OPTIMAL CONTROL OF JUMP PROCESSES

by

Rene Kamiel Boel

Memorandum No. ERL-M448

3 July 1974

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

ACKNOWLEDGEMENT

The author expresses his sincere thanks to Professor P. Varaiya. Without his enthusiasm during countless enlightening discussions, this work would have been impossible. Sincere thanks are also due for the many interesting conversations with Professor E. Wong. Finally, I thank J. van Schuppen, with whom I extensively discussed the research reported here.

Financial support for this work was provided by an ESRO-NASA International Fellowship in 1971-73 and in 1973-74 by a research assistantship at the Electronics Research Laboratory, the funds coming from NSF-10656X3 and AFOSR-F44620-71-C-0087.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

The results presented here deal with the optimization of dynamical systems with random perturbations. First, a mathematical model is developed, where we consider a family of stochastic processes, with the same sample paths but different probability distributions. These distributions depend on actions taken by a decision-maker, at different points in time, using some information about the past of the process. One can therefore say that the family of stochastic processes is indexed by "control laws," that is rules for choosing a certain action, depending on time and available information. With each stochastic process, there also corresponds a cost-function, parameterized by the same control law, and which allows us to define an <u>optimal control law</u>. Afterwards this mathematical model will be shown to yield significant theorems, for more specific stochastic processes.

This model is different from the approach usually taken for optimization of Markov processes, as described in the review paper by Fleming [13]. There, a probability space $(\Omega,\mathcal{F},\mathcal{P})$ is given a priori, and different processes have different sample paths defined by having control-dependent coefficients in some differential equation. This causes a number of problems. First, if the available observations depend on the stochastic process $(x_t,\mathcal{F}_t,\mathcal{P}_o)$ defined by

$$dx_t = f(t,x_s,s\leq t,u(t),n(t)) \tag{1}$$

where $n(t)$ is the perturbation defined a priori, then a σ-field like $\mathcal{F}_t^y = \sigma(g(x_s),s\leq t)$ will depend on the control law u chosen, making variational analysis very difficult. Secondly, to derive any criterion for optimality, one needs existence of a solution to (1), which imposes strong continuity conditions on f. Since optimal control laws are often discontinuous (e.g. bang-bang control) this is unacceptable for optimal control. And third, while (1) may be reinterpreted for discontinuous processes x_t (see for example Skorokhod [29]), this poses some mathematical problems.

Quite a different description has been used by Blackwell [3], Ross [27] among others, in problems where $(x_t, \mathcal{F}_t, \mathcal{P})$ is a Markov (or semi-Markov) chain. There it is assumed that at each time an action is taken which specifies the transition probabilities, and the cost associated with each type of jumps. If at each time only a finite number of actions are possible, this leads to a dynamic programming equation, with very nice properties. However, when the action space is a continuum, more care must be taken. Then relations between probability measures must be specified, which requires knowledge of the Radon-Nikodym derivatives.

One case where this derivative is known is for translations in Wiener space (Girsanov [17]). This led Stroock-Varadhan [31] to the following definition of the solution of an Ito differential equation:

$$x_t = x_o + \int_o^t f(x,s,u)\, ds + w_t \tag{2}$$

If $(x_t, \mathcal{F}_t, \mathcal{P}_o)$ is a Wiener process, then under the measure $\mathcal{P}_u$,

$$\frac{d\mathcal{P}_u}{d\mathcal{P}_o} = \exp\left(\int_o^t f(x,s,u)\, dw_s - \frac{1}{2}\int_o^t f^2(x,s,u)\, ds\right)$$

$(w_t, \mathcal{F}_t, \mathcal{P}_u)$ is a Brownian motion, and $(x_t, \mathcal{F}_t, \mathcal{P}_u)$ is a solution to the differential equation (2). This solution exists without the objectionable continuity conditions on f, and since the sample paths of x_t are unchanged the partial information σ-field $\mathcal{F}^y_t$, will also be unchanged if f depends on some parameter u. This result was used almost immediately by Beneš, [1], [2], Duncan-Varaiya [10] to prove existence of an optimal solution if $f(x_s,s,u)$ depends on a control law u, which also determines a cost J(u).

Davis-Varaiya [8], also used it to derive the principle of optimality of dynamic programming. This takes roughly the following form. The minimal expected cost after t is smaller than the expected value, given the information at time t, of the cost of using any control law in [t,t+h) and thereafter returning to an optimal control law. Using supermartingale decomposition theorems it is possible to transform this into a necessary and sufficient condition for optimality.

The model described above is obviously a special case of the model described in the first paragraph. It turns out consequently, that all the results mentioned so far, will hold for arbitrary processes. This is shown in Chapter 2, except for the existence result which requires a more careful derivation.

The results in §2 remain rather abstract, because it is difficult to relate the Radon-Nikodym derivative $\frac{d\mathcal{P}_u}{d\mathcal{P}_o}$ with the dynamics of the system. From the results of van Schuppen-Wong [34] it follows that this will be possible if all martingales on $(\mathcal{F}_t, \mathcal{P}_o)$ can be represented as stochastic integrals with respect to some basic process. In [4] it has been shown that this holds for σ-fields $\mathcal{F}_t$ generated by processes that are piecewise constant, and have finitely many totally inaccessible jumps, of different types, in a finite interval. This includes many processes of practical interest, such as point processes, branching processes, Markov chains, queueing processes. This general class has been called jump processes. For all properties on jump processes, used hereafter, the reader is referred to [4] and [5].

For these jump processes, it is now possible, just as in the Brownian motion case, to bring the necessary and sufficient condition for optimality into the form of a Hamilton-Jacobi equation. This can be summarized as follows: there exists a function $H(x,p,t,u)$, the Hamiltonian, depending on the observations about the past of x_t, on a "costate" p_t, which is $\mathcal{F}^y_t$-adapted and can depend on u and on the "dynamics" of the process. Then, $u^*(t,\omega)$ is optimal if and only if

$$\min_{u \in U} H(x,p,t,u) = H(x,p,t,u^*) = 0.$$

Unfortunately, the result is not a Pontryagin maximum principle because the costate (parameter) p does not satisfy a known equation. This, together with the special cases of complete information and of a Markov process, is discussed in Chapter 3.

Finally, in Chapter 4, it will be shown how the previous results can be applied to practical problems. After a few theorems, useful for computational purposes, attention is given to the problem of modeling a system as a jump process. First of all, a good probability measure $\mathcal{P}_o$ has to be chosen on $(\Omega,\mathcal{F})$. It has to be sufficiently complicated, so that all reasonable probability measures $\mathcal{P}_u$ are absolutely continuous with respect to it, while at the same time it has to be mathematically tractable. Also the way in which the control law u influences the measure $\mathcal{P}_u$ and the cost function J(u) has to be described as simple as possible.

It will be obvious from §4 that only for jump processes can one hope to get an explicit, or even approximate, optimal control law, since the criteria become much simpler. In [5], it has also been ovserved that a number of detection and estimation problems can be solved for jump processes. In fact, all the results of van Schuppen [33] for Poisson processes, extend to jump processes. This indicates the usefulness of jump processes in the study of discontinuous stochastic processes.

This paper is a continuation of the work reported in [4], [5]. The reader is therefore referred to these reports for all the theory on jump processes. In particular, the reader unfamiliar with the work of Meyer on stochastic integration, [26], can find a summary of this in §2 of [5]. Some of the notation used in §2-3 will be introduced below:

Notation-Conventions

A stochastic process $(x_t,\mathcal{F}_t,\mathcal{P})$, on a given probability space $(\Omega,\mathcal{F},\mathcal{P})$ will always be assumed right-continuous. For submartingales $\{X_t\}$, such that EX_t is right-continuous, one can always choose such a version anyway.

A <u>uniformly integrable martingale</u> ($\in \mathcal{M}^1(\mathcal{F}_t,\mathcal{P})$) is a martingale such the $\sup_t E|m_t| < \infty$, and $m_o = 0$ a.s. Similarly $\mathcal{M}^2(\mathcal{F}_t,\mathcal{P})$ denotes (uniformly) <u>square integrable martingales</u>; and $\mathcal{M}^1_{loc}(\mathcal{F}_t,\mathcal{P})$, $\mathcal{M}^2_{loc}(\mathcal{F}_t,\mathcal{P})$ require the boundedness condition only up to stopping times s_n, $s_n \uparrow \infty$ as $n \uparrow \infty$.

The family $\mathcal{A}^+(\mathcal{F}_t,\mathcal{P})$ contains all non-decreasing processes a_t, such that $a_o = 0$ and $\sup_t E\, a_t < \infty$, that is processes of <u>integrable variation</u>. $\mathcal{A} = \mathcal{A}^+ - \mathcal{A}^+$ and $\mathcal{A}_{loc}$ is as usual obtained by a stopping time.

Also $L^1(A,\mathcal{Q},\mu) = \{x(a)\colon A \to \mathbb{R},\ \int_A |x(a)|\mu(da) < \infty\}$ are called <u>integrable functions</u>, for different measurable space $(A,\mathcal{Q})$. This should not be confused with $L^1(Q^x) = L^1(P^x) = L^1(\tilde{P}^x)$

$$=\{f \mid f\colon Z\times I\times\Omega \to \mathbb{R},\ f \text{ predictable}$$

$$E\int_Z \int_I |f(z,t)|P(dz,dt) < \infty\}$$

Here <u>predictable</u> means that for fixed $z \in Z$, $f(z,\cdot,\cdot)\colon I \times \Omega \to \mathbb{R}$ is measurable with respect to the σ-field $\mathcal{P}$ generated by the left-continuous real-valued functions $g(t,\omega)$ on $I \times \Omega$.

CHAPTER 2

OPTIMALITY CRITERIA: GENERAL CASE

In this section optimality criteria will be derived for a system where the control acts through a change of the probability measure and of the cost. The structure of the underlying stochastic process is completely unspecified. The price one pays for this generality, is that all the criteria involve a process, defined through a functional minimization, as difficult as the original problem.

2.1. Mathematical framework

The following stochastic processes are defined over a closed subset I of the real line, usually [0,1] or [0,∞] or the natural numbers. All the results will be written in the notation of continuous time (integrals, derivatives). Since the discrete time results are simpler in most cases, the required changes are obvious

On the probability space $(\Omega,\mathcal{F},\mathcal{P}_o)$ a stochastic process $(x_t,\mathcal{F}_t,\mathcal{P}_o)$, over I, is given. To simplify the notation, assume the initial time is 0, and $\mathcal{F}_o = \{\phi,\Omega\}$, the trivial σ-field. The control action consists in changing this process, to a different process $(x_t,\mathcal{F}_t,\mathcal{P}_u)$, having the same sample paths, but a different probability measure $\mathcal{P}_u$, absolutely continuous with respect to $\mathcal{P}_o$. Note that the requirement of absolutely continuity is very restrictive when $I = [0,\infty)$. With each such control action is associated a cost $J(u) \geq 0$, which is to be minimized. This notation requires that there exists one single parameter u, the control law, which determines the probability measure $\mathcal{P}_u$ and the cost J(u).

The following assumptions will be made about the control law u. There is a given increasing family $\{\mathcal{F}_t^y\}$ of sub-σ-algebras of $\mathcal{F}_t(\mathcal{F}_t^y \subset \mathcal{F}_t,\ t \in I)$. This expresses the fact that the decision-maker does not have all the information $\mathcal{F}_t$, available. Hence, the control law u is a function $u(t,\omega)$: $I \times \Omega \to U$, with U a fixed set of control values; this function is assumed $\mathcal{F}_t^y$-predictable. $u(t,\omega)$ is defined without reference to a probability measure. It can then be used to define $\mathcal{P}_u$, which makes $(u(t),\mathcal{F}_t,\mathcal{P}_u)$ a stochastic process on $(\Omega,\mathcal{F},\mathcal{P}_u)$. It is always assumed that if $A \in \mathcal{F}_t$ then $\mathcal{P}_u(A)$ depends only on the control law restricted to the interval $[0,t]$. The class of admissible control laws, $\mathcal{U}$, contains all $\mathcal{F}_t^y$-predictable functions $u : I \times \Omega \to U$, such that $\mathcal{P}_u << \mathcal{P}_o$ and $\mathcal{P}_u(\Omega) = 1$. Throughout this paper the following assumption is made. Let $u,v \in \mathcal{U}$, then $(u,t,v) \in \mathcal{U}$ where $(u,t,v)(s) \begin{cases} = u(s) & s \leq t \\ = v(s) & s > t \end{cases}$

The following notation will also we used:

$$L(u) = \frac{d\mathcal{P}_u}{d\mathcal{P}_o}, \text{ the Radon-Nikodym derivative and } L_o^t(u) = E_o[L(u)|\mathcal{F}_t]$$

Then $(L_o^t(u),\mathcal{F}_t,\mathcal{P}_o)$ is a positive uniformly integrable martingale $\mathcal{M}^1(\mathcal{F}_t,\mathcal{P}_o))$.

The influence of the control law u on the cost $J(u)$ will be described as follows:

$$J(u) = E_u\left[\int_I r_o^s \cdot c(s,u(s))dA^u(s) + r_o^f \cdot T_f\right] \tag{1}$$

[1] E_o denotes integration over $\mathcal{P}_o$. Similarly E_u will be used for integration over $\mathcal{P}_u$.

$$E_o\left[L(u)\left(\int_I r_o^s\cdot c(s,u(s))\cdot dA^u(s) + r_o^{t_f}\cdot T_f\right)\right]$$

or $J : \mathcal{U} \to \mathbb{R}_+$

Here t_f denotes the final time of the set I(usually 1 or ∞), and the following functions are given:

(i) <u>instantaneous cost function</u> $c(t,u,\omega) : I \times U \times \Omega \to \mathbb{R}_+$ which for each fixed u, is $\mathcal{F}_t$-adapted and $\mathcal{B}_I \otimes \mathcal{F}$ measurable on $I \times \Omega$ ($\mathcal{B}_I$ the Borel sets on I). This is a measure of the increase in cost in an infinitesimal time interval $[t,t+\Delta t)$, given $u \in \mathcal{U}$ takes the value $u(t,\omega) = u$ and depending on the past of the proces (Note that u is used for both the control law and its present value. The exact meaning will always be obvious from the context).

(ii) an $\mathcal{F}_t$-adapted increasing process $A^u(t,\omega)$, which can be continuous or discontinuous, allowing the Stieltjes integral in (1) to be discontinuous. The most useful processes are:

a) $A^u(t) = t$; whenever $A^u(s)$ is absolutely continuous with respect to Lebesgues measure, this case can be obtained.

b) $A^u(t) = \sum 1_{\{t \geq \tau_i\}}$, a counting process ($\tau_i$ is an $\mathcal{F}_t$-stopping time); here the cost increases only at certain random points in time.

c) $A^u(t)$ is the predictable increasing process associated with the counting process in b), and which can replace the counting process, if $c(t,u,\omega)$ is predictable, since the values of the integrals are equal (see Meyer [26]).

(iii) <u>discounting rate</u> $r_s^t(\omega) : I \times I \times \Omega \to \mathbb{R}_+$, $s \leq t$ which for each s is $\mathcal{F}_t$-adapted, jointly $\mathcal{B}_I \otimes \mathcal{F}$ measurable and uniformly integrable ($E_o r_s^t < K$).

Moreover: $r_{t_1}^{t_3} = r_{t_1}^{t_2} \cdot r_{t_2}^{t_3}$ a.s. $(\mathcal{P}_o)$ for $t_1 \le t_2 \le t_3$

$r_t^t = 1$ a.s. $(\mathcal{P}_o)$

This function expresses the fact that future costs are weighed differently (usually less heavily) than present costs. Also, if $I = \mathbb{R}_+$, a discounting rate $r_s^t < 1$ for all $t > s$, is necessary to make the total cost finite.

(iv) <u>final cost</u> $J_f(\omega) : \Omega \to \mathbb{R}_+$, an $\mathcal{F}$ measurable function. This measures the cost incrrred at or after the final time t_f. If $I = [0,\infty)$ it is logical to take $J_f = 0$, since $r_o^{t_f} = 0$ anyway. The optimization problem considered here, is: find a control law $u^* \in \mathcal{U}$ such that

$$J(u^*) = J^* = \bigwedge_{u \in \mathcal{U}} J(u) \tag{2}$$

where J^*, the infimum of $\{J(u) | u \in \mathcal{U}\}$ is well defined since $J(u) \ge 0$.

<u>Remarks</u>: 1. A random time interval $[0,T] \subset I$, with T an $\mathcal{F}_t$-stopping time, can be considered by making $c(t,u,\omega) = 0$, $t \ge T(\omega)$ or $A^u(t,\omega) = A^u(T(\omega),\omega)$, $t \ge T(\omega)$. If the stopping time T is independent of the processes influenced by the control, a discounting rate $r_o^t = \mathcal{P}_o(T \ge t)$ also transforms the stopped problem into an infinite time problem.

2. The discounting rate $r_s^t(\omega)$ is not allowed to depend explicitly on the control law u. In an economic example this means that the decision-maker cannot decide on the interest rates. However the distribution of $r_s^t(\omega)$ depends on $\mathcal{P}_u$, and can thus be influenced.

3. Except for the special results with complete information or Markovian assumptions, the final cost J_f can depend explicitly on the control law $u \in \mathcal{U}$.

4. The generalization, where the instantaneous cost $c(t,u,\omega)$ depends on the control law u, used in [o,t] can be included by letting $u(t,\omega) \in U_t$, a time-dependent set of control values (functions $[o,t] \to U$ here). This would only burden the notation.

2.2. Principle of optimality

From now on the class $\mathcal{U}$ of admissible control laws, will be further restricted: it contains the previously defined control laws u, such that

$$E_o\left\{L(u)\left[\int_I r_o^s \cdot c(s,u(s))dA^u(s) + J_f\right]\right\} < \infty$$

We assume that this new class $\mathcal{U}$ is non-empty (otherwise the problem is trivial $J(u) = \infty, \forall u$). Then for all $u,v \in \mathcal{U}$ the following processes are well-defined and integrable:

$$\phi(u,v,t) = E_o\left[L(u,t,v)\left(\int_t^{t_f} r_t^s \cdot c(s,v)dA^v(s) + r_f^{t_f} \cdot J_f\right) \middle| \mathcal{F}_t^y\right]$$

$$\in L^1(\Omega, \mathcal{F}_t^y, \mathcal{P}_o) \tag{3}$$

$$\psi(u,v,t) = E_{u,t,v}\left[\int_t^{t_f} r_t^s \cdot c(s,v)dA^v(s) + r_t^{t_f} \cdot J_f \middle| \mathcal{F}_t^y\right]$$

$$\in L^1(\Omega, \mathcal{F}_t^y, \mathcal{P}_u) \tag{4}$$

related by

$$\psi(u,v,t) = \frac{\phi(u,v,t)}{E_o\left[L_o^t(u) \middle| \mathcal{F}_t^y\right]}$$

Remark 1: It is implicitly assumed here that $\mathcal{P}_u$, restricted to $\mathcal{F}_t$, depends only on the values of the control law on the interval $[0,t]$, Hence $E_o[L(u,t,v)|\mathcal{F}_t^y] = E_o[L^t{}_o(u)|\mathcal{F}_t^y] = E_o[L(u)|\mathcal{F}_t^y]$. Therefore the notation $L_1(\ ,\mathcal{F}_t^y,\mathcal{P}_u)$ makes sense. This result will be used repeatedly in the following definitions, and in lemma 2.1 and theorem 2.1.

The value of $\phi(u,v,t)$ (resp. $\psi(u,v,t)$ is the expected unnormalized (resp. normalized) future cost, as evaluated at time t, when control law u is used up to time t, and control law v is used thereafter. To evaluate the expected cost after time t, evaluated at 0, given the information at t, one multiplies $\phi(u,v,t)$ (resp. $\psi(u,v,t)$ by r_o^t. Since $L_1(\Omega,\mathcal{F}_t^y,\mathcal{P}_o)$ is a complete lattice with the natural partial ordering for real-valued functions ([11], IV-8-22), the following infima exist:

$$v(u,t) = \bigwedge_{v\in\mathcal{U}_t} \phi(u,v,t) \in L_1(\Omega,\mathcal{F}_t^y,\mathcal{P}_o) \tag{5}$$

$$w(u,t) = \bigwedge_{v\in\mathcal{U}_t} \psi(u,v,t) = \frac{v(u,t)}{E_o[L_o^t(u)|\mathcal{F}_t^y]} \in L_1(\Omega,\mathcal{F}_t^y,\mathcal{P}_u) \tag{6}$$

Remark 2: $\mathcal{U}_t$ is obtained from $\mathcal{U}$ by restricting the domain I, in the definition of the control laws, to $[t,t_f]$.

V(u,t) is the <u>unnormalized value function</u>, while W(u,t) is the <u>normalized value function</u>. They represent the lowest possible cost after the present time t, given the present information, and depending on the control law u that has been used in the past.

Definition (see [8]). The class $\mathcal{U}$ is called <u>$\mathcal{F}_t^y$-relatively complete</u> with respect to the unnormalized (resp. normalized) value function if for all $u \in \mathcal{U}$, all $t \in I$, all $\varepsilon > 0$, there exists a $v \in \mathcal{U}_t$ such that:

$$\phi(u,v,t) \leq V(u,t) + \varepsilon \quad \text{a.s.}(\mathcal{P}_o)$$

(resp. $\psi(u,v,t) \leq W(u,t) + \varepsilon$ a.s.$(\mathcal{P}_u)$

Lemma 2.1: $\mathcal{U}$ is $\mathcal{F}_t^y$-relatively complete w.r.t. both $V(u,t)$ and $W(w,t)$, under the previously made assumptions.

proof: For $V(u,t)$ see Davis-Varaiya [8], lemma 3.1

For $W(u,t)$ the same proof can be used, since remark 1 above implies that for any $M_{v_\alpha} \in \mathcal{F}_t^y$, $\mathcal{P}_{u,t,v_\alpha}(M_{v_\alpha}) = \mathcal{P}_u(M_{v_\alpha})$

¤

The following theorem gives necessary and sufficient conditions for optimality, assuming $V(u,t)$ (resp. $W(u,t)$) is known. The result is in the form of a "dynamic programming" principle of optimality. Checking the conditions of the theorem directly however requires a minimization $\bigwedge_{v \in \mathcal{U}_t} \phi(u,v,t)$, as difficult as the original problem. However the theorem can be used to prove better criteria later on.

Theorem 2.1: (Principle of optimality)

a) For all $t \in I$, all $h \geq 0$ (s.t. $t + h \in I$), $\forall u \in \mathcal{U}$:

$$W(u,t) \leq E_u\left(\int_t^{t+h} r_t^s \cdot c(s,u(s))dA^u(s) \middle| \mathcal{F}_t^y\right) + E_u\left(r_t^{t+h} \cdot W(u,t+h) \middle| \mathcal{F}_t^y\right) \quad (7a)$$

$$W(u,t_f) = E_u(J_f | \mathcal{F}_{t_f}^y) \quad (7b)$$

A control law $u \in \mathcal{U}$ is optimal if and only if equality holds in (7a)

b) The same statements hold, if (7a), (8a) are replaced by:

$$V(u,t) \leq E_o\left[L(u)\int_t^{t+h} r_t^s \cdot c(s,u(s))dA^u(s)\,|\mathcal{F}_t^y\right]$$

$$+ E_o[r_t^{t+h}\cdot V(u,t+h)\,|\mathcal{F}_t^y] \tag{8a}$$

$$V(u,t_f) = E_o[L(u)J_f\,|\mathcal{F}_{t_f}^y] \tag{8b}$$

<u>proof</u>: The equalities (8a), (8b) are obvious from the definitions (3) - (6). The rest of the proof will be given in 4 steps.

i) (7a) is derived as follows: using (4) and (6)

$$W(u,t) \leq E_u\left[\int_t^{t+h} r_t^s \cdot c(s,u(s))dA^u(s)\,|\mathcal{F}_t^y\right]$$

$$+ \bigwedge_{v\in\mathcal{U}_{t+h}} E_{u,t+h,v}[r_t^{t+h}\psi(u,v,t+h)\,|\mathcal{F}_t^y]$$

It remains to prove that the infimum and conditional expectation operators can be interchanged, i.e.

$$\bigwedge_{v\in\mathcal{U}_{t+h}} E_u[r_t^{t+h}\psi(u,v,t+h)\,|\mathcal{F}_t^y] = E_u[\bigwedge_{v\in\mathcal{U}_{t+h}} r_t^{t+h}\psi(u,v,t+h)\,|\mathcal{F}_t^y] \tag{9}$$

The inequality $\geq$ in (9) is obvious since for all $v \in \mathcal{U}_{t+h}$:

$$: E_u[r_t^{t+h}\psi(u,v,t+h)\,|\mathcal{F}_t^y] - E_u[\bigwedge_{v\in\mathcal{U}_{t+h}} r_t^{t+h}\psi(u,v,t+h)\,|\mathcal{F}_t^y]$$

The inequality $\leq$ follows from the relative completeness, i.e. $\forall\varepsilon > 0$ $\exists\ v^\varepsilon \in \mathcal{U}_{t+h}$ such that:

$$\psi(u,v^\varepsilon,t+h) \leq \bigwedge_{v\in\mathcal{U}_{t+h}} \psi(u,v,t+h) + \varepsilon$$

or $\bigwedge_{\varepsilon>0} E_u[r_t^{t+h}\psi(u,v^\varepsilon,t+h)\,|\mathcal{F}_t^y] \leq E_u[r_t^{t+h} \bigwedge_{v\in\mathcal{U}_{t+h}} \psi(u,v,t+h)\,|\mathcal{F}_t^y] + \varepsilon$

The proof of (7b) is analogous.

ii) We now prove that $u \in \mathcal{U}$ is optimal if and only if for all

$t \in I : V(u,t) = \phi(u,u,t)$ a.s. $(\mathcal{P}_o)$

(resp. $W(u,t) = \psi(u,u,t)$ a.s. $(\mathcal{P}_u)$)

If equality holds, then applying it at $t = 0$ shows optimality.

Conversely, if u is optimal, then applying definition (6) gives

$$W(u,0) = J^* = E_u\left[\int_0^t r_o^s \cdot c(s,u(o))dA^u(s)\right] + E_u[\psi(u,u,t)]$$

while applying (7a) for $t = 0$, $h = t$, gives

$$W(u,0) \leq E_u\left[\int_0^t r_o^s . c(s,u(s))dA^u(s)\right] + E_u[W(u,t)]$$

Substracting gives $E_u[\psi(u,u,t) - W(u,t)] \leq 0$

From definition (6) $\psi(u,u,t) - W(u,t)$ is a positive random variable,

hence $\psi(u,u,t) = W(u,t)$ a.s. $(\mathcal{P}_u)$

iii) If u is optimal then equality in (7b) follows from ii) as

follows:

$$V(u,t) = \phi(u,u,t)$$

$$= E_o\left[L(u)\int_t^{t+h} r_t^s \cdot c(s,u(s))dA^u(s)\,\mathcal{F}_t^y\right] + E_o[r_t^{t+h}\phi(u,u,t+h)|\mathcal{F}_t^y]$$

$$= E_o\left[L_o^{t+h}(u)\int_t^{t+h} r_t^s \cdot c(s,u(s))dA^u(s)|\mathcal{F}_t^y\right]$$

$$+ E_o[r_t^{t+h} \cdot V(u,t+h)|\mathcal{F}_t^y]$$

The proof of (7a) is even easier.

iv) If equality holds in (7b) then take $t = 0$, $h = t_f$, to get

$$V(u,0) = J^* = E_o\left[L(u)\left(\int_I r_o^s \cdot c(s,u(s))dA^u(s) + V(u,t_f)\right)\right]$$

$$= E_o\left[L(u)\left(\int_I r_o^s \cdot c(s,u(s))dA^u(s) + r_o^{t_f} J_f\right)\right]$$

by (8b). Hence $\phi(u,u,0) = J^*$ and u is optimal. The normalized value function can be used in the same way. ⌑

Corollary 2.1. For all $u \in \mathcal{U}$ the process

$$r_o^t \cdot V(u,t) + E_o\left[L(u)\int_0^t r_o^s \cdot c(s,u(s))dA^u(s)\,|\,\mathcal{F}_t^y\right] \tag{10a}$$

$$\text{(resp. } r_o^t \cdot W(u,t) + E_u\left[\int_0^t r_o^s \cdot c(s,u(s))dA^u(s)\,|\,\mathcal{F}_t^y\right] \tag{10b}$$

is a $\mathcal{P}_o$ (resp. $\mathcal{P}_u$) uniformly integrable sub-martingale. Any $u \in \mathcal{U}$ is optimal if and only if the processes defined in (10a), (10b) are martingales.

proof: Add $E_o\left[L(u)\int_0^t r_o^s \cdot c(s,u(s))dA^u(s)\,|\,\mathcal{F}_t^y\right]$

(resp. $E_u\left[\int_0^t r_o^s \cdot c(s,u(s))dA^u(s)\,|\,\mathcal{F}_t^y\right]$ to both sides of (7a) (resp (7b)), after multiplying by r_o^t. ⌑

Since $E_o\left[L(u)\int_I r_o^s \cdot c(s,u(s))dA^u(s) + r_o^{t_f} \cdot J_f\,|\,\mathcal{F}_t^y\right] \in \mathcal{M}^1(\mathcal{F}_t^y, \mathcal{P}_o)$

(resp. $E_u\left[\int_I r_o^s c(s,u(s))dA^u(s) + r_o^{t_f} J_f\,|\,\mathcal{F}_t^y\right] \in \mathcal{M}^1(\mathcal{F}_t^y, \mathcal{P}_o)$

the following processes will be supermartingales:

$$v(u,t) = E_o\left[L(u)\left(\int_t^{t_f} r_o^s\cdot c(s,u(s))dA^u(s) + r_o^{t_f}J_f\right)\Big|\mathcal{F}_t^y\right] - r_o^t\cdot V(u,t) \tag{11a}$$

$$= r_o^t(\phi(u,u,t) - V(u,t))\ .$$

$$w(u,t) = E_u\left[\int_t^{t_f} r_o^s\cdot c(s,u(s))dA^u(s) + r_o^{t_f}J_f\Big|\mathcal{F}_t^y\right] - r_o^t\cdot W(u,t) \tag{11b}$$

$$= r_o^t(\psi(u,u,t) - W(u,t))$$

Corollary 2.2: For all $u \in \mathcal{U}$, the process $v(u,t)$ (resp. $w(u,t)$) is a potential with respect to the σ-fields $\mathcal{F}_t^y$, under the measure $\mathcal{P}_o$ (resp. $\mathcal{P}_u$). $u^* \in \mathcal{U}$ is optimal if and only if $v(u^*,t) \equiv 0$ a.s. ($\mathcal{P}_o$) resp. $w(u^*,t) \equiv 0$ a.s. ($\mathcal{P}_u$)

Remarks 3: Theorem 2.1 summarizes and extends the results of Davis-Varaiya [8, theorem 3.1 and 4.1].

4: Corollary 2.2 is implicit in the methods of proof used in Kushner [20], for Markov processes. Also for a Markov process, and by Meyer's supermartingale decomposition theorem probably for general processes, the potentials $v(u,t)$ and $w(u,t)$ could be used as Lyapunov functions to prove stability, if the assumption at the beginning of this paragraph had not been made. This extension, to allow control laws u that make the process unstable, has not been pursued here.

5. The submartingales defined in (10a) and (10b) have an interesting heuristic interpretation. Their value is the expected total cost, evaluated with the information $\mathcal{F}_t^y$ available at time t, given control law u was used up to t, and an optimal control law is used afterwards. The expected value will increase if the non-optimal

control law u is used for a longer time, explaining the submartingale. However if u is optimal this expected value will remain constant on the average, leading to the martingale property.

6: The processes $v(u,t) = r_o^t(\phi(u,u,t) - V(u,t))$

and $w(u,t) = r_o^t(\psi(u,u,t) - W(u,t))$

express the loss incurred by using u after time t, compared with the optimal control.

7: Theorem 2.1 can be rederived from corollary 2.1. Hence, the optional sampling theorem implies that t in (7a), (7b) can be replaced by any $\mathcal{F}_t^y$-stopping time.

Since the instantaneous cost c(t,u) is always non-negative, it follows that for an optimal control law u, V(u,t) (resp. W(u,t)) should decrease in expected value, i.e. be a supermartingale. This also follows from (7a), (7b) with equality. This suggests the following definition, first used by Davis-Varaiya [8].

<u>Definition</u>: A control law $u \in \mathcal{U}$ is <u>value-decreasing</u> if $r_o^t \cdot V(u,t)$ is an $\mathcal{F}_t^y$-supermartingale under $\mathcal{P}_o$. Then $r_o^t \cdot W(u,t)$ is an $\mathcal{F}_t^y$-supermartingale under $\mathcal{P}_u$.

To make the optimization problem non-trivial we reduce the class of admissible control laws further to include value-decreasing control laws only. If there exists an optimal control u*, it will be in $\mathcal{U}$ still.

By the <u>supermartingale decomposition theorem</u> of Meyer ([24],VII T 29) we therefore can assure the existence of a predictable increasing, uniformly integrable process $\Lambda_o^t V(u) \in \mathcal{A}^+(\mathcal{F}_t^y,\mathcal{P}_o)$ (resp. $\Lambda_o^t W(u) \in \mathcal{A}^+(\mathcal{F}_t^y,\mathcal{P}_u)$

$$r_o^t V(u,t) = J_o - \Lambda_o^t V(u) + m_V^u(t) \tag{12a}$$

resp. $$r_o^t W(u,t) = J_o - \Lambda_o^t W(u) + m_W^u(t) \tag{12b}$$

Then (7a), (7b) can be rewritten as:

$$E_o[\Lambda_o^{t+h} V(u) - \Lambda_o^t V(u) | \mathcal{F}_t^y] \leq E_o\left[L(u) \int_t^{t+h} r_o^s \cdot c(s,u(s)) dA^u(s) | \mathcal{F}_t^y\right] \tag{13a}$$

$$E_u[\Lambda_o^{t+h} W(u) - \Lambda_o^t W(u) | \mathcal{F}_t^y] \leq E_u\left[\int_t^{t+h} r_o^s \cdot c(s,u(s)) dA^u(s) | \mathcal{F}_t^y\right] \tag{13b}$$

Obviously equality holds if and only if $u \in \mathcal{U}$ is optimal.

To simplify the notation, for an increasing process Λ_o^t, we denote $\Lambda_t^s = \Lambda_o^s - \Lambda_o^t$. This corresponds to considering a positive measure $d\,{}_o^t$ instead.

The following <u>necessary and sufficient condition</u> for optimality does no longer require the advance knowledge of the value function, and is therefore easier to check than theorem 2.1. It is similar to the Hamilton-Jacobi equation in deterministic optimal control.

<u>Theorem 2.2</u>: a) $u^* \in \mathcal{U}$ is optimal if and only if there exist

i) a constant J_o

ii) a process $\Lambda_o^t(u) \in \mathcal{A}^+(\mathcal{F}_t^y, \mathcal{P}_o)$

such that 1) $E\,\Lambda_o^{t_f}(u) = J_o - E_o(L(u) r_o^{t_f} J_f) \quad \forall u \in \mathcal{U}$ (13a)

2) for all $t \in I$, all $h > 0$ (s.t. $t + h \in I$)

$$E_o\left[- \Lambda_t^{t+h}(u) + L(u) \int_t^{t+h} r_o^s \cdot c(s,u(s)) dA^u(s) | \mathcal{F}_t^y\right] \geq 0 \tag{14a}$$

with equality holding for u^*

Then $J_o = J^* = J(u^*)$, the cost of the optimal control law, and

$$r_o^t \cdot V(u^*,t) = E_o(\Lambda_t^{t_f}(u^*) + L(u^*)J_f | \mathcal{F}_t^y) \tag{15a}$$

b) The same statements hold if $\Lambda_o^t(u)$ is replaced by $\Lambda'^t_o(u) \in \mathcal{A}^+(\mathcal{F}_t^y, \mathcal{P}_u)$ and (13a), (14a), (15a) by

1) $$E_u \Lambda'^{t_f}_o(u) = J_o - E_u(r_o^{t_f} J_f) \qquad \forall u \in \mathcal{U} \tag{13b}$$

2) $$E_u\left[-\Lambda'^{t+h}_t(u) + \int_t^{t+h} r_o^s \cdot c(s,u(s))dA^u(s) \middle| \mathcal{F}_t^y\right] \geq 0 \tag{14b}$$

and $$r_o^t \cdot W(u^*,t) = E_{u^*}(\Lambda'^{t_f}_t(u^*) + r_o^{t_f} J_f | \mathcal{F}_t^y) \tag{15b}$$

<u>proof</u>: The necessity part follows from equations (12) and (13) by putting $\Lambda_o^t(u) = \Lambda_o^t V(u)$. To prove sufficiency consider the uniformly integrable process:

$$Z(u,t) = E_o(\Lambda_t^{t_f}(u) | \mathcal{F}_t^y) + E_o(L(u) r_o^{t_f} J_f | \mathcal{F}_t^y) \tag{16}$$

and compare with

$$r_o^t \phi(u,u,t) = E_o\left[L(u)\left(\int_t^{t_f} r_o^s \cdot c(s,u(s))dA^u(s) + r_o^{t_f} J_f\right) \middle| \mathcal{F}_t^y\right] \tag{17}$$

Substracting (16) from (17) gives

$$r_o^t \cdot \phi(u,u,t) - Z(u,t) = E_o\left[-\Lambda_t^{t_f}(u) + L(u)\int_t^{t_f} r_o^s \cdot c(s,u(s))dA^u(s) \middle| \mathcal{F}_t^y\right]$$

$r_o^t \cdot \phi(u,u,t) - Z(u,t) \geq 0$ for all $u \in \mathcal{U}$ by (14) and

$r_o^t \cdot \phi(u^*,u^*,t) - Z(u^*,t) = 0$

In particular $\phi(u,u,0) \geq Z(u,0) = J_o = Z(u^*,0) = \phi(u^*,u^*,0)$

This proves that u* is optimal, and

$$J^* = J(u^*) = \phi(u^*,u^*,0) = J_o.$$

Moreover: $Z(u^*,t) = \phi(u^*,u^*,t) = W(u^*,t)$ by ii) in the proof of theorem 2.1. ¤

Remark 8: Even if the normalized value function $W(u,t)$ were independent of u (which is the case for a Markov process and for the complete information case with an additional constraint on $L(u)$) the equation $\inf_{u \in \mathcal{U}_t} E_u[-\Lambda_t'^{t+h}W + \int_t^{t+h} r_o^s \cdot c(s,u(s))dA^u(s) | \mathcal{F}_t^y]$ is still dependent on u, since $\Lambda_t'^{t+h}W$ depends on u through the measure $\mathcal{P}_u$. Hence it cannot be solved by the usual technique for Hamilton-Jacobi equations. Better results will be obtained for jump processes, in the next chapter.

2.3. Local optimality conditions

Some additional assumptions will make it possible to derive in this paragraph, results corresponding with $h \downarrow 0$ in theorem 2.2. The results will be somewhat simpler.

The following lemma is very useful. It shows the existence for some supermartingales, of an operator similar to the differential generator for Markov processes, as studied by Dynkin [12]. It has previously been used in filtering problems (see [5] and [16]).

Lemma 2.2. Let f_t be a supermartingale w.r.t. the increasing family of σ-fields $\mathcal{G}_t$, on a probability space $(\Omega, g, \mathcal{P})$. Let $A^g_{t,h}f = \frac{1}{h}[f_t - E(f_{t+h}|\mathcal{G}_t)]$. Suppose that for all n and for some $h_o > 0$ there exists a constant K_n such that $|A^g_{t,h}f| \leq K_n$, $\forall h \in [0,h_o]$

$\forall t \in [o,n]$. Then $A_t^g f = w \lim_{h\to 0} A_{t,h}^g f$ exists, is integrable and $\mathcal{G}_t$-adapted.

<u>proof</u>: First, let $I = [o,n]$. Then $\{A_{t,h}^g f \mid h \in [o,h_o]\}$ is a weakly compact subset of $L_1(\Omega \times I, \mathcal{G} \otimes \mathcal{B}_I, \mathcal{P} \otimes \lambda)$ (λ = Lebesgues measure) because of the uniform boundedness ([11], IV-8-9 and 12). Hence there exists a weakly convergent ($\sigma(L_1-L_\infty)$ sense) subsequence. Since $E(A_{t,h}^g f \mid g_t)$ decreases, this weak limit must be unique, and all convergent subsequences tend a.s. to the same limit $A_t^g f$. For $I = [0,\infty)$, first construct limits on each finite interval $[o,n]$. By uniqueness these limits can be extended to $[0,\infty)$. ¤

<u>Remark 1</u>: This proof is taken from Davis-Varaiya ([8], lemma 4.2).

The following <u>assumptions</u> are now made

a) $A^u(t) = t$ (or absolutely continuous w.r.t. Lebesgues measure).

b) $c(t,u,\omega) \in [0,K]$ $\forall t \in I$, $\forall u \in U$, $\forall \omega \in \Omega$

c) $r_s^t(\omega) \in [0,1]$ $\forall s,t \in I$, $\forall \omega \in \Omega$.

$$\text{Then } 0 \le \frac{r_o^t}{h} \; V(u,t) - E_o(r_t^{t+h} V(u,t+h) \mid \mathcal{F}_t^y \le \frac{1}{h} E_o\left[L(u)\int_t^{t+h} r_o^s \cdot c(s,u(s))ds \mid \mathcal{F}_t^y\right]$$

$$\le K \cdot E_o[L_o^t(u) \mid \mathcal{F}_t^y]$$

$$\text{and } \; 0 \le \frac{r_o^t}{h} \; W(u,t) - E_u(r_t^{t+h} W(u,t+h) \mid \mathcal{F}_t^y] \le \frac{1}{h}\left[E_u \int_t^{t+h} r_o^s \cdot c(s,u(s)) \cdot ds \mid \quad {}_t^y\right]$$

$$\le K$$

Lemma 2.2 is immediately applicable to the normalized value function $W(u,t)$. For the unnormalized case, the following stopping times, together with uniqueness, will prove the existence of a limit

$A_t^y(r_o^{\cdot}V(u,\cdot))$ (where A_t^y stands for $A_t^{\mathcal{F}^y}$). Let τ_n be the $\mathcal{F}_t^y$-stopping time: $\tau_n = \inf_{t\in I} \{E_o(L_o^t(u)|\mathcal{F}_t^y) \geq n\}$

By the uniform integrability of $\{L_o^t(u)\}$, it is clear that $\tau_n \to \infty$ w.p.1.

By the proof of Meyer's decomposition theorem ([24], VII- T28 and 29)

$$
{}_o^tV(u) = w\lim_{h\to 0} \int_0^t \frac{1}{h}[r_o^sV(u,s) - E_o(r_o^{s+h}V(u,s+h)|\mathcal{F}_s^y)]ds
$$

$$
= \int_0^t A_s^y(r_oV(u,\cdot))ds \tag{18a}
$$

where the interchange of weak limits (in $\sigma(L_1,L_\infty)$ topology) is justified by the boundedness assumption b) and c) above and by an argument given by Davis-Varaiya ([8], lemma 4.2). By the previous arguments $A_t^y(r_oV(u,\cdot))$ is integrable over $(\Omega \times I, \mathcal{P}_o \otimes \lambda)$ similarly.

$$
\Lambda\, {}_o^tW(u) = \int_0^t A_s^y(r_o^{\cdot}W(u,\cdot))ds \tag{18b}
$$

Theorem 2.2 can then be rewritten as:

Theorem 2.3: a) $u^* \in \mathcal{U}$ is optimal if and only if there exist

i) a constant J_o

ii) for each $u \in \mathcal{U}$ (value decreasing) a positive process $\lambda_t(u)$, $\mathcal{F}_t^y$-adapted

such that 1) $\int_0^t \lambda_s(u)\,ds$ is uniformly integrable

2) $E_o(\int_0^{t_f} \lambda_s(u)ds) = J_o - E_o(L(u)r_o^{t_f}J_f)$ (19a)

3) for all $t \in I$, all $u \in \mathcal{U}$

$$- \lambda_t(u) + E_o(L_o^t(u) r_o^t \cdot c(t,u(t)) | \mathcal{F}_t^y) \geq 0 \tag{20a}$$

with equality holding for u*.

b) The normalized result is completely equivalent.

proof: From theorem 2.2 and the definition of weak limits one has for all positive, bounded, $\mathcal{F}_t^y$-measurable random variables .

$$\lim_{h\to 0} E_o\left[\frac{1}{h} \int_t^{t+h} (-\lambda_s(u) + L(u) r_o^s \cdot c(s,u(s)))ds \;|\mathcal{F}_t^y\right] = 0.$$

This is true if and only if (20a) is satisfied ¤

Remark 2: The fixed time t in the previous theorem can be replaced by any $\mathcal{F}_t^y$-stopping time τ: use remark 7 in §2 and the fact that Meyer's decomposition theorem for supermartingales ([24],VII T29) is stated for any stopping time.

3: This optimality criterion has been derived before for Markov processes (see Kushner [20]),for conditional Markov processes by Stratonovich [30], and for processes on a Wiener space by Davis-Varaiya [8].

2.4. Markov controls.

In this section we will prove the intuitively obvious fact that if $(x_t, \mathcal{F}_t^x, \mathcal{P}_o)$ is a Markov process, then a control depending on the present value x_t only, should be optimal in the class of complete information control laws. A first problem is that a control law $u(t,\omega)$ is assumed to be predictable, and can therefore not depend on x_t (unless it is left-continuous). Therefore we consider control laws $v(t, x_{t-}(\omega)) = u(t,\omega)$ and assume $(x_{t-}, \mathcal{F}_t^x, \mathcal{P}_o)$ to be Markovian.

If $\{\mathcal{F}_t^x\}$ is quasi-left continuous, then $(x_t,\mathcal{F}_t^x,\mathcal{P}_o)$ being Markovian implies $(x_{t-},\mathcal{F}_t^x,\mathcal{P}_o)$ being Markovian (note: $\mathcal{F}_t^x = \sigma(x_s, s \leq t) = \sigma(x_{s-}, s \leq t)$), for sufficiently smooth transition functions. We now make the following <u>assumptions</u>:

i) $(x_t,\mathcal{F}_t^x,\mathcal{P}_o)$ is a Markov process.

ii) $\mathcal{F}_{t-}^x = \mathcal{F}_t^x = \sigma(x_{s-}, s \leq t) = \sigma(x_s, s \leq t)\ \forall t \in I$.

iii) Let $L_s^{t-}(u) = \dfrac{L_o^{t-}(u)}{L_o^{s-}(u)}$, $s \leq t$ and assume $L_s^{t-}(u)$ depends only on the values of the control law on $[s,t)$.

iv) $0 \leq c(t,u,\omega) = \tilde{c}(t,y,x_{t-}(\omega)) \leq K$ for some K and let J_f be a bounded function of x_{t_f}.

v) for fixed $u \in \mathcal{M}$ and $t \in I$, $(L_s^t(u),\mathcal{F}_s^x,\mathcal{P}_o)$ is Markovian (i.e. $L_s^t(u)$ depends only on $\{x_\tau, s \leq \tau < t\}$). Here $\mathcal{M} = \mathcal{U} \cap \{u(t,\omega) = v(t,x_{t-}(\omega))\}$ = class of <u>Markov controls</u>. Assume U is a separable metric space.

vi) $\mathcal{P}_o(x_1,t_1;dx_2,t_2) \xrightarrow{t_1 \to t_2} \delta_{x_1 x_2}$ w.p. 1 for all t. (i.e. $x_t = x_{t-}$ w.p. 1 for all t).

<u>Lemma 2.3</u>. Under the previous assumptions $(x_{t-},\mathcal{F}_t^x,\mathcal{P}_o)$, $(x_t,\mathcal{F}_t^x,\mathcal{P}_u)$, $(x_{t-},\mathcal{F}_t^x,\mathcal{P}_u)$ are all Markov processes (for any $u \in \mathcal{M}$).

<u>proof</u>. Let $f : X \to R$ be any bounded continuous function, let $s < t$, then

$$E_o(f(x_{t-})|\mathcal{F}_s^x) = E_o(f(x_t)|\mathcal{F}_s^x) = E_o(f(x_t)|x_s) \tag{21}$$

and

$$E_o(f(x_{t-})|x_{s-\varepsilon}) = \int_X E_o(f(x_{t-})|x_s)\cdot P(x_{s-\varepsilon},s-\varepsilon;dx_s,s)$$

which for $\varepsilon \downarrow 0$ gives :

$$E_o(f(x_{t-})|x_{s-}) = E_o(f(x_{t-})|x_s) \tag{22}$$

(by assumption vi)). Putting (21) and (22) together:

$$E_o(f(x_{t-})|\mathcal{F}_s^x) = E_o(f(x_{t-})|x_{s-}).$$

By the monotone class theorem, this result can be extended to any measurable functional on $\{x_{\tau-}, s < \tau\}$, which implies $(x_{t-}, \mathcal{F}_t^x, \mathcal{P}_o)$ is a Markov process (see Meyer [25]).

For the 2nd part, consider ($s < t$ again):

$$E_u(f(x_t)|\mathcal{F}_s^x) = \frac{E_o[L(u)f(x_t)|\mathcal{F}_s^x]}{E_o[L(u)|\mathcal{F}_s^x]} \quad \text{(see Loève, [22], p.344)}$$

$$= E_o[L_s^t(u)f(x_t)|\mathcal{F}_s^x]$$

$$= E_o[L_s^t(u)f(x_t)|x_s] \quad \text{since } v \in \mathcal{U} \text{ and } E_o[L_f^{t_f}(u)|\mathcal{F}_s^x] = 1$$

$$= \frac{E_o[L_s^{t_f}(u)f(x_t)|x_s]}{E_o[L_s^{t_f}(u)|x_s]} \cdot \frac{E_o[L_o^s(u)|x_s]}{E_o[L_o^s(u)|x_s]}$$

$$= \frac{E_o[L(u)f(x_t)|x_s]}{E_o[L(u)\ x_s]} \quad \text{(past and future are conditionally independent)}$$

$$= E_u[f(x_t)|x_s]$$

By the monotone class theorem this is sufficient to show that $(x_t, \mathcal{F}_t^x, \mathcal{P}_u)$ is a Markov process. ¤

Using the previous lemma, it becomes obvious that:

$$\psi(u,v,t,\omega) = E_{uv}\left[\int_t^{t_f} r_t^s \cdot c(s,v(s))dA_s^v + J_f \middle| \mathcal{F}_t^x\right]$$

$$= E_v\left[\int_t^{t_f} r_t^s \cdot c(s,v(s))dA_s^v + J_f \middle| x_{t-}\right]$$

$$= \psi'(t,v,x_{t-}(\omega)) \quad \text{for } v \in \mathcal{M}$$

and $U(t,x_{t-}(\omega)) = \bigwedge_{v\in\mathcal{M}} \psi'(t,v,x_{t-}(\omega))$, the <u>Markovian value function</u> is independent of u, the control law on [o,t).

We also note that, if assumption iii) above is satisfied, then for any control $u(t,\omega)$ which is $\mathcal{F}_{t-}^x$ predictable, the <u>complete information value function</u>:

$$W(t,\omega) = \bigwedge_{v\in\mathcal{U}} E_{uv}\left[\int_t^{t_f} r_t^s \cdot c(s,v(s))dA_s^v + J_f \middle| \mathcal{F}_t^x\right]$$

$$= \bigwedge_{v\in\mathcal{U}} E_o\left[L_t^{t_f}(v)\left(\int_t^{t_f} r_t^s \cdot c(s,v(s))dA_s^v + J_f\right) \middle| \mathcal{F}_t^x\right] \tag{25}$$

is also independent of the past control law u(on [o,t]). To prove the statment that Markov controls are optimal, one has to show $W(t,\omega) = U(t,x_{t-})$ for all t (see ii) in proof of principle of optimality). To prove this, we need a principle of optimality for the Markovian value function. Since lemma 2.1 requires increasing σ-fields, it is not applicable here. Therefore a limiting argument, starting with a discrete backward dynamic programming, will be used.

First we introduce the concept of a discrete Markov control: $u \in \mathcal{M}^d$, the class of <u>discrete Markov controls</u>, if $u(t,\omega)$ is constant except for jumps at $t_o = 0, t_1,\ldots,t_n = t_f$, and $u(t,\omega)$ is a function of $x_{t_{\bar{i}}}$ on the interval $[t_i, t_{i+1})$. Note that the t_i's are constant times, not stopping times.

The following additional assumptions are now made:

vii) $c(t,u,x)$ is uniformly continuous in $u \in U$ on $I \times X$ and A_s^u is independent of u.

viii) $\{L(u) : u \in \mathcal{M}^d\}$ is dense in $\{L(u) : u \in \mathcal{M}\}$ in the following sense: for each $u \in \mathcal{M}$ there exists a sequence of $u_n \in \mathcal{M}^d$ such that

a) $u_n(t,\omega) \to u(t,\omega)$ pointwise

b) $L(u_n) \overset{W}{\to} L(u)$ (weak convergence in L_1).

<u>Lemma 2.4.</u> If assumptions i) to viii) are satisfied, then for all $u \in \mathcal{M}$ there exists a $v \in \mathcal{M}^d$ such that: $J(v) \leq J(u) + \varepsilon$ (where $I = [0,1]$).

<u>proof</u>: Let $u_n \in \mathcal{M}^d$ such that $u_n \to u$ pointwise and $L(u_n) \overset{W}{\to} L(u)$. Then:

$$J(u_n) - J(u) = E_{u_n}\left(\int_0^1 r_o^s c(s,u_n(s))dA_s + J_f\right)$$

$$- E_u\left(\int_0^1 r_o^s c(s,u(s))dA_s + J_f\right)$$

$$= E_o\left[L(u_n)\left(\int_0^1 r_o^s \cdot c(s,u_n(s))dA_s + J_f\right)\right]$$

$$- E_o\left[L(u)\left(\int_0^1 r_o^s \cdot c(s,u(s))dA_s + J_f\right)\right]$$

$$= E_o\left[L(u_n)\left(\int_0^1 r_o^s(c(s,u_n(s)) - c(s,u(s))dA_s + J_f\right)\right]$$

$$+ E_o[(L(u_n) - L(u))(\int_0^1 r_o^s\cdot c(s,u(s))dA_s + J_f)\,|$$

$$\to 0 \quad \text{as } n \to \infty \qquad \square$$

Theorem 2.4. Under assumptions i) to viii), and for all $u \in \mathcal{M}$ for all $t \in I$, all $h > 0$ (s.t. $t + h \in I$)(where $I = [0,1]$)

$$U(t,x_{t-}) \leq E_u[\int_t^{t+h} r_t^s\cdot c(s,u(s))dA_s|x_{t-}] + E_u[r_t^{t+h}U(t+h,x_{t+h}-)|x_{t-}] \tag{24}$$

$$U(t_f,x_{t_f-}) = J_f \tag{25}$$

$u \in \mathcal{M}$ is optimal if and only if equality holds in (24).

proof: The only part of the proof of theorem 2.1 that does not carry over immediately is the fact that:

$$\bigwedge_{v\in\mathcal{M}_{t+h}} E_{u,v}[r_t^{t+h}\psi(u,v,t+h)|x_{t-}]$$

$$= E_u[\bigwedge_{v\in\mathcal{M}_{t+h}} r_t^{t+h}\psi(u,v,t+h)|x_{t-}]$$

$$= E_u[r_t^{t+h}U(t+h,x_{t+h-})|x_{t-}]$$

Let $\mathcal{M}^n$ be the subclass of $\mathcal{M}^d$, containing all discrete control laws having no more than n jumps. Then $U^n(t,x_{t-}) = \bigwedge_{v\in\mathcal{M}_t^n} \psi(u,v,t)$ and by lemma 2.4: $U^n(t,x_{t-}) \downarrow U(t,x_{t-})$ as n tends to infinity for $I = [o,n]$, any bounded interval. (Note that in the proof of lemma 2.4, no use is made of the fact that the initial state is fixed). We have

$$E_{ut+h,v}[r_t^{t+h}\cdot\psi(u,v,t+h)|x_{t-}]$$

$$\leq E_u[r_t^{t+h}U^n\cdot(t+h,x_{t+h}-)|x_{t-}] + \varepsilon$$

since this operation involves only n minimizations and each infinum can be approximated up to ε/n. This argument can be repeated for each n, with fixed ε, so that there exists $v \in \mathcal{M}^n_{t+h}$ such that

$$E_{u,t+h,v}[r_t^{t+h}\cdot\psi(u,v,t+h)|x_{t-}] \leq E_u[r_t^{t+h}U(t+h,x_{t+h-})|x_{t-}] + 2\varepsilon$$

by taking $n > N$ such that:

$$U^n(t+h,x_{t+h-}) - U(t+h,x_{t+h-}) \leq \varepsilon.$$

The rest of the proof proceeds as in theorem 2.1.

If $I = [0,\infty)$, then there has to be a T such that

$E_u[\int_T^\infty r_o^t\cdot c(t,u(t),x_{t-})dA_t^u|x_{t-}] \leq \varepsilon$ for all $x_{t-} \in X$ since $J(u) < \infty$

Note that $r_o^\infty = 0$, by assumption $r_s^t < 1$. Hence the result holds for infinite time intervals too. ◻

Remark 1: Compare this proof with Davis-Varaiya [9].

2: $\mathcal{M}_t$, $\mathcal{M}_t^n$, $\mathcal{M}_t$ are derived from $\mathcal{M}^d$, $\mathcal{M}^n$, $\mathcal{M}$ in the same way as $\mathcal{U}_t$ from $\mathcal{U}$.

3: E_u in this theorem depends on the control values between t and t+h, only.

As in §2, one can write

$$r_o^tU(t,x_{t-}) = J_M - \Lambda_o^tU(u) + m_U^u(t) \tag{26}$$

where m_U^u is an $\mathcal{F}_{t-}^x$ martingale (under the measure $\mathcal{P}_u$), and $\Lambda_o^t U(u)$ is an $\mathcal{F}_{t-}^x$ predictable increasing process, which depends on u through the measure $\mathcal{P}_u$. We can show that $\Lambda_s^t U(u) = \Lambda_o^t U(u) - \Lambda_o^s U(u)$ depends only on $\mathcal{F}_s^t = \sigma(x_\tau, \tau \leq \tau < t)$, as follows: by Meyer's supermartingale decomposition theorem ([24])

$$\Lambda_s^t U(u) = \text{w.} \lim_{h\to 0} \int_s^t [U(\tau,x_{\tau-}) - E_u(U(\tau+h,x_{\tau+h-})|\mathcal{F}_\tau^x] \times \frac{d\tau}{h}$$

$$= \text{w.} \lim_{h\to 0} \int_s^t \frac{1}{h} [U(\tau,x_{\tau-}) - E_u(U(\tau+h,x_{\tau+h-})|x_{\tau-})]d\tau$$

For each h the integral depends of $\mathcal{F}_s^{t+h}$ and the limit depends on $\cap \mathcal{F}_s^{t+h} = \mathcal{F}_s^t$ (straightforward from assumption ii) at the beginning of the chapter). This result immediately implies that $m_U^u(t) - m_U^u(s)$ depends only on $\mathcal{F}_s^t$, and hence the martingale is an additive functional on the Markov process $(x_{t-}, \mathcal{F}_t^x, \mathcal{P}_u)$ (for a definition see Kunita-Watanabe [19]).

The Markov version of theorem 2.2 is now obvious:

<u>Theorem 2.5</u> $u^* \in \mathcal{M}$ is optimal if and only if there exist

i) a constant J_M

ii) for each $u \in \mathcal{M}$ (value decreasing) a process $\Lambda_o^t(u) \in \mathcal{A}^+(\mathcal{F}_t^x, \mathcal{P}_u)$ such that $\Lambda_s^t(u)$ is $\mathcal{F}_s^t$-measurable

such that 1) $E_u \Lambda_o^{t_f}(u) = J_o - E_u(r_o^{t_f} J_f) \ \forall u \in \mathcal{M}$

2) for all $t \in I$, all $h > 0$ (s.t. $t+h \in I$), all $u \in \mathcal{M}$, $E_u[(-\Lambda_t^{t+h}(u) + \int_t^{t+h} r_o^s \cdot c(s,u(s),x_{s-})dA_s | x_{t-}] \geq 0$ with equality holding for u*.

Then $J_M = \bigwedge_{u \in \mathcal{M}} J(u) = J(u^*)$ and

$$r_o^t U(t, x_{t-}) = E_{u^*}(\Lambda_t^{t_f}(u) + r_o^{t_f} \cdot J_f | x_{t-})$$

Theorem 2.6: The optimal Markov control is also optimal in the class of nonanticipative controls $\mathcal{U}$, i.e. $\bigwedge_{u \in \mathcal{M}} J(u) = J^*$ (with complete observation).

proof: from theorem 2.5 and 2.2, since the $\Lambda_0^t(u)$ and J_M which exist by (26), can also be used in theorem 2.2. Because $\mathcal{F}_t^y = \mathcal{F}_t^x$, the equation (14b) now only depends on the restriction of u to $[t, t+h)$, i.e. if $\mathcal{C} = \{u : [t, t+h) \to U$, Borel measurable$\}$, then (14b) can be written as (for sufficiently small h).

$$\inf_{u \in \mathcal{C}} E_u \left[-\Lambda_t^{t+h}(u) + \int_t^{t+h} r_o^s \cdot c(s, u(s)) dA^u(s) | \mathcal{F}_t^x \right] = 0.$$

Hence $J_M = J_o = J^*$ ¤

CHAPTER 3

OPTIMALITY CRITERIA: FUNDAMENTAL JUMP PROCESS

In this section, it will be shown how the martingale representation theorem for fundamental jump processes, leads to simplified versions of the theorems obtained in Chapter 2. In particular the case of complete observation and the case of a Markov process, will lead to a true Hamilton-Jacobi equation, that can be used to prove existence of solutions.

For most definitions and properties used below, see [5], §2.

3.1. Mathematical Model

The model of Chapter 2, section 1 is now used with a more detailed structure for the stochastic process $(x_t, \mathcal{F}_t, \mathcal{P}_o)$. Let $(x_t, \mathcal{F}_t^x, \mathcal{P}_o)$ be a fundamental jump process on the probability space $(\Omega, \mathcal{F}^x, \mathcal{P}_o)$. Here Ω is a space of piecewise constant functions $\omega\colon I \to Z$, where $(Z, \mathcal{Z})$ is a Blackwell space, with a finite number of jumps in every finite time interval. Then $x(t,\omega) = \omega(t)$ is the evaluation function. The σ-field from now on is always $\mathcal{F}_t^x = \sigma(x_s, s \leq t)$, the generated σ-fields (completed with respect to the measure $\mathcal{P}_o$). It is assumed that the time of occurrence of the nth jump $T_n(\omega)$ is a totally inaccessible $\mathcal{F}_t^x$-stopping time. This implies that the family $\{\mathcal{F}_t^x\}$ is free of times of discontinuity ([4], lemma 3.1).

Remark: 1. When x_t is a counting process, that is only jumps of size +1 are allowed, then Chou and Meyer [6] have shown that the assumption of total inaccessibility of $T_n(\omega)$ is unnecessary for all of the following results (excluding continuity of $\tilde{P}^x(B,t)$).

To $(x_t, \mathcal{F}^x_t, \mathcal{P}_o)$ one can associate a family of <u>counting processes</u>

$P^x(B,t) = \sum_{s \le t} I_{x_{s-} \ne x_s} \cdot I_{x_s \in B}$, for all $B \in \mathcal{Z}$, and a family of predictable increasing processes $\tilde{P}^x_o(B,t)$ such that

$$Q^x(B,t) = P^x(B,t) - \tilde{P}^x_o(B,t) \in \mathcal{M}^2_{loc}(\mathcal{F}^x_t, \mathcal{P}_o).$$

This predictable increasing process can be written as:

$\tilde{P}^x(B,t) = \int_o^t \int_B n^x_o(dz,s)\ \Lambda^x_o(ds)$. Then $(n^x_o(dz,s),\ \Lambda^x_o(ds))$ is called the local description.

The change of probability measure ($\mathcal{P}_o$ to $\mathcal{P}_u$) can now be described as follows: $\{L^t_o(u)\}$ is a uniformly integrable $\mathcal{F}^x_t$-martingale, that is positive. By [5], theorem 3.1, there exists, for each $u \in U$ a real-valued, predictable function $\phi(z,t,u,\omega) \cdot I_{\{t \le T_k\}} \in L^1_{loc}(\tilde{\ }^{x}_{o})$, where $T_k = \inf\ \{t\colon L^t_o(u) \le \frac{1}{k}\}$. This function ϕ is called the <u>rate process</u>.

$$L^t_o(u) = \prod_{\substack{x_{s-} \ne x_s \\ s \le t}} [1 + \phi(x_s,s,u)] \cdot \exp\left[-\int_o^t \int_Z \phi(z,s,u)\ \tilde{P}^x_o(dz,ds)\right] \tag{1}$$

$$= 1 + \int_o^t L^{s-}_o(u) \int_Z \phi(z,s,u)\ Q^x_o(dz,ds) \tag{2}$$

As in remark 4 of §2-1, it is only a small loss of generality to assume that $\phi(z,t,u)$ depends only on the present value $u(t,\omega)$ of the control law. Hence, from now on $\phi\colon Z \times I \times U \times \Omega \to [-1,\infty)$, which is predictable for each fixed u. The lower bound -1, follows from

$L_o^t(u) \geq 0$, hence $1 + \phi(z,s,u) \geq 0$, $\forall z,s,u$.

From [5], §3-1, it follows that $(x_t, \mathcal{F}_t^x, \mathcal{P}_u)$ is a fundamental jump process with counting processes $P^x(B,t)$, associated predictable increasing process

$$\tilde{P}_u^x(B,t) = \int_o^t \int_B [1 + \phi(z,s,u)]\, \tilde{P}_o^x(dz,ds). \tag{3}$$

and local description $([1 + \phi(z,s,u)]\, n_o^x(dz,s),\ \Lambda_o^x(ds))$. This can be interpreted as follows: the rate of jumps of type $[z,z+dz)$ at time t is changed from $n_o^x(dz,t)\ \Lambda_o^x(dt)$ to $[1 + \phi(z,t,u)]\ n_o^x(dz,t)\ \Lambda_o^x(dt)$. For example, let there be 2 types of jumps, both occurring with rate $\frac{1}{2}$; then $\phi_1 = \frac{1}{2}$, $\phi_2 = -\frac{1}{2}$, changes this to a process with jumps of type 1 occurring with rate 1, and no jumps of type 2 (w.p.1). This example shows that is is <u>not</u> always true that $\mathcal{P}_o << \mathcal{P}_u$.

<u>Remarks: 2.</u> This analysis can be reversed, by starting with a given family of functions $\phi(z,t,u)$, $u \in U$, satisfying all requirements of ϕ, imposed above. Then:

$$\mathcal{U} = \{\text{possible control laws } u\colon I \times \Omega \to U \,|\, E_o L(u) = 1\},$$

and define $\mathcal{P}_u$ by $L(u) = \frac{d\mathcal{P}_u}{d\mathcal{P}_o}$. It is difficult then to specify $\mathcal{U}$, since no necessary and sufficient conditions are known for $E_o L(u) = 1$. The sufficient conditions in [5], §3-3 may be too restrictive for some applications.

3. In this model only rates of transition are changed, corresponding to $\mathcal{P}_u << \mathcal{P}_o$. A number of interesting problems, such as

inventory control, seem therefore excluded. However by making $\mathcal{P}_o$ sufficiently complicated, this can also be included (see example in Chapter 4). Note that this problem does not occur in the finite state Markovian problem, because optimality criteria can be written directly in terms of transition probabilities (see Blackwell [3], Ross [27]).

In order to write the following optimality criteria in the simplest possible form, the total cost J(u) is supposed to be of the following form:

$$J(u) = E_u\left[\int_T \int_Z r_o^s \cdot c(z,s,u(s)) \cdot \tilde{P}_u^x(dz,ds) + r_o^{t_f} \cdot J_f\right] \qquad (4)$$

This does in fact include most other reasonable cost structures:

i) if the cost increases only when a jump occurs then

$$J(u) = E_u\left[\int_I \int_Z r_o^s \cdot c(z,s,u(s))\, P^x(dz,ds) + r_o^{t_f} J_f\right]$$

$$= E_u\left[\int_I \int_Z r_o^s \cdot c(z,s,u(s))\, \tilde{P}_u^x(dz,ds) + r_o^{t_f} J_f\right]$$

ii) if Lebesgues measure t is absolutely continuous with respect to $\tilde{P}_u^x(Z,t)$ then

$$E_u \int_I r_o^s \cdot \tilde{c}(s,u(s))\, ds = E_u \int_I \int_Z r_o^s \cdot \tilde{c}(s,u(s)) \frac{ds}{d\tilde{P}_u^x(Z,ds)} \cdot \tilde{P}_u^x(dz,ds)$$

This is not possible if $\tilde{P}_u^x(Z,t)$ remains constant in a non-zero interval, but this corresponds to no jumps being possible, and it is not unreasonable not to assign any cost to this trivial part of the process.

Moreover the equations following are easily transformed to costs involving integration over ds or $\tilde{P}_o^x(dz,ds)$.

Remark 4. The results of Chou-Meyer [5] and Jacod [18] suggest the following improvement on the results of [4]: if

$$G_n(t,B) = \mathcal{P}(T_{n+1}-T_n \leq t,\ x_{T_{n+1}} \in B|\mathcal{F}^x_{T_n})$$

then:

$$\tilde{P}^x(B,t) = \sum_{T_i\leq t} \int_o^{T_i-T_{i-1}} \frac{G_n(ds,B)}{G_n[s,\infty),Z)} + \int_o^{t-T_n} \frac{G_n(ds,B)}{G_n([s,\infty),Z)} \tag{5}$$

To prove this, let $\mu_s = P^x(T_n+s,B) - P^x(T_n,B) - \int_o^s \frac{G_n(dx,B)}{G_n([s,\infty),Z)}$

or (with $S_n = T_{n+1} - T_n$): $\mu_{s\wedge S_n} = I_{s\geq S_n} \cdot I_{\{x_{S_n}\in B\}} - \int_o^s \frac{G_n(dx,B)}{G_n([x,\infty),Z)}$.

Let $T \geq T_n$ be any stopping time, then by lemma 2.1 of [4], there exists a random variable $R, \mathcal{F}^x_{T_n}$-measurable, s.t. $T \wedge T_{n+1} = (T_n+R)\wedge T_{n+1}$. The result will then be proven if

$$0 = E\mu_{R\wedge S_n} = \mathcal{P}(R\geq S_n, x_{S_n}\in B) - E\int_o^{R\wedge S_n} \frac{G_n(dx,B)}{G_n([x,\infty),Z)} \tag{6}$$

(since the process defined by (5) is obviously a predictable, integrable increasing process). (6) follows from:

$$E\int_0^{R\wedge S_n} \frac{G_n(dx,B)}{G_n([x,\infty),Z)} = \int_0^{\infty} G_n(ds,Z) \int_0^{R\wedge s} \frac{G_n(dx,B)}{G_n([x,\infty),Z)}$$

$$= \int_0^{R} G_n(dx,B) \int_x^{\infty} \frac{G_n(ds,Z)}{G_n([x,\infty),Z)} = \mathcal{P}(R \geq S_n, x_{s_n} \in B).$$

This result also shows that $\tilde{P}^x(B,t)$ is absolutely continuous iff all the conditional distributions $G_n(dt,B)$ are absolutely continuous, and $G_n([0,t),B) < 1$ for all $t < \infty$.

3.2. Optimality Conditions with Partial Information

The simplest way of defining the partial observation field, $\mathcal{F}_t^y$, is as follows: let $(Y,\mathcal{Y})$ be a Blackwell space, let $\gamma : Z \to Y$ be measurable. Then $y_t = \gamma(x_t)$ and $\mathcal{F}_t^y = \sigma(y_s, s \leq t)$ (completed for $\mathcal{P}_o$) define the observations available to the decisionmaker. The function can be used to express that only some of the x-coordinates are observed (if $Z = \mathbb{R}^n$, $Y = \mathbb{R}^m$, $m < n$) or are classified in a finite or countable number of categories (if $Z = \mathbb{R}^n$, $Y = \{0,1,\ldots,N\}^n$ or $\mathbb{N}^n$), or any other noise-free function of the present value of the state. Note that γ cannot be used to express randomness of the measurements; all the randomness in the system has to be included in the fundamental jump process $\{x_t, \mathcal{F}_t^x, \mathcal{P}_o\}$. This can always be achieved by extending the state space Z. Extension of the state space may also be necessary if $y_t = \gamma(x_s, s \leq t)$. Then including y_t as part of the fundamental process, x_t, is always possible, without changing $(\Omega, \mathcal{F}, \mathcal{P}_o)$ (except a trivial change in $(\Omega, \mathcal{F})$ strictly speaking, since the range of the functions in Ω is changed). This is illustrated by the following example: let $x_t = (N_t^1, N_t^2)$, N_t^1, N_t^2 Poisson processes independent of each other, let $dy_t = dN_t^1 - 1_{y_t \geq 0} \cdot dN_t^2$ then $x_t' = (N_t^1, N_t^2, y_t)$ generates a probability space $(\Omega', \mathcal{F}_t^{x'}, \mathcal{P}_o')$ which is isomorphic to $(\Omega, \mathcal{F}_t^x, \mathcal{P}_o)$ obtained for the jump process x_t. The previous example is useful in the theory of queues.

It is obvious that $(y_t, \mathcal{F}_t^y, \mathcal{P}_o)$ is again a fundamental jump process. For any $C \in \mathcal{Y}$ the counting process $P^y(C,t)$ can be written as:

$$P^y(C,t) = \sum_{T_i \leq t} I_{\{\gamma(x_{T_i}) \in C\}}$$

$$= \int_Z I_{\{\gamma(s) \in C\}} \cdot P^x(dz,t) \tag{7}$$

To simplify the notation, form now in $I_c(z) = I_{\{\gamma(z) \in C\}}$. Also, assume $\tilde{P}_o^x(B,t) = \int_o^t \int_B f(z,s)\, \mu(dz,ds)$ where $\mu(B,t)$ is $\mathcal{F}_t^y$-predictable for all $B \in Z$ (in many practical applications it will be deterministic), and $f(z,t)$ is $\mathcal{F}_t^x$-predictable. For an arbitrary $\mathcal{F}_t^x$-adapted process g_t, the notation $\hat{g}_t = E_o(g_t | \mathcal{F}_t^y)$ will be used. Then

$$P^y(C,t) - \int_o^t \int_Z \widehat{I_C f}\,(z,s)\, \mu(dz,ds)$$

$$= \int_o^t \int_Z I_c(z)\, [P^x(dz,ds) - f(z,s)\, \mu(dz,ds)]$$

$$+ \int_o^t \int_Z [I_c(z) f(z,s) - \widehat{I_c(z) f(z,s)}]\, \mu(dz,ds) \in \mathcal{M}^1_{loc}(\mathcal{F}_t^y, \mathcal{P}_o).$$

By the uniqueness of the predictable increasing process associated with $P^y(C,t)$:

$$\tilde{P}_o^y\,(C,t) = \int_o^t \int_Z \widehat{I_C(z)\ f(z,s)}\ \mu(dz,ds) \tag{8}$$

Similarly one can prove

$$\tilde{P}_u^y\,(C,t) = \int_o^t \int_Z \widehat{(1 + \phi(z,s,u))\ I_C(z)\ f(z,s)}\ \mu(dz,ds)$$

Remark 1. In some cases, the extension of the state space Z can be avoided by defining

$$y_t = \int_o^t \int_Z h(z,s)\ P^x(dz,ds) \text{ where } h\colon Z \times I \to \mathbb{R}^n$$

is an $\mathcal{F}_t^x$-predictable process. All the results of this chapter continue to hold if I_C is replaced by h. However if Y is not a vector space, this is not possible.

The argument used above, to find the $\mathcal{F}_t^y$-predictable increasing process associated with $P^y(C,t)$ can also be applied to the increments in cost in the time interval [t,t+h) as expressed by the right hand side of (2.13b):

$$E_u[\int_t^{t+h} \int_Z r_o^s \cdot c(z,s,u(s))\ \tilde{P}_u^x(dz,ds)|\mathcal{F}_t^y]$$

$$= E_u[\int_t^{t+h} \int_Z r_o^s \cdot c(z,s,u(s))(1+\phi(z,s,u(s))\ f(z,s)\mu(dz,ds)|\mathcal{F}_t^y].$$

$$= E_u[\int_t^{t+h}\int_Z \widehat{r_o^{\cdot} \cdot c \cdot (1+\phi) \cdot f(z,s,u) \cdot \mu(dz,ds)} | \mathcal{F}_t^y]$$

Inequality (2.13b) can then be written as

$$0 \leq E_u[- \Lambda_t^{t+h}W(u) + \int_t^{t+h}\int_Z \widehat{r_o^{\cdot}\cdot c\cdot (1+\phi)\cdot f(z,s,u)\cdot\mu(dz,ds)} |\mathcal{F}_t^y] \tag{9}$$

This implies that

$$A_t = - \Lambda_o^t W(u) + \int_o^t\int_Z \widehat{r_o^{\cdot}\cdot c\cdot (1+\phi)\cdot f(z,s,u)}\ \mu(dz,ds)$$

is a submartingale under $\mathcal{P}_u$. At the same time it is clearly $\mathcal{F}_t^y$-predictable. Hence

$A_t = C + B_t + m_t$, B_t an $\mathcal{F}_t^y$-predictable increasing process, C a constant, m_t an $\mathcal{F}_t^y$-martingale. Hence m_t is a predictable martingale with respect to the family of σ-fields $\mathcal{F}_t^y$ which is free of times of discontinuity. This implies

$$m_t \overset{a.s. \mathcal{P}_u}{=} D, \text{ a constant.}$$

Therefore A_t is itself an increasing process and (9) can be replaced by:

$$0 \leq - \Lambda_t^{t+h}W(u) + \int_t^{t+h}\int_Z \widehat{r_o^{\cdot}\cdot c\cdot (1+\phi)f(z,s,u)}\ \mu(dz,ds) \tag{10}$$

with equality holding if and only if u is optimal (this inequality holds a.s. $\mathcal{P}_u$).

Theorem 2.2b can now be rewritten as:

Theorem 3.1. $u^* \in \mathcal{U}$ is optimal if and only if there exist

i) a constant J_o .

ii) a process $\bar{\Lambda}_o^t(u) \in \mathcal{A}(\mathcal{F}_t^y, \mathcal{P}_u)$ for all $u \in \mathcal{U}$ (value decreasing)

iii) a family of processes $\eta(z,t,u) \in L^1(Q_u^y)$

such that:

$$\text{a)}\quad J_o - \bar{\Lambda}_o^{t_f}(u) + \int_o^{t_f}\int_Z \eta(z,s,u)\, P^y(dz,ds) = E_u(r_o^{t_f} J_f | \mathcal{F}_{t_f}^y) \qquad (11)$$

$$\text{b)}\quad -\bar{\Lambda}_t^{t+h}(u) + \int_t^{t+h}\int_Z \left[\widehat{(\eta\cdot I_y + \dot{r}_o \cdot c)(1+\phi)f}\right](z,s,u)\mu(dz,ds)$$

$$\geq 0 \text{ a.s. } \mathcal{P}_u,\ \forall u \in \mathcal{U}, \qquad (12)$$

Then $J_o = J^* = J(u^*)$ and

$$r_o^t\, W(u^*,t) = J_o + \bar{\Lambda}_o^t(u^*) + \int_o^t\int_Z \eta(z,s,u^*)\, P^y(dz,ds) .$$

Proof Applying (2.12b) and the martingale representation theorem, one obtains the existence of $\eta(t,z,u) \in L^1(Q_u^y)$ for all value-decreasing admissible u, and such that

$$r_o^t\, W(u,t) = J_o - \Lambda_o^t\, W(u) + \int_o^t\int_Z \eta(z,s,u)\, Q_u^y(dz,ds)$$

$$= J_o - \Lambda_o^t\, W(u) - \int_o^t\int_Z \eta(z,s,u)\, \widehat{I_Y(1+\phi)f}(z,s,u)\, \mu(dz,ds)$$

$$+\int_o^t\int_Z \eta(z,s,u)\; P^y(dz,ds) \tag{13}$$

which implies the "necessity" part of the theorem (after identifying $\bar{\Lambda}_o^t(u) = \Lambda_o^t W(u) + \int_o^t\int_Z \eta(z,s,u)\, \widehat{L_Y(1+\phi)f(z,s,u)}\; \mu(dz,ds)$.

The theorem then follows from theorem 2.2b.

Remarks:3. A local version of this theorem, corresponding to theorem 2.3, can easily be written down. This version will be stated in some of the applications in Chapter 4.

3. The previous arguments do not apply to the unnormalized case, because the conditional expectation

$$E_o\left[L(u)\int_o^t\int_Z r_o^s\; c(z,s,u(s))\;\tilde{P}_u^x(dz,ds)\,\middle|\,\mathcal{F}_t^y\right]$$

is not necessarily predictable. However, limiting arguments, as in [8], theorem 4.3 lead to similar results.

4. The previous result is similar to the minimization of a Hamiltonian in the deterministic optimal control problem, where $\bar{\Lambda}_o^t(u)$ plays the role of a costate. However because of the closed loop nature of the stochastic control problem, the costate now depends on the control applied in the past.

3.3 Optimality Conditions with Complete Information

In this section it is assumed that $y_t = x_t$, i.e. $h(z,t) = z\ \forall t$, and then $\mathcal{F}_t^y = \mathcal{F}_t^x$. This is called the complete observation case, because the decision-maker knows the whole past of the process. The results then are considerably simpler. Observe that

$$W(u,t) = \bigwedge_{v\in\mathcal{U}_t} \frac{E_o[L_o^t(u)L_t^{t_f}(v)(\int_t^{t_f}\int_Z r_o^s\cdot c(z,s,v(s))\tilde{P}_u^x(dz,ds))|\mathcal{F}_t^x]}{E_o[L^t(u)L_t^{t_f}(v)|\mathcal{F}_t^x]}$$

$$= \bigwedge_{v\in\mathcal{U}_t} E_o[L_t^{t_f}(v)(\int_t^{t_f}\int_Z r_o^s\cdot c(z,s,v(s))\ \tilde{P}_u^x(dz,ds))|\mathcal{F}_t^x)$$

$= W(t)$, independent of u, the control law used before t.

<u>Remark 1</u>. $L_t^{t_f}(v) = \dfrac{L(u,t,v)}{L_o^t(v)} = \dfrac{L(v)}{L_o^t(v)}$, independent of u.

Then the processes $\Lambda_o^t\ W(u)$ and $\eta(z,t,u)$ in (13) can still depend on u, because they are defined for each probability measure $\mathcal{P}_u$. However

$$r_o^tW(t) = J_o - \Lambda_o^t\ W(u_1) + \int_o^t\int_Z \eta(z,s,u_1)\ Q_{u_1}^x\ (dz,ds)$$

$$= J_o - \Lambda_o^t\ W(u_2) + \int_o^t\int_Z \eta(z,s,u_2)\ Q_{u_2}^x\ (dz,ds)$$

Identifying the jumps (and using predictability of $\eta(z,t,u)$), gives:

$$\int_o^t\int_Z \eta(z,t,u_1)\ P^x(dz,ds) = \int_o^t\int_Z \eta(z,t,u_2)\ P^x(dz,ds)$$

Therefore $\eta(z,t,u) = \nabla W(z,t)$, independent of u. Then

$$\bar{\Lambda}_o^t\, W = \Lambda_o^t\, W(u) + \int_o^t \int_Z \nabla W(z,s)(1+\phi(z,s,u))\, \tilde{P}_o^x(dz,ds)$$

is also independent of u. Consequently theorem 3.1 can be simplified to:

<u>Theorem 3.2</u>. $u^* \in \mathcal{U}$ is optimal if and only if there exist:

i) a constant J_o

ii) a process $\bar{\Lambda}_o^t \in \mathcal{A}(\mathcal{F}_t^x, \mathcal{P}_o)$

iii) a process $\eta(z,t) \in L^1(Q_o^x)$

such that

$$\text{a)}\quad J_o - \bar{\Lambda}_t^{t_f} + \int_o^{t_f} \int_Z \eta(z,s)\, P^x(dz,ds) = r_o^{t_f} \cdot J_f \tag{14}$$

$$\text{b)}\quad - \bar{\Lambda}_t^{t+h} + \int_t^{t+h} \int_Z [\eta(z,s) + r_o^s \cdot c(z,s,u(s))]\, [1+\phi(z,s,u(s))]$$

$$\tilde{P}_o^x(dz,ds) \geq 0 \ (\text{a.s. } \mathcal{P}_u) \quad \forall\, u \in \mathcal{U} \tag{15}$$

and

$$- \bar{\Lambda}_t^{t+h} + \int_t^{t+h} \int_Z [\eta(z,s) + r_o^s \cdot c(z,s,u^*(s))]\, [1+\phi(z,s,u^*(s)]$$

$$\times\, \tilde{P}_o^x(dz,ds) = 0 \ (\text{a.s. } \mathcal{P}_{u^*}) \tag{16}$$

Then $J_o = J^* = J(u^*)$ and

$$r_o^t \cdot W(t) = J_o - \bar{\Lambda}_o^t + \int_o^t \int_Z \eta(z,s)\, P^x(dz,ds)$$

Proof. Immediate by identifying $\bar{\Lambda}_o^t = \bar{\Lambda}_o^t\, W$ and $\eta(z,t) = \nabla W(z,t)$. ¤

Remark 3. This criterion now is a true Hamilton-Jacobi equation, since $\bar{\Lambda}_o^t\, W$ and $\nabla W(z,t)$ are known explicitly as functions of W (i.e. $\bar{\Lambda}$ and ∇ are known operators), then an optimal u^* can be found as follows: solve the minimization problem (for some small h)

$$\inf_Y \{-\bar{\Lambda}_t^{t+h} W + \int_t^{t+h} \int_Z [\nabla W(z,s) + r_o^s \cdot c(z,s,Y(\bar{\Lambda}W,\nabla W))]$$

$$\times\ [1+\phi(z,s,Y(\bar{\Lambda}W,\nabla W,s))]\ P^x(dz,ds)\}$$

and obtain $Y^*(\bar{\Lambda}W,\nabla W,s)$. Replacing u^* in (16) by Y^* then gives a complicated operator equation, to be solved for W. Finally u^* is obtained by replacing W by its value in Y^*.

4. The method explained above would be considerably easier, if an adjoint equation were known for the costate $\eta(z,t)$ (as in Pontryagin's maximum principle). In some very special cases, this has been done by Fleming [14], Kushner [21] and Sworder [32].

5. Since $V(u,t) = L_o^t(u)\, W(t)$, depends explicitly on u, no genuine Hamiltonian can be obtained for the unnormalized value function.

The following corollaries are easy consequences of theorem 3.2.

Corollary 3.1. Let $\tilde{P}_o^x(dz,ds) = n_o^x(dz,s)\lambda(s)\ ds$, and make assumptions b) and c) of 2.3, then $u^* \in \mathcal{U}$ is optimal if and only if there exist

i) a constant J_o

ii) a predictable process $\alpha(t)$ s.t. $\int_o^t \alpha(s)ds \in \mathcal{A}(\mathcal{F}_t^x, \mathcal{P}_o)$

iii) a family of processes $\eta(z,t) \in L^1(Q_o^x)$.

such that:

$$\text{a)}\quad J_o - \int_o^{t_f} \alpha(s)\ ds + \int_o^{t_f} \int_Z \eta(z,s)\ P^x(dz,ds) = r_o^{t_f} J_f \tag{17}$$

$$\text{b)}\quad -\alpha(t) + \int_Z [\eta(z,t) + r_o^t \cdot c(z,t,u)][1+\phi(z,t,u)]\ n_o^x(dz,t) \times \lambda(t) \geq 0, \quad \forall t \in I,\ \forall u \in U \tag{18}$$

with equality holding for $u^*(t)$.

Proof. This follows immediately from theorems 2.3 and 3.2, with

$$\alpha(t) = A_t^x W(u) + \int_Z \nabla W(z,t)\ n_u^x(dz,t)\ \lambda(t) \qquad \square$$

Remark 6. The importance of this theorem is, that it gives a minimization over U instead of over a function space $\mathcal{U}$.

Corollary 3.2. Suppose at each time only n different values of a jump are possible (see [4], §4). Let $\lambda_i(t)$ be the rate for each of these jumps under $\mathcal{P}_o$, $\phi_i(t,u)$ the rate process for jump type i.

Let $c(z,s,u) = c_i(s,u)$ for z a jump of the $\underline{i}$th type. Assume the conditions of corollary 3.1 are satisfied. Then $u^* \in \mathcal{U}$ is optimal if and only if there exist a constant J^*, a process $\alpha(t)$ s.t. $\int_0^t \alpha(s)\, ds \in \mathcal{A}(\mathcal{F}_t^x, \mathcal{P}_0)$ and n processes $\eta_i(t) \in L^1(Q_0^x)$, such that

$$J_0 - \int_0^t \alpha(s)\, ds + \sum_{i=1}^{n} \int_0^t \eta_i(s)\, p_i(ds) = J_f \tag{19}$$

and

$$-\alpha(t) + \sum_{i=1}^{n} [\eta_i(t) + r_0^t \cdot c_i(t,u)][1+\phi_i(t,u)]\lambda_i(t) \geq 0 \qquad \forall t \in I,\ \forall u \in U \tag{20}$$

with equality holding for $u^*(t)$.

Remark 7. Corollary 3.1 and 3.2 could have been stated in the partial information case but the expressions take a rather complicated form, and are probably not very useful.

In some special cases, it will now be shown that the heuristic method given in remark 3 to construct an optimal control law, leads to a formal proof of the existence of an optimal control law. The method of proof is adapted from Davis [7].

The following assumptions are made throughout the rest of this section:

i) $I = [0,1]$

ii) U is a separable, metric space, $\mathcal{B}_U$ is the class of all Borel sets (i.e. the σ-algebra generated by all sets in the topology).

iii) $E_o \exp(M\, P(Z,1)) < \infty$ for all $M > 0$ and $\tilde{P}(Z,t) < \mu(t)$ a.s. where $\mu(t)\colon R_+ \to R_+$ is an increasing deterministic function.

iv) $\phi(z,t,u,\omega)\colon Z \times I \times U \times \Omega \to R$ is jointly measurable (i.e. with respect to $\mathcal{Z} \otimes \mathcal{B}[0,1] \otimes \mathcal{U} \otimes \mathcal{F}^x$) and $\phi(\cdot,\cdot,u,\cdot)$ is $\mathcal{F}_t^x$-predictable for each fixed u

v) $\phi(z,t,\cdot,\omega)\colon U \to R$ is continuous on U.

vi) $c(z,t,u,\omega)$ is jointly measurable, $\mathcal{F}_t^x$-predictable for fixed u and continuous on U for fixed (z,t,ω).

vii) there exist $\alpha > 0$, K, $K^1 < \infty$ such that:

a) $$\int_Z [1 + \phi(z,t,u)]^\alpha\, n_o^x(dz,t) < K + K^1[P^x(Z,t) + \tilde{P}_o^x(Z,t)] \text{ a.s.}$$

b) $$\int_Z |\ln\,(1+\phi(z,t,u))|\, P^x(dz,t) < K + K^1[P^x(Z,t) + \tilde{P}_o^x(Z,t)] \text{ a.s.}$$

Condition iii) together with vii,a) implies that $\{L^t(u)\}$ is a family of uniformly integrable martingales, by proposition 3.4 in [5], while iii) and vii,b) imply $L(u) > 0$ for all $u \in \mathcal{U}$ (since iii) implies that all moments of $P(Z,1)$ exist under the measure $\mathcal{P}_o$, and hence also $E_o[\exp M(P(Z,1))^2]$). Condition vii) could be replaced by any other condition insuring the above properties.

Note that $L(u) > o$ implies that $\mathcal{P}_o << \mathcal{P}_u$ and hence the only probability measures $\mathcal{P}_u$ allowed are mutually absolutely continuous with respect to $\mathcal{P}_o$. This is a very strong condition (only 1 of the examples in §4 satisfies it), but is seems unavoidable for both lemma 3.1 and theorem 3.3.

The following notation is now introduced:

$\Phi = \{\phi(z,t,\omega): Z \times I \times \Omega \to \mathbb{R} \mid \phi$ satisfies conditions iv), v) and vii) (take a set U containing 1 point)$\}$

$$\Phi_N = \{\phi \mid \phi \in \Phi, -1 + \frac{1}{N} \leq \phi(z,t,\omega) \leq N\}$$

If $\phi \in \Phi$ then $L_o^t(\phi) = \prod_{\substack{s \leq t \\ x_{s-} \neq x_s}} [1+\phi(x_s,s)] \exp[-\int_o^t \int_Z \phi(z,s)\tilde{P}_o^x(dz,ds)]$

i.e. $L_o^t(u) = L_o^t(\phi(\cdot,\cdot,u,\cdot))$ with the old notation. Let $\mathcal{D}(\Phi) = \{L_o^1(\phi) \mid \phi \in \Phi\}$ and similarly define $\mathcal{D}(\Phi_N)$.

Lemma 3.1. $\mathcal{D}(\Phi)$ is convex and weakly compact (i.e. compact in $\sigma(L_1,L_\infty)$-topology).

Proof. Convexity is immediate from Meyer's differentiation rule, since for ϕ_1, $\phi_2 \in \Phi$

$$\rho_t = \lambda_1 \cdot L_o^t(\phi_1) + \lambda_2 L_o^t(\phi_2) = L_o^t \left(\frac{\lambda_1 \cdot L_o^t(\phi_1)\ \phi_1 + \lambda_2\ L_o^t(\phi_2)\phi_2}{\lambda_1 L_o^t(\phi_1) + \lambda_2 L_o^t(\phi_2)}\right)$$

and the argument of the last $L_o^t(\cdot)$ is obviously in Φ.

Furthermore by prop. 3.4 of [5], $\mathcal{D}(\Phi)$ is uniformly integrable and hence weakly sequentially compact. The lemma will then be proven if we show that $\mathcal{D}(\Phi)$ is strongly closed. First consider $\mathcal{D}(\Phi_N)$. Since

$$L_o^t(\phi)^2 = \prod_{\substack{x_{s-}\neq x_s \\ s \leq t}} [1+2\phi(x_s,s)] \exp[-\int_o^t \int_Z 2\phi(z,s)\, \tilde{P}_o^x(dz,ds)$$

$$+ [\prod_{\substack{x_{s-}\neq x_s \\ s \leq t}} \phi^2(x_s,s)] \times \exp[-\int_o^t \int_Z 2\phi(z,s)\, \tilde{P}_o^x(dz,ds)]$$

$$\leq e^{\mu(t)+2\ln N \cdot P(Z,1)}$$

and hence $E\, L_o^t(\phi)^2 \leq f_o(N)$ (f_o an increasing function). Therefore $\mathcal{D}(\Phi_N)$ is L_2-bounded; L_2-closure is proven in the same way as lemma 3 of [10].

Now let $\{\phi_n\}$ be a sequence in Φ such that

$$\lim_{n\to\infty} L_o^1(\phi_n) \to \rho \text{ a.s. and } L_1(\Omega, \mathcal{F}^x, \mathcal{P}_o).$$

Then $\mathcal{D}(\phi)$ will be L_1-closed if $\rho = L_o^1(\phi)$ for some $\phi \in \Phi$. To show this, let

$$\phi_n^N(z,t,\omega) = \begin{cases} \phi_n(z,t,\omega) & \text{if } -1+\frac{1}{N} \leq \phi_n(z,\tau,\omega) \leq N \text{ for all } z, \text{ all } \tau \leq t \\ 0 & \text{otherwise} \end{cases}$$

Since $\phi_n^N \in \mathcal{D}(\Phi_N)$ is an L_2-closed, bounded and therefore weakly compact set (see theorem V-4-7 of [11]), there exists for each N a $\phi^N \in \Phi_N$ such that $L_o^1(\phi^N) = w \cdot \lim L_o^1(\phi_n^N)$ where the weak limit is taken along some subsequence. As in lemma 7 of [10] one can show that $\phi^{N+i}(z,t,\omega) = \phi^N(z,t)$ whenever both are non-zero. Condition vii) then implies that a $\phi(z,t,\omega) \in \Phi$ exists such that

$$L_o^1(\phi) = w \cdot \lim L_o^1(\phi_n) \qquad \square$$

To prove the existence theorem we introduce the following Hamiltonian:

$$H(t,u,p,\omega) = \int_Z [p(z,\omega)+r_o^t(\omega)\cdot c(z,t,u,\omega)][1+\phi(z,t,u,\omega)] \times n_o^x(dz,t)$$

where $p(z,\omega)$ $Z \times \Omega \to R$ is in $L_1(\Omega,\mathcal{F}^x,\mathcal{P}_o)$.

Remark 8. The Hamiltonian is defined above, assuming the conditions of corollary 3.1 to simplify the notation in the following theorem. However theorem 3.3 still holds in the general case, with

$$H(t,u,p,\omega) = \int_t^{t+h} \int_Z [p(z)+r_o^t\cdot c(z,t,u)][1+\phi(z,t,u)] \; \tilde{P}_o^x(dz,ds)$$

for some h. However, then the minimization is always over a function space (functions $[t,t+h] \to U$).

Theorem 3.3. Suppose that for each t, p, ω the Hamiltonian $H(t,u,p,\omega)$ achieves its minimum over U, i.e. $\exists u_o \in U$ such that:

(H) $$H(t,u_o,p,\omega) = \bigwedge_{u\in U} H(t,u,p,\omega) = \mathcal{H}(t,p,\omega).$$

Then an optimal admissible control law $u^*(t)$ exists.

Proof. By assumptions v) and vi) $H(t,u,p,\omega)$ is continuous in U, and for S a countable, dense (compact) subset of U one has:

$$\mathcal{H}(t,p,\omega) = \bigwedge_{u\in S} H(t,u,p,\omega)$$

This implies that $\mathcal{H}(t,p,\omega)$ is jointly measurable and $\mathcal{F}_t^x$-predictable, such that by (H)

$$\mathcal{H}(t,p,\omega) \in H(t,U,p,\omega).$$

By an extension of Filippov's lemma (see [1], lemma 1) there exists a mapping $y(t,p,\omega)$, jointly measurable and $\mathcal{F}_t^x$-predictable, such that $\mathcal{H}(t,p,\omega) = H(t,y(t,p,\omega),p,\omega)$

In the comments preceding theorem 3.1, it is shown that there exists a process $\nabla W(z,t)$, integrable for fixed t and $\mathcal{F}_t^x$-predictable. Hence $u^*(t,\omega) = y(t,\nabla W(z,t,\omega),\omega)$ is a well-defined predictable process taking values in U, i.e. it is an admissible control. To prove that u^* is an optimal control law, it suffices to show (by corollary 3.1):

$$\bar{\Lambda}_o^t\, W = \int_o^t \beta(s)\, ds \quad \text{a.s. } \mathcal{P}_{u^*}, \text{ for all } t \in I \tag{22}$$

where $\beta(t,\omega) = \lambda(t,\omega) \cdot \mathcal{H}(t,\nabla W(z,t,\omega),\omega)$

$$= \lambda(t,\omega) \int_Z [\nabla W(z,t) + r_o^s \cdot c(z,t,u^*(s))] \cdot [1+\phi(z,t,u^*(t))]$$

$$\times\, n_o^x(dz,t)$$

From corollary 3.1 it is now obvious that

$$J(u^*) - J^* = E_{u^*}\left[\int_o^1 \beta(s)\, ds - \bar{\Lambda}_o^1\, W\right]$$

and $J(u) - J^* = E_u \left[\int_o^1 \int_Z [\nabla W(z,s) + r_o^s \cdot c(z,s,u(s))] \, \tilde{P}_u^x(dz,ds) - \bar{\Lambda}_o^1 W\right]$

$$\geq E_u \left[\int_o^1 \beta(s) \, ds - \bar{\Lambda}_o^1 W\right]$$

By definition of J^*, there exists a sequence of controls $\{u_n\}$ such that $J(u_n) \downarrow J^*$. By assumption iv) and vii) $\{\phi(u_n)\}$ is a sequence in ϕ, hence there exists a $\phi \in \Phi$ such that:

$$L_o^1(\phi) = \underset{n\to\infty}{w\cdot\lim} \, L_o^1(\phi(u_n))$$

Therefore, for each N:

$$E_o\left[L_o^1(\Phi)\cdot\left(\left(\int_o^1 \beta(s) \, ds - \bar{\Lambda}_o^1 W\right) \quad N\right)\right]$$

$$= \lim_{n\to\infty} E_o\left[L_o^1(\phi(u_n))\left(\left(\int_o^1 \beta(s) \, ds - \bar{\Lambda}_o^1 W\right) \wedge N\right)\right]$$

$$= 0$$

Since $L_o^1(\phi) > 0$ a.s. $\mathcal{P}_o$ (comment before lemma 3.1), and since (15) gives $\int_o^t \beta(s) \, ds - \bar{\Lambda}_o^t W \geq 0$, it follows that

$$\int_o^1 \beta(s) \, ds - \bar{\Lambda}_o^1 W = 0$$

Since $\int_o^t \beta(s) \, ds - \bar{\Lambda}_o^t W$ is increasing, this is sufficient to prove (22). ◻

Remark 9. This proof, adapted from Davis [7], is very similar to proofs of the existence of optimal controls on diffusion processes (see Fleming [13]), but has the advantage that it does not make continuity assumptions on the value functions. An alternative proof can be given along the lines of Benes [2] and Duncan-Varaiya [10], if one assumes that $\phi(z,t,U)$ is convex.

3.4. Markovian Jump Processes.

Suppose $(x_t, \mathcal{F}_t^x, \mathcal{P}_o)$ is a Markov process, $\mathcal{F}_t^y = \mathcal{F}_t^x$, and: $\phi(z,s,u(s),\omega) = \phi(z,s,u(s),x_{s-}(\omega))$ (ϕ denotes 2 different functions really), and $c(z,s,u(s),\omega) = c(z,s,u(s),x_{s-}(\omega))$. Then assumptions i) to vi) of §2.4 are clearly satisfied. Moreover assume $c(z,t,u,x_{t-})$ and $\phi(z,t,u,x_{t-})$ are uniformly continuous, on the separable metric space U, fixed z,t,x_{t-}. Since $\{L(u): u \in \mathcal{M}\}$, $\mathcal{M}$ the class of Markov controls, is uniformly integrable, there must exist a sequence of stopping time τ_n such that with :

$$\begin{aligned} u^n(t,\omega) &= u(t,\omega) \qquad t \le \tau_n, \int_o^t |u(s,\omega)|^2\, ds \le n \text{ a.s.} \\ &= 0 \qquad\qquad\quad t > \tau_n \end{aligned}$$

(and for notational simplicity take $\phi(z,t,0,x_{t-}) = 0$) then

$$L(u^n) \to L(u) \text{ in } L_1\ (\Omega, \mathcal{P}_o).$$

By [23], VIII,2, lemma 1, there exists a sequence of discrete control laws $\{u_{nk}\}$(discrete in sense of §2.4) such that

$$\lim_{k\to\infty} E_o \int_o^{\tau_n} |u_n^{(s)} - u_{nk}^{(s)}|^2\, ds = 0$$

$$\text{and } \lim_{k\to\infty} E_o \int_o^{\tau_n} |\phi(z,s,u_n(s),x_{s-}) - \phi(z,s,u_{nk}(s),x_{s-})|^2 \, ds = 0$$

by uniform continuity. Hence $L(u_{nk}) \to L(u)$ in probability, which together with uniform integrability and the L_1-convergence theorem in [22], p. 163, shows that

$$L(u_{nk}) \overset{n,k\to\infty}{\to} L(u) \text{ in } L_1(\Omega,\mathcal{P}_o)$$

So, condition viii) of §2.4 is satisfied too. All the results of §2.4 now apply to Markovian jump processes as above.

Remark 1. The convergence argument above is taken from Davis-Varaiya [9].

The previous results can now be summarized in the following optimality criterion. In agreement with most of the literature on the optimization of Markov processes, it will be assumed that transition densities exist (i.e. $\tilde{P}_o^x(B,t)$ is absolutely continuous, see remark 4 in §3.1).

Theorem 3.4. Assume $(x_t,\mathcal{F}_t^x,\mathcal{P}_o)$ is a Markovian jump process as described above, and let

$$\tilde{P}_o^x(B,t) = \int_o^t \int_B n^x(dz,s,x_{s-}) \cdot \lambda(s,x_{s-}) \, ds$$

Then

$$\bigwedge_{u\in\mathcal{M}} J(u) = \bigwedge_{u\in\mathcal{U}} J(u). \tag{24}$$

A Markovian control law $u^* \in \mathcal{M}$ is optimal if and only if there exist:

i) a constant J_o

ii) a function: $\alpha(t,x_{t-}):I \times Z \to R$ such that

$$\int_o^t \alpha(s,x_{s-})\, ds \in \mathcal{A}(\mathcal{F}_t^x, \mathcal{P}_o)$$

iii) a family of functions $\eta(z,t,x_{t-})$: $Z \times I \times Z \to Z$ such that

$$E_o \left| \int_o^t \int_Z \eta(z,s,x_{s-})\, P^x(dz,ds) \right| < \infty$$

which satisfy:

$$\text{a)} \quad J_o - \int_o^{t_f} \alpha(s,x_{s-})\, ds + \int_o^{t_f} \int_Z \eta(z,s,x_{s-})\, P^x(dz,ds) = J_f \qquad (25)$$

$$\text{b)} \quad -\alpha(t,x_{s-}) + \int_Z [\eta(z,t,x_{t-}) + r_o^t \cdot c(z,t,u,x_{t-})]$$

$$\times\ [1+\phi(z,t,u,x_{t-})]\ n_o^x(dz,t,x_{t-})\lambda(t,x_{t-}) \geq 0\ \forall t \in I\ \forall u \in U \quad (26)$$

with equality holding for $u^*(t)$.

Proof. Immediate from corollary 3.3 and theorems 2.5 and 2.6, except for the special form of $\eta(z,t,\omega)$: this is obvious however from the remark that m_U^x is an additive functional on the Markov process (compare with Kunita-Watanabe, §5, [19])

Remark 2. $\alpha(t,x_{t-})$ can be identified as:

$$\alpha(t,x_{t-}) = A_t\, U\,(x_{t-},u) + \int_Z \eta(z,t,x_{t-})[1+\phi(z,t,u,x_{t-})]$$

$$\times\, n_o^x(dz,t)\ \lambda(t,x_{t-})$$

where $A_t U$ is the differential generator corresponding to the process $U(t,x_{t-})$ on $(\Omega,\mathcal{F}^x,\mathcal{P}_u)$. No expression is known for $\eta(z,t,x_{t-})$ for a general jump process, unlike the Wiener process where it can be identified with a gradient if $U(t,x_{t-})$ is sufficiently smooth. However, the inequality (26) can be rewritten as:

$$- A_t\, U(x_{t-},u) + \int_Z r_o^t \cdot c(z,t,u,x_{t-})[1+\phi(z,t,u,x_{t-})]$$

$$\times\, n_o^x(dz,t,x_{t-}) \cdot \lambda(t,x_{t-}) \geq 0 \qquad (27)$$

with equality holding for $u^*(t)$. This equation will be used in the next chapter to solve some examples.

CHAPTER 4

APPLICATIONS

In this chapter, it will be shown how the results obtained in Chapter 3, can be applied to some real problems. First, an explicit form will be given for the increasing process, introduced in Chapter 2. This will then be used to prove some simple results on systems, linear in the control. Finally a number of examples from varying fields, will be modeled as jump processes, and explicit forms of the optimal control derived.

1. Infinitesimal Generators.

In [12], Dynkin shows that for a wide class of functions $f(x_t)$, where x_t is a Markov jump process, the following limit exists in the weak topology on L_∞:

$$w \cdot \lim_{h \to 0} \frac{1}{h}\{f(x_t) - E[f(x_{t+h})\ x_t]\} = \lambda[f(x_t) - \int_Z f(z)\ \pi(x_t, dz)]$$

where λ is the rate at which jumps occur, while $\pi(x,dz)$ is the distribution of jumps starting from x, given a jump occurs. This result can be rewritten for the time-dependent function $U(t,x_t)$, in terms of $\tilde{P}^x(B,t)$:

Proposition 4.1. Let $\tilde{P}_o^x(B,t) = \int_o^t \int_B n_o^x(dz,s) \cdot \lambda(s) \cdot ds$

Let $U(t,z) \leq M$, $\forall\ z \in Z$; let $U(t,z)$ be differentiable in t.

Then:

$$w \cdot \lim_{h\to 0} \frac{1}{h} \{U(t,x_{t-}) - E_u\, [U(t+h,x_{t+h-}) | \mathcal{F}_t^x]$$

$$= -\frac{\partial U(t,x_t)}{\partial t} + \lambda(t) \times \int_Z [U(t,x_{t-})-U(t,z)][1+\phi(z,t,u)]n^x(dz,t) \tag{1}$$

with boundary condition $U(t_f,x_{t_f-}) = J_f$

Proof. $E_u\, [U(t+h,x_{t+h-}) | x_{t-}]$

$$= U(t+h,x_{t-}) \times \mathcal{P}_u\ (\text{no jump in } [t,t+h) | x_{t-})$$

$$+ \int_Z U(t+h,z) \times \mathcal{P}_u \left(\left| \begin{matrix} \text{jump from } x_t \text{ to} \\ (z,z+dz) \text{ in } [t,t+h) \end{matrix} \right| x_{t-} \right)$$

$$+ f(M) \times o(h)$$

Let $t \in (T_n,T_{n+1}]$, then (G_n as in remark 4 of §3.1):

$$\mathcal{P}_u(\text{no jump in } [t,t+h) | x_{t-}) = 1 - \frac{G_n(Z,[t,t+h))}{G_n(Z,[t,\infty)}$$

and

$$\mathcal{P}_u(1 \text{ jump: } x_t \text{ to } (z,z+dz) \text{ in } [t,t+h) | x_{t-}) = \frac{G_n(dz,[t,t+h))}{G_n(Z,[t,\infty))}$$

which is $\mathcal{F}_{T_n}$ measurable, and $x_{t-} = x_{T_n}$ for the ω's considered. By remark 4 of §3.1, the result is then immediate. ¤

Consider now a jump process $(x_t,\mathcal{F}_t^x,\mathcal{P}_o)$, not necessarily Markovian. By cor. 2.4 of [4],

$$W(t,\omega) = W(t,x_{T_i\wedge t},T_i\wedge t,\ i=1,2,\ldots).$$

Denote:

$$W_n(t,x_{T_i},T_i;i=1,\ldots,n) = W(t,x_{T_1},\ldots,x_{T_n},x_{T_n},\ldots,T_1,\ldots,T_n,T_n,\ldots).$$

Then on $\{\omega: T_n \leq t < T_{n+1}\}$ we have:

$$W(t,\omega) = W_n(t,x_{T_i},T_i;\ i=1,\ldots,n).$$

The previous jump process can be made Markovian by imbedding it into another process that has the whole past $(x_{T_i \wedge t}, T_i \wedge t) = x_t$ as state. The extension of proposition 4.1 is now obvious:

<u>Theorem 4.1</u>. Let $\tilde{P}_o^x(B,t) = \int_o^t \int_B n_o^x(dz,s)\ \lambda(s)\ ds$

Let $W(t,\omega) \leq M$, $\forall\omega \in \Omega$. Let $W(\cdot,\omega)$ be differentiable in t. Then on the set $\omega : T_n(\omega) \leq t < T_{n+1}(\omega)\}$

$$w \cdot \lim \frac{1}{h}\{W(t,\omega) - E_u[W(t+h,\omega)|\mathcal{F}_t^x]\}$$

$$= -\frac{\partial W(t,\omega)}{\partial t} + \lambda(t,\omega) \int_Z [W_n(t,x_{T_i},T_1,\ldots\ i=1\ldots,n) \tag{2}$$

$$- W_{n+1}(t,x_{T_i},z,T_i,t;\ i=1,\ldots,n)]\, n_o^x(dz,t)$$

with boundary condition $W(t,\omega) = J_f$.

<u>Remark 1</u>. The boundedness condition on the value function could probably be replaced by an integrability condition. Since boundedness does not seem to pose a problem for the finite state examples below, this has not been done.

2. In the partial information case the result is:

$$w \cdot \lim \frac{1}{h} \{W(u,t,\omega) - E_u[W(u,t+h,\omega)|\mathcal{F}_t^y]\}$$

$$= -\frac{\partial W}{\partial t}(u,t,\omega) + \lambda(t,\omega)\int_Z [W_n(t,y_{T_i},T_i) - W_{n+1}(t,y_{T_i},y,T_i,t)]$$

$$\times \widehat{[1+\phi(z,t,u)]\; I_C(z)\; f(z,t)}\; \mu(dz,t)$$

where ^ denotes $E_u(\cdot|\mathcal{F}_t^y)$. This is too complicated to be very useful.

Theorem 4.2. Let the conditions of theorem 4.1 and corollary 3.1 be satisfied. Then the optimal control law $u^*(t,\omega)$ can be found by solving:

$$0 = \min_{u\in U} \Big[\frac{\partial[r_o^t W(t,\omega)]}{\partial t} + r_o^t\cdot\lambda(t)\Big[\int_Z [-W_n(t,x_{T_i},T_i) + W_{n+1}(t,x_{T_i},z,T_i t)$$

$$+ c(t,z,u)]\;[1+\phi(t,z,u)]\; n_o^x(dz,t)\Big] \tag{3}$$

Proof. Immediate by using remark 2 of §3.4 (which holds for non-Markovian processes too). ¤

2. Bang-Bang Control.

It is well known that, in deterministic optimal control, for a cost linear in the control, the optimal control is bang-bang, i.e. takes a finite number of values only, and jumps from one of these values to another, a finite number of times. We now prove a similar result in the stochastic case.

Theorem 4.2. Let U be a compact subset of $\mathbb{R}^m$, defined by $A \cdot u + B \geq 0$, where $A, B \in \mathbb{R}^{n\times m}$. Let either $c(z,t,u,\omega)$ or $\phi(z,t,u,\omega)$ be of the form $A_c(z,t,\omega) \cdot u + B_c(z,t,\omega)$ (i.e. affine function of u) and the other (c or ϕ) independent of u. Then the optimal control law $u^*(t,\omega)$ jumps a finite number of times (in every finite time interval) between the vertices of U (finite in number).

Proof. The criterion of theorem 4.2 takes the form

$$0 = \min_{u\in U} [f(z,t,U(t,x_{t-}),\omega) \times u + g(z,t,U(t,x_{t-}),\omega)] \tag{5}$$

This is a linear programming problem, so the optimal values of u lie on one of the vertices of U. ¤

Remark 1. If the cost is of the form

$$E_u[\int_o^t (A(s)\ u + B(s))\ ds + J]$$

then criterion (3) is transformed to:

$$0 = \min_{u\in U} [\frac{\partial[r_o^t\ W(t,\omega)]}{\partial t} + r_o^t\ \lambda(t) \int_Z (-W_n + W_{n+1})(1+\phi)\ n_o^x(dz,t)$$

$$+ A(t)\ u + B]$$

Hence, the optimal control law is bang-bang, if both c and ϕ are affine in u.

2. If c and ϕ are more complicated functions of u, it may still be possible to solve (5) for u, as a function of $U(t,x_{t-})$ and then

solve the highly non-linear partial differential equation. Then an optimal control law will be constant, equal to the value of a vertex, over some intervals, but can change continuously between those intervals.

3. Examples

a. The simplest case imaginable is controlling the rate of a point process $(N_t, \mathcal{F}_t^N, \mathcal{P}_o)$, with rate 1. The decision-maker can change the rate to any value $u \in [a,b]$, $b > a \geq 0$. Since the process is Markovian under $\mathcal{P}_o$, and assuming $\mathcal{F}_t^N$ is observed, the control u only has to depend on t and the present value N_t of the jump process.

Observing that $\phi(t,u) = u - 1$, the optimality criterion can be written as:

$$0 = \min_{a \leq u \leq b} \left[\frac{\partial U(t,N_{t-})}{\partial t} + u(U(t,N_{t-}+1) - U(t,N_{t-})) + c(t,u,N_{t-})\right] \tag{6}$$

together with $U(t_f,N_{t_f}) = J_f(N_{t_f})$

One possible cost structure suggested by D. Snyder, related to minimizing damage to a sample in electron microscopy, is:

$$\max_{u \in \mathcal{U}} \mathcal{P}_u(N_{t_f} = k) = \max_{u \in \mathcal{U}} E_u(I_{\{N_{t_f}=k\}}).$$

Then $c(t,u,N_{t-}) = 0$ and $J_f = I_{\{N_{t_f}=k\}}$.

The optimal control law is

$$u^*(t) = b \quad \text{if } U(t,N_{t-}+1) > U(t,N_{t-})$$

$$u^*(t) = a \quad \text{if } U(t,N_{t-}+1 \leq U(t,N_{t-}).$$

Note that (6) has to be maximized for this application. Equation (6) becomes:

$$0 = \frac{\partial U(t,N_{t-})}{\partial t} + b \cdot \max\,(U(t,N_{t-}+1) - U(t,N_{t-}),0)$$

$$+ a \cdot \min\,(U(t,N_{t-}+1) - U(t,N_{t-}),0)$$

which can easily be solved starting from

$$U(t_f,N_{t_{f-}}) = 1 \quad \text{if } N_{t_f} = k$$

$$= 0 \quad \text{otherwise.}$$

Remark 1. Suppose there were a second, independent Poisson process M_t which can neither be observed nor controlled, and one wants to maximize $\mathcal{P}_u\,(N_t + M_t = k)$. The problem now is one with partial information. However because of the independence assumed everything said above still works, if one replaces the final condition by:

$$U(t_f,N_{t_f}) = I_{\{N_{t_f}=i\}} \cdot e^{-t_f} \cdot \frac{t_f^{k-i}}{(k-i)!} \qquad i=0,1,\ldots,k \tag{7}$$

$$= 0 \text{ otherwise.}$$

This follows from

$$\mathcal{P}_u(N_t + M_t = k) = \sum_{i=0}^{k} \mathcal{P}_u(N_t = i) \cdot \mathcal{P}_o(M_t = k-i).$$

Note that this analysis fails as soon as M_t can either be controlled or observed, or is not independent of N_t, since then the expression replacing (7) depends explicitly on the control law u. Therefore, it seems that only an iterative solution can be hoped for. This illustrates some of the problems of partial information optimization.

b. queues: An interesting class of problems is the optimization of networks of queues, as occur in traffic control, management of a time-sharing computer, etc. Only the simplest possible cases will be considered here.

Consider the following simple queue: on a probability space $(\Omega, \mathcal{F}, \mathcal{P}_o)$ are defined 2 independent Poisson processes A_t (rate λ) and D_t (rate 1), representing the arrivals up to time t, and the potential departures up to time t. The queue-length Q_t can then be expressed by:

$$Q_t = A_t - \int_o^t I_{\{Q_{t-}>0\}}\, d\, D_t$$

since no departure is possible if the system is empty. Then Q_t is a jump process, having jumps of size +1 with rate 1, jumps of size -1 with rate 1 if $Q_{t-} > 0$, with rate 0 if $Q_{t-} = 0$. This is a Markovian jump process.

We now suppose the service rate of this system can be changed from 1 to any integer $u = 0,1,\ldots,N$, while the arrival rate is unchanged. Then $\phi_1(t,u) = 0$, $\phi_2(t,u,Q_{t-}) = u - 1$ if $Q_{t-} > 0$. If $J(u) = E_u[\int_o^{t_f} c(s,u,Q_s)\, ds + f(Q_{t_f})]$ then the optimum is obtained by:

$$0 = \min_{u=0,1,\ldots,N} [c(t,u,Q_{t-}) + \frac{\partial U(t,Q_{t-})}{\partial t} + \lambda(U(t,Q_{t-}+1) - U(t,Q_{t-}))$$

$$+ u \cdot I_{\{Q_{t-}>0\}} \cdot (U(t,Q_{t-}-1) - U(t,Q_{t-}))]$$

with $U(t_f, Q_{t_f}) = f(Q_{t_f})$.

If $c(t,u,Q_{t-}) = a \cdot u + Q_{t-}$ $(a>0)$ and $f(Q_{t_f}) = 0$ then the optimal solution is very easy between 0 and t_f - a use

$$u^*(t) = N \text{ if } Q_{t-} > 0$$

$$u^*(t) = 0 \text{ if } Q_{t-} = 0$$

and in $(t_f-a, t_f]$: $u^*(t) = 0$ always. This corresponds to

$$U(t,Q_{t-}) = (t_f-t)\, Q_{t-} + \lambda \frac{(t_f-t)^2}{2} \quad \text{for } t \in (t-a, t_f).$$

It is not necessary to find $U(t,Q_{t-})$ on $[0,t_f-a]$, since $U(t,Q_{t-}) - U(t,Q_{t-}-1)$ is clearly decreasing with increasing time, and hence larger than a throughout the interval.

A slightly more complicated problem is that of choosing between serving one of 2 queues (e.g. traffic lights at an intersection). Each of the queues Q_t^1 and Q_t^2 are described as above, and the control values (u_1,u_2) possible are (1,0) and (0,1). Then the optimality criterion becomes:

$$\begin{aligned}
0 = \inf_{(u_1,u_2)\left\{\substack{=(0,1)\\ =(1,0)}\right.} \Bigg\{ & c(t,u_1,u_2,Q_{t-}^1,Q_{t-}^2) + \frac{\partial U(t,Q_{t-}^1,Q_{t-}^2)}{\partial t} \\
& + \lambda_1 (U(t,Q_{t-}^1+1,Q_{t-}^2) - U(t,Q_{t-}^1,Q_{t-}^2)) \\
& + \lambda_2 (U(t,Q_{t-}^1,Q_{t-}^2+1) - U(t,Q_{t-}^1,Q_{t-}^2)) \\
& + u_1 (U(t,Q_{t-}^1-1,Q_{t-}^2) - U(t,Q_{t-}^1,Q_{t-}^2)) \cdot I_{Q_{t-}^2 > 0} \\
& + u_2 (U(t,Q_{t-}^1,Q_{t-}^2-1) - U(t,Q_{t-}^1,Q_{t-}^2)) \cdot I_{Q_{t-}^2 > 0} \Bigg\}
\end{aligned}$$

with $U(t_f, Q^1_{t_{f-}}, Q^2_{t_{f-}}) = 0$

Assuming c is independent of the control value, one gets $u^*(t)=(1,0)$ if

$$(U(t,Q^1_{t-}-1,Q^2_{t-}) - U(t,Q^1_{t-},Q^2_{t-})) \cdot I_{Q^1_{t-}>0}$$

$$< (U(t,Q^1_{t-},Q^2_{t-}-1) - U(t,Q^1_{t-},Q^2_{t-})) \cdot I_{Q^2_{t-}>0}$$

c. Optimal investment. An example of a jump process with a large (but finite) number of types of jumps, can be obtained by considering the price $\pi_i(t)$ of some stocks $(i=1,\ldots,n)$. We assume each of the prices $\pi_i(t)$ can change by jumps over a fraction $\alpha_{ij}, j=1,\ldots,m_i$ and $\alpha_{ij} \geq -1\ \forall i, j$, (i.e. $\pi_i(T+) = \pi_i(T-) + \alpha_{ij}\ \pi_i(T-)$) occurring at random time T_{ij} with rate $\lambda_{ij\ell}(t)$. Then, according to Merton [23], if somebody has wealth $W(t)$ (≥ 0 always), and has invested a fraction $w_i(t)$ of this in stock i, his wealth will also be a jump process, described by:

$$dW_t = \sum_{i=1}^{n} w_i(t) \cdot W_{t-} \cdot \frac{d\pi_i(t)}{\pi_i(t-)} \quad , \quad \sum_{i=1}^{n} w_i(t) = 1 \tag{8}$$

This is a jump process that has jumps of size $\alpha_{ij} w_i(t) \cdot W_t$ with rate $\lambda_{ij}(t)$. Here, as before, the probability measure depends on the choice of the fractions $w_i(t)$, which therefore are called the controls. In this form, Varaiya [35] has shown that the problem

$\max\limits_{\substack{0 \leq w_i(t) \leq 1 \\ \Sigma w_i(t)=1}} E_u\, J(W_{t_f})$, with $J(W_{t_f})$ the utility of the wealth at the final time t_f, can be solved by a simple non-dynamical optimization.

By using the results of Chapter 2 it is possible to solve a slightly more general problem: suppose the investor also has a wage income $y_t \cdot dt$ in $[t,t,+dt)$ (beyond his control) and can consume $C_t \cdot dt$ of his wealth, in $[t,t+dt)$, where $C_t \in \mathbb{R}_+$ can be chosen freely, and is thus a new control variable. Then:

$$dW_t = (y_t - C_t)\, dt + \sum_{i=1}^{n} w_i(t) \cdot W_{t-} \cdot \frac{d\pi_i(t)}{\pi_i(t-)} \quad .$$

We want to maximize

$$E_u \left[\int_0^{t_f} J(t,C_t)\, dt + J_f(W_{t_f}) \right]$$

where J and J_f represent the utility of consumption at t, resp. of the final wealth. Now the probability measure $\mathcal{P}_u$ is determined by the control variable $w_i(t)$, $i=1,\ldots,n$ and C_t. In the present case W_t is no longer a jump process, because of $\int_0^t (y_s - C_s) ds$. However if we assume $\lambda_{ij}(t)$ to depend only on the present value W_{t-}, then $(W_t, \mathcal{F}_t^W, \mathcal{P}_o)$ will be Markovian ($\mathcal{P}_o$ any fixed choice of $w_i(t)$ and C_t). It is then easy to calculate the infinitesimal generator of the value function

$$U(t,W_{t-}) = \inf_{w_i(t),C_t} E_u \left[\int_t^{t_f} J(s,C_s)\, ds + J_f(W_{t_f}) \,|\, W_{t-} \right]$$

as

$$w \cdot \lim_{h \to 0} \frac{1}{h} (U(t,W_{t-}) - E_u(U(t+h,W_{t+h-})|W_{t-})]$$

$$= -\frac{\partial U(t,W_{t-})}{\partial t} + (C_t - y_t) \cdot \frac{\partial U(t,W_{t-})}{\partial W} \tag{9}$$

$$+ \sum_{i=1}^{n} \sum_{j=1}^{m_i} \lambda_{ij}(t) \; [U(t,W_{t-}) - U(t,(1+w_i(t)\alpha_{ij})W_{t-})]$$

Remark 4. The above calculation can be extended to the case where $\pi_i(t)$, $i=1,\ldots,n$ includes a Brownian motion component, stochastically independent of the jumps. See Merton [23].

Using criterion 2.3, the optimality criterion for the system described above becomes:

$$0 = \max_{\substack{C_t \geq 0 \\ w_i(t) \geq 0 \\ \Sigma w_i(t)=1}} [J(t,C_t) + \frac{\partial U(t,W_{t-})}{\partial t} + (y_t - C_{t-}) \cdot \frac{\partial U(t,W_{t-})}{\partial W}$$

$$- \sum_{i=1}^{n} \sum_{j=1}^{m_i} \lambda_{ij}(t)(U(t,W_{t-}) - U(t,(1+w_i(t)\alpha_{ij})W_{t-}))] \tag{10}$$

with $U(t_f,W_{t_{f-}}) = J_f(W_{t_f})$.

Assume $y_t = 0$ and $J(t,C_t) = \frac{C_t^\gamma}{\gamma}$, $J_f(W_{t_f}) = a \cdot W_{t_f}^\gamma/\gamma$, with $a \geq 0$ and $\gamma \in (0,1]$, then one gets by inspection the following solution of (10).

$$U(t,W_{t-}) = f(t) \cdot \frac{W_{t-}^{\gamma}}{\gamma}$$

$$C_t^* = \frac{W_{t-}}{f(t)^{1/1-\gamma}}$$

and $w_i^*(t)$ is obtained by the maximization:

$$\max_{\substack{w_i(t)\geq 0 \\ \sum_{i=1}^{n} w_i(t)=1}} \sum_{i=1}^{n} \sum_{j=1}^{m_i} \lambda_i(t) \cdot \frac{f(t)}{\gamma} \cdot W_{t-}^{\gamma} \cdot [(1+w_i\ (t)\alpha_{ij})^{\gamma} - 1] = A.$$

Here $f(t) = [(1+a-\gamma)\ \exp(A \cdot \frac{t_f - t}{\gamma}) + \gamma - 1]^{1-\gamma}$

This is a very easy structure for the optimum: $w_i^*(t)$ is the same as in the simple case treated in [35], while C_t^* is proportional to W_{t-}, the constant of proportionality being a predetermined function of time.

It should be noted that this simple solution depends very heavily on the assumptions made about the utility functions (it can be extended to $J(t,C_t) = \frac{(a(t)C_t + b(t))^{\gamma}}{\gamma}$ and $J(t,C_t) = \ln(a(t)C_t + b(t))$ with J_f having the same form) and on the assumption that there is no wage income. In all other cases it seems that (10) will be more useful as a starting point for approximations.

Remark 5. One could try to model consumption and wage income as jump processes. However it is difficult then to find a reasonable form for the utility of consumption. Moreover, it would then be reasonable to let jumps occur at fixed times, which would violate the assumption of total inaccessibility of the jump times.

d. <u>Modeling problems</u>: As mentioned in Chapter 2-3 it may be necessary to choose a complicated probability measure $\mathcal{P}_o$, to make all physically realizable or approximately realizable $\mathcal{P}_u$ absolutely continuous with respect to $\mathcal{P}_o$. A simple example is given by inventory problems (for a review, see Scarf [28]). Suppose an inventory can contain $X_t = -\infty,\ldots,0,1,\ldots,N$ items of one product. The number of items decreases by 1 at random times T_i, with a given interarrival distribution. At some points in time S_i, it is decided to increase the inventory by some number K_i items. Those arrive s time units after the decision is made, s also having a given distribution. The problem then is:

$$\min E_u \left[\int_0^{t_f} L(\tau,X_\tau)d\tau + \sum_{\substack{i=1 \\ S_\ell \le t_f}}^{\ell} f(K_i) + J_f(X_{t_f}) \right]$$

The control here consists in choosing the times S_i and the amounts K_i to order. Note that $X_t \le 0$ is allowed because orders keep arriving even if no item is available. This situation should be avoided, therefore make $L(s,X_s)$ very large for negative X_s.

Let X_t take all integer values $\le N$. If $x_t = k$, then it can jump, under the measure $\mathcal{P}_o$, to values $k-1,k+1,\ldots,N$ with rate 1 for each jump. Under measure $\mathcal{P}_u$ all but one of those jumps, have rate 0 (i.e. $\phi_{k\ell} = -1$). The one jump allowed with rate 1, under $\mathcal{P}_u$ is from k to k-1 if one decides not to order ($\phi_{k,k-1}=0$) or from X_{S_i+s} to $X_{S_i+s} + K_i$; if one decided to order K_i items at time S_i, this jump will occur with rate $p_i(s)$ (where $p_i(s)$ is the density of s, the time the order takes to arrive), at each time $S_i + s$. Then

$$\phi_{X_{S_i+s}X_{S_i+s} + K} = p_i(s) - 1.$$

The control problem then takes the form described in §2 and 3, if $p_i(s) > 0$, $\forall s \geq 0$. However it is usually reasonable to assume that s only takes a discrete number of values (say arrive 1 or 2 or 3 days after ordering). Then the jump times of X_t are not all totally inaccessible; therefore it is not known whether the theory of §3 applies to this case. Also, no case has been found where the optimality criteria can be analytically solved or where the simple (s,S)-ordering policy (see Scarf [28]) can be rederived.

Another class of problems, which can be modeled as an optimization of the form considered in §2, are optimal stopping problems. Then, let $\phi(t,\tau(\omega)) = - I_{\{t>\tau(\omega)\}}$ where $\tau(\omega)$ is any stopping time, which now plays the role of a control. The problem then is, to find a $\tau(\omega)$ such that it maximizes:

$$E_\tau \; J(\tau(\omega),\omega).$$

Application of theorem 2.3 unfortunately leads to a partial differential-difference equation with moving boundary, which is difficult to solve.

CHAPTER 5

CONCLUSION

In this thesis a fairly general form of the stochastic optimal control problem is discussed. It is shown that the mathematical framework chosen (absolutely continuous changes of the probability measure) leads to some abstract necessary and sufficient conditions for optimality. In the case of a jump process, these conditions are simple enough, to be used in deriving some properties of the optimal control, even though they do not give a computationally feasible algorithm. Therefore, the thesis can be considered as a study on the application of jump processes.

REFERENCES

1. V. Beneš: "Existence of Optimal Strategies Based on Specified Information, for a Class of Stochastic Decision Problems," SIAM J. Control 8, 179-188, 1970.

2. V. Beneš: "Existence of Optimal Stochastic Control Laws," SIAM J. Control 9, 446-475, 1971.

3. D. Blackwell: "Discounted Dynamic Programming," Ann. Math. Statist. 36, 226-235, 1965.

4. R. Boel, P. Varaiya and E. Wong: Martingales on Jump Processes I: Representation Results, Electronics Res. Lab., Univ. of Calif., Berkeley, Calif. Memo #M-407, September 1973, submitted to SIAM J. Control.

5. R. Boel, P. Varaiya and E. Wong: Martingales on Jump Processes II: Applications, Electronics Res. Lab., Univ. of Calif., Berkeley, Calif. Memo #M-409, December 1973, submitted to SIAM J. Control.

6. C. S. Chou - P. A. Meyer: "Sur la Représentation des Martingales comme Intégrales Stochastiques dans les Processus Ponctuels," to appear in: Séminaire de Probabilitiés VIII, Lecture Notes in Mathematics, Springer Verlag.

7. M. H. A. Davis: "On the Existence of Optimal Policies in Stochastic Control," SIAM J. Control 11, 587-594, 1973.

8. M. H. A. Davis - P. Varaiya: "Dynamic Programming Conditions for Partially Observable Stochastic Sysyems," SIAM J. Control 11, 226-261, 1973.

9. M. H. A. Davis - P. Varaiya: "Information States for Linear Stochastic Systems," J. Math Analysis of Appl. 37, 384-402, 1972.

10. T. E. Duncan - P. Varaiya: "On the Solutions of a Stochastic Control System," SIAM J. Control 9, 354-371, 1971.

11. N. Dunford - J. T. Schwartz: Linear Operators I, Interscience Publishers, Inc., New York, 1958.

12. E. B. Dynkin: Markov Processes, vol. I, New York: Academic Press, 1965.

13. W. Fleming: "Optimal Continuous-Parameter Stochastic Control," SIAM Rev. 11, 470-509, 1969.

14. W. Fleming: "Stochastic Lagrange Multipliers," in: Mathematical Theory of Control, Balakrishnan and Neustadt, eds., Proc. Symp. USC 1967, Academic Press, New York, 1967.

15. W. Fleming: "Optimal Control of Partially Observable Diffusions," SIAM J. Control 6, 194-214, 1968.

16. M. Fujisaki, G. Kallianpur, H. Kunita: "Stochastic Differential Equations for the Non-linear Filtering Problem," Osaka J. Math, 9, 19-40, 1972.

17. I. V. Girsanov: "On Transforming a Certain Class of Stochastic Processes by Absolutely Continuous Substitution of Measures," Theory of Prob. & Appl. 5, 285-301, 1960.

18. J. Jacod: "On the Stochastic Intensity of a Random Point Process over the Half-line," Department of Statistics, Princeton University, April 1974.

19. H. Kunita - S. Watanabe: "On Square Integrable Martingales," Nagoya Math. J. 30, 209-245, 1967.

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

21. H. Kushner: "On the Stochastic Maximum Principle: Fixed Time of Control," J. of Math. Analysis of Appl. 11, 78-92, 1965.

22. M. Loève: Probability Theory, Princeton, N. J.: D. Van Nostrand Cy., 3rd ed., 1963.

23. R. Merton: "Optimum Consumption and Portfolio Rules in a Continuous Time Model," J. Econ. Theory 3, 373-413, 1971.

24. P. A. Meyer: Probability and Potentials, Blaisdell, Waltham, Mass., 1966.

25. P. A. Meyer: Processus de Markov, Lecture Notes in Mathematics, 26, Springer-Verlag, Berlin, 1967.

26. P. A. Meyer: Intégrales Stochastiques, Séminaire de Probabilités I, Lecture Notes in Mathematics 39, Springer-Verlag, Berlin, 1967.

27. S. Ross: Applied Probability Models with Optimization Applications, Holden-Day, San Francisco, 1970.

28. H. Scarf: A Survey of Analytic Techniques in Inventory Theory, in: Multistage Inventory Models and Techniques, Stanford University Press, Stanford, Calif., 1963.

29. A. V. Skorokhod: Studies in the Theory of Random Processes, Addison Wesley, Reading, Mass., 1965.

30. R. L. Stratonovich: Conditional Markov Processes and their Application to the Theory of Optimal Control, American Elsevier, New York, 1968.

31. D. Stroock - S. Varadhan: "Diffusion Processes with Continuous Coefficients," Comm. Pure & Appl. Math 22, 345-400, 1969.

32. D. Sworder: "Feedback Control of a Class of Linear Systems with Jump Parameters," IEEE Trans. Automatic Control AC-14 9-14, 1969.

33. J. van Schuppen: Estimation Theory for Continuous Time Processes, a Martingale Approach, Electronics Res. Lab., Univ. of Calif., Berkeley, Calif., Memo #M-405, September 1973.

34. J. van Schuppen - E. Wong: Transformation of Local Martingales under a Change of Law, Electronics Res. Lab., Univ. of Calif., Berkeley, Calif. Memo #M-385, May 1973.

35. P. Varaiya: "The Martingale Theory of Jump Processes," to appear in IEEE Trans. on Automatic Control, February 1975.