)\}$ is an optimal time. See [4], and the references there. If we define $X_t = \phi(x_t)$ and $W_t = \psi(x_t)$ then τ maximizes $E_x X$ and $\tau = \inf\{t: X_t = Z_t\}$. Thus the optimal stopping problem generalizes naturally as follows.

Let (Ω, F, P) be a probability space and $(F_t)_{t \geq 0}$ be an increasing, right-continuous, completed family of sub-σ-fields of F. Let T denote the set of F_t-stopping times and X_t be a given positive, bounded optional process defined on $[0,\infty]$. The optimal stopping problem is then to find $T \in T$ such that

$$EX_T = \max_{S \in T} EX_S$$

This problem is studied by Bismut and Skalli in [8]. The simplest case occurs when X_t satisfies the following hypothesis:

(7.11) *Let* $\{T_n, T\}$ *be stopping times such that* $T_n \uparrow T$ *or* $T_n \downarrow T$. *Then* $EX_{T_n} \to EX_T$.

Criteria under which (7.11) holds are given in [8].

An essential role in this problem is played by the *Snell envelope* of X_t, introduced by Mertens [62, Theorem 4]. He shows that the set of all supermartingales which majorize X_t has a smallest member, denoted W_t, which is characterized by the property that for any stopping time T and σ-field $G \subset F_t$,

$$E[W_T | G] = \operatorname*{ess\,sup}_{S \geq T} E[X_S | G]$$

Thus in particular for each fixed time t

$$W_t = \operatorname*{ess\,sup}_{S \geq t} E[X_S | F_t]$$

so that W_t is the *value function* for the optimal stopping problem. Under condition (7.11) W_t is *regular* [63, VII D33] and hence has the Meyer decomposition

$$W_t = M_t - B_t$$

where M_t is a martingale and B_t a *continuous* increasing process with $B_0 = 0$. Now define

$$D_0' = \inf\{t>0: B_t > 0\}$$

and

$$A = \{(t,\omega): X_t(\omega) = W_t(\omega)\}$$

The *debut* of A is the stopping time $D_0^A = \inf\{t: (t,\omega) \in A\}$. It is shown in [8] that $D_0^A \leq D_0'$ and that:

(7.12) *A stopping time* T *is optimal if and only if the graph of* T *is contained in* A *and* $T \leq D_0'$

In particular, both D_0^A and D_0' are optimal.

This result implies an optimality criterion similar to (5.3): if T is optimal then $B_{t \wedge T} = 0$ so that $W_{t \wedge T} = M_{t \wedge T}$ is a martingale, and conversely if $W_{t \wedge T}$ is a martingale then it is easily seen that T must satisfy the conditions of (7.12).

Analogous results can be obtained for processes more general than those satisfying (7.11); the details are more involved and only ε-optimal stopping times may exist.

Impulse control: Space precludes any detailed discussion of this topic, but it should be mentioned that a martingale treatment has been given by Lepeltier and Marchal [59]. In the simplest type of problem one has a stochastic differential equation

$$dx_t = f(x_t)dt + \sigma(x_t)dw_t$$

A *strategy* $\delta = \{T_n, Y_n\}$ consists of an increasing sequence of stopping times T_n and a sequence of random variables Y_n such that Y_n is F_{T_n}-measurable. The corresponding trajectory is x_t^δ defined by

$$x_0^\delta = x \text{ (given)}$$

$$\left.\begin{aligned} dx_t^\delta &= f(x_t^\delta) + \sigma(x_t^\delta)dw_t \\ x_{T_n}^\delta &= x_{T_n^-}^\delta + Y_n \end{aligned}\right\} \quad t \in [T_n, T_{n+1}[$$

The strategy δ is to be chosen to minimize

$$J(\delta) = E[\sum_n I_{(T_n \leq 1)} + \int_0^1 c(x_s^\delta)ds]$$

A value function and conditions for optimality can be obtained along the lines of §5. It is worth pointing out that the above system obviously has a Markovian flavor about it, and indeed it is shown in [59] that the value function is Markovian (i.e., at time t it depends on x^δ only through x_t^δ) even though the controls δ are merely assumed to be non-anticipative. Some further remarks on this are given in the next section.

7.4 Markovian systems

Let us return to the problem of §4 and suppose that the system equation and cost are

$$dx_t = f(t,x_t,u_t)dt + \sigma(t,x_t)dw_t^u$$

$$J(u) = E_u[\int_0^1 c(t,x_t,u_t)dt + \Phi(x_1)],$$

i.e., we have a diffusion model as considered in §2. In §4 the admissible controls $\mathcal{U}$ were general non-anticipative functionals but here it seems clear that feedback controls of the form $u(t,x_t)$ should be adequate. Denote by $\mathcal{M}$ the set of measurable functions $u\colon [0,1] \times R^n \to U$; then $\mathcal{M} \subset \mathcal{U}$ if we identify $u \in \mathcal{M}$ with the process $u_t = u(t,x_t)$, and x_t is a Markov process under measure P_u. Thus we can define the *Markovian value function* $W^M(t,x)$ as (with obvious notation)

$$W^M(t,x) = \bigwedge_{u \in \mathcal{M}} E_{t,x}^u[\int_t^1 c(s,x_s,u_s)ds + \Phi(x_1)]$$

The conjecture then is that $W^M(t,x_t) = W_t$ a.e. (W_t being defined as in §4) so that in particular

$$\inf_{u \in M} J(u) = \inf_{u \in U} J(u)$$

This is easily established (see [25, §6]) if it can be shown that W^M satisfies a principle of optimality similar to (4.7). However this is not clear, as there is still, to my knowldge, no direct proof that the class M satisfies the ε-lattic property. An argument along the lines given in §5 fails because it involves "mixing" two controls $u_1, u_2 \in M$ to form a control v by taking

$$v_s = \begin{cases} u_1(s,x_s) I_A \\ u_2(s,x_s) I_{A^c} \end{cases}$$

where $s \geq t$ and $A \in F_t$. But then v_s is of course no longer Markov. Thus the results presented in §6 of [23] must be regarded as incomplete.

This problem has been dealt with in the case of controlled Markov jump processes by Davis and Wan [26]. There it is possible to "mix" two controls in a more ingenious way which, however, uses the special structure of the sample paths very explicitly and hence does not generalize to other problems. An alternative approach would be to start with the value process W_t as previously defined and to show directly that $W_t = \tilde{W}(t,x_t)$ for some function $\tilde{W}$. This has been done by Lepeltier and Marchal [59] for impulse control problems but again the argument is very problem-specific.

My general conclusion from the above is that the direct Martingale approach is not particularly well adapted to Markovian problems, and that more information can be obtained from methods such as those of Bismut [5] which are specially tailored for Markov processes.

8. CONCLUDING REMARKS

The successes of martingale methods in control are twofold: firstly the essence of the optimality principle is revealed in the general formulation (5.3), and in particular the fundamental difference between the situations of complete and of incomplete observations is clearly brought out; and secondly, the power of the submartingale decomposition provides, in effect, a weak form of differentiation which enables minimum principles and existence of optimal controls to be established with few technical restrictions. The drawbacks of the method are that it does not lead naturally to computational techniques, and there are difficulties in handling Markovian systems and problem formulations of the "separation principle" type.

Here are a few suggestions for further research.

(8.1) Obtain a more explicit characterization of the "adjoint process" g_t of §4. Comparisons with deterministic optimal control theory and other forms of stochastic minimum principle [6], [53] suggest that it should satisfy some form of "adjoint equation," yet little is known about this unless the optimal control is smooth [44].

(8.2) To my knowledge martingale methods have not been applied seriously to infinite-time problems (see Kushner [55] for some results using methods similar to those of Bismut [5]).

(8.3) The partially-observable problem continues to elude a satisfactory treatment. In particular there are no good existence theorems, and experience with the separation theorem (§6) suggests that these may be hard to get. My feeling is that the proper formulation of partially-observable problems must explicitly include filtering, since it is the conditional distribution of the state given the observations that is the true "state" of the system. A lot of information about nonlinear filtering is available [60] but, again using the separation principle as a cautionary tale, it is far from clear how to incorporate this into the martingale framework. Possibly some entirely different approach, such as Nisio's nonlinear semigroup formulation, will turn out to be more appropriate. See [20] for a step in this direction.

(8.4) Show that the ε-lattice property holds in some generality for Markovian systems with Markov controls (cf. §7.4).

(8.5) Give a constructive treatment of Uchida's result [73] on the existence of Nash equilibirum points in stochastic differential games.

(8.6) Is *mutual* absolute continuity of the measures P_u really necessary for the existence result (4.20)? If not then better existence results could possibly be obtained for problems such as controlled jump processes (§7.1) where mutual absolute continuity does not arise so naturally.

9. REFERENCES

SPm/LNMn Université de Strasbourg Séminaire de Probabilités m, Lecture Notes in Mathematics vol. n, Springer-Verlag, Berlin-Heidelberg-New York

SICON SIAM Journal on Control (and Optimization)

ZW Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete

[1] V.E. Beneš, Existence of optimal strategies based on specified information, for a class of stochastic decision problems, SICON 8 (1970) 179-188

[2] V.E. Beneš, Existence of optimal stochastic control laws, SICON 9 (1971) 446-475

[3] V.E. Beneš, Full "bang" to reduce predicted miss is optimal, SICON 15 (1976) 52-83

[4] A. Bensoussan and J.L. Lions, Applications des inéquations varationelles en contrôle stochastique, Dunod, Paris, 1978

[5] J.M. Bismut, Théorie probabiliste du contrôle des diffusions, Mem. Americ. Math. Soc. 4 (1976), no. 167

[6] J.M. Bismut, Duality methods in the control of densities, SICON 16 (1978) 771-777

[7] J.M. Bismut, An introductory approach to duality in optimal stochastic control, SIAM Review 20 (1978) 62-78

[8] J.M. Bismut and B. Skalli, Temps d'arrêt optimal, théorie général des processus et processus de Markov, ZW 39 (1977) 301-313

[9] R. Boel and M. Kohlmann, Stochastic control over double martingales, in "Analysis and Optimization of Stochastic Systems" ed. O.L.R. Jacobs, Academic Press, New York/London 1979

[10] R. Boel and M. Kohlmann, Semimartingale models of stochastic optimal control with applications to double martingales, preprint, Institut fur Angewandte Mathematik der Universität Bonn, 1977

[11] R. Boel and P. Varaiya, Optimal control of jump processes, SICON 15 (1977) 92-119

[12] R. Boel, P. Varaiya and E. Wong, Martingales on jump processes I and II, SICON 13 (1975) 999-1061

[13] P. Brémaud and J. Jacod, Processus ponctuels et martingales: resultats recents sur la modelisation et le filtrage, Adv. Appl. Prob. 9 (1977) 362-416

[14] P. Brémaud and J.M. Pietri, The role of martingale theory in continuous-time dynamic programming, Tech. report, IRIA, Le Chesnay, France, 1978

[15] J.M.C. Clark, The representation of functionals of Brownian motion by stochastic integrals, Ann. Math. Stat. 41 (1970) 1285-1295

[16] M.H.A. Davis, On the existence of optimal policies in stochastic control, SICON 11 (1973) 507-594

[17] M.H.A. Davis, The representation of martingales of jump processes, SICON 14 (1976) 623-638

[18] M.H.A. Davis, The separation principle in stochastic control via Girsanov solutions, SICON 14 (1976) 176-188

[19] M.H.A. Davis, Functionals of diffusion processes as stochastic integrals, submitted to Math. Proc. Camb. Phil. Soc.

[20] M.H.A. Davis, Nonlinear semigroups in the control of partially-observable stochastic systems, in Measure Theory and Applications to Stochastic Analysis, ed. G. Kallianpur and D. Kölzow, Lecture Notes in Mathematics, Springer-Verlag, to appear

[21] M.H.A. Davis and J.M.C. Clark, "Predicted Miss" problems in stochatic control, Stochastics 2 (1979)

[22] M.H.A. Davis and R.J. Elliott, Optimal control of a jump process, ZW 40 (1977) 183-202

[23] M.H.A. Davis and M. Kohlmann, Stochastic control by measure transformation: a general existence result, preprint, Institut fur Angewandte Mathematik der Universitat Bonn (1978)

[24] M.H.A. Davis and P.P. Varaiya, Information states for linear stochastic systems, J. Math. Anal. Appl. 37 (1972) 387-402

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

[26] M.H.A. Davis and C.B. Wan, The principle of optimality for Markov jump processes, in "Analysis and Optimization of Stochastic Systems" ed. O.L.R. Jacobs, Academic Press, New York/London, 1979

[27] C. Dellacherie, Capacités et processus stochastiques, Springer-Verlag, Berlin, 1972

[28] C. Doléans-Dade, Quelques applications de la formule de changement de variables pour les semimartingales, ZW 16 (1970) 181-194

[29] T.E. Duncan, Dynamic programming criteria for stochastic systems in Riemannian manifolds, Applied Math. & Opt. 3 (1977) 191-208

[30] T.E. Duncan and P.P. Varaiya, On the solutions of a stochastic control system, SICON 9 (1971) 354-371

[31] T.E. Duncan and P.P. Varaiya, On the solutions of a stochastic control system II, SICON 13 (1975) 1077-1092

[32] R.J. Elliott, The existence of value in stochastic differential games, SICON 14 (1976) 85-94

[33] R.J. Elliott, Double martingales, ZW 34 (1976) 17-28

[34] R.J. Elliott, The optimal control of a stochastic system, SIAM J. Control & Optimization 15 (1977) 756-778

[35] R.J. Elliott, The existence of optimal strategies and saddle points in stochastic differential games, in Differential Games and Applications, ed. P. Hagedorn, Lecture Notes in Control and Information Sciences 3, Springer-Verlag, Berlin, 1977

[36] R.J. Elliott, Levy systems and absolutely continuous changes of measure, J. Math. Anal. Appl. 61 (1977) 785-796

[37] R.J. Elliott, The optimal control of a semimartingale, 3rd Kingston Conference on Control Theory, Kingston, RI (1978)

[38] R.J. Elliott, The martingale calculus and its applications, this volume

[39] R.J. Elliott and P.P. Varaiya, A sufficient condition for the optimal control of a partially observed stochastic system, in "Analysis and Optimization of Stochastic Systems," ed. O.L.R. Jacobs, Academic Press, New York/London, 1979

[40] W.H. Fleming, Optimal continuous-parameter stochastic control, SIAM Rev. 11 (1969) 470-509

[41] W.H. Fleming and R.W. Rishel, Deterministic and stochastic optimal control, Springer-Verlag, New York, 1975

[42] M. Fujisaki, G. Kallianpur and H. Kunita, Stochastic differential equations for the nonlinear filtering problem, Osaka J. Math. 9 (1972) 19-40

[43] I.V. Girsanov, On transforming a certain class of stochastic processes by absolutely continuous substitution of measures, Theory of Prob. and Appls. 5 (1960) 285-301

[44] U.G. Haussmann, On the stochastic maximum principle, SICON 16 (1978) 236-251

[45] U.G. Haussmann, Functionals of Itô processes as stochastic integrals, SICON 16 (1978) 252-269

[46] U.G. Haussmann, On the integral representation of functionals of Ito processes, Stochastic 2 (1979)

[47] J. Jacod, Multivariate point processes: predictable projection, Radon-Nikodym derivatives, representation of martingales, ZW 31 (1975) 235-253

[48] J. Jacod and J. Memin, Caracteristiques locales et conditions de continuité absolue pour les semimartingales, ZW 35 (1976) 1-37

[49] M. Kohlmann, A game with Wiener noise and jump process disturbances, submitted to Stochastics

[50] M. Kohlmann, On control of jump process, preprint, Institut für Angewandte Mathematik der Universitat Bonn, 1978

[51] N.V. Krylov, Control of a solution of a stochastic integral equation theory of prob. and Appls. 17 (1972) 114-131

[52] H. Kunita and S. Watanabe, On square-integrable martingales, Nagoya Math. Journal 30 (1967) 209-245

[53] H.J. Kushner, Necessary conditions for continuous-parameter stochastic optimization problems, SICON 10 (1972) 550-565

[54] H.J. Kushner, Probability Methods for Approximations in Stochastic Control and for Elliptic Equations, Academic Press, New York, 1977

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

[56] E. Lenglart, Transformation des martingales locales par changement absolument continu de probabilités, ZW 39 (1977) 65-70

[57] D. Lepingle and J. Memin, Sur l'integrabilité uniforme des martingales exponentielles, ZW 42 (1978) 175-203

[58] J.P. Lepeltier and B. Marchal, Sur l'existence de politiques optimales dans le contrôle integro-differentiel, Ann. Inst. H. Poincaré 13 (1977) 45-97

[59] J.P. Lepeltier and B. Marchal, Techniques probabilistes dans le contrôle impulsionelle, Stochastics 2 (1979)

[60] R.S. Liptser and A.N. Shiryayev, Statistics of Random Processes, vols I and II, Springer-Verlag, New York, 1977

[61] J. Memin, Conditions d'optimalité pour un problème de contrôle portant sur une famille de probabilités dominées par une probabilité P, preprint, University of Rennes, 1977

[62] J.F. Mertens, Processus stochastiques géneraux, applications aux submartingales, ZE 22 (1972) 45-68

[63] P.A. Meyer, Probability and Potentials, Blaisdell, Waltham, MA, 1966

[64] P.A. Meyer, Un cours sur les intégrales stochastiques, SP10/LNM 511 (1976)

[65] R.E. Mortensen, Stochastic optimal control with noisy observations, Int. J. Control 4 (1966) 455-464

[66] M. Nisio, On a nonlinear semigroup attached to stochastic control, Pub. Res. Inst. Math. Sci., Kyoto, 12 (1976) 513-537

[67] A.A. Novikov, On an identity for stochastic integrals, Theor. Probability Appl. 17 (1972) 717-720

[68] R. Rishel, Necessary and sufficient conditions for continuous-time stochastic optimal control, SICON 8 (1970) 559-571

[69] J.H. van Schuppen and E. Wong, Transformations of local martingales under a change of law, Ann. Prob. 2 (1974) 879-888

[70] C. Striebel, Martingale conditions for the optimal control of continuous-time stochastic systems, 5th Symposium on nonlinear estimation and applications, San Diego, CA, 1974

[71] D.W. Stroock and S.R.S. Varadhan, Diffusion processes with continuous coefficients, Comm. Pure and Appl. Math. 22 (1969) 345-400, 479-530

[72] B.S. Tsyrelson, An example of a stochastic equation having no strong solution, Theory of Prob. and Appls. 20 (1975) 427-430

[73] K. Uchida, On the existence of a Nash equilibrium point in N-person nonzero sum stochastic differential games, SICON 16 (1978) 142-149

[74] K. Uchida, A note on the existence of a Nash equlibrium point in stochastic differential games, to appear in SICON

[75] P. Varaiya, Differential games, Proc. VI Berkeley Symposium on Math. Stat. and Prob., vol. 3, 687-697, Univ. of California Press, Berkeley, 1972

[76] P.P. Varaiya, N-person stochastic differential games, SICON 14 (1976) 538-545

[77] C.B. Wan and M.H.A. Davis, existence of optimal controls for stochastic jump processes, to appear in SICON

[78] W.M. Wonham, On the separation theorem of stochastic control, SICON 6 (1968) 312-326

[79] C. Yoeurp, Decompositions des martingales locales et formules exponentielles, SP10/LNM 511 (1976) 432-480

SR gewidmet
Dezember 1978