# OPTIMAL CONTROL OF JUMP PROCESSES

by

R. Boel and P. Varaiya

Memorandum No. ERL-M553

October 1975 (Revised)

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

# OPTIMAL CONTROL OF JUMP PROCESSES[1]

by

R. Boel and P. Varaiya[2]

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

ABSTRACT

The paper proposes an abstract model for the problem of optimal control of systems subject to random perturbations, for which the principle of optimality takes on an appealing form. This model is specialized to the case where the state of the controlled system is realized as a jump process. The additional structure permits operationally useful optimality conditions. Some illustrative examples are solved.

[1] Research supported in part by the Joint Services Electronics Program Contract F44620-71-C-0087 and the National Science Foundation Grant GK-43024X and for R. Boel, by an ESRO-NASA International Fellowship.

[2] Part of this paper was completed when one of the authors (P. V.) was visiting the Department of Electrical Engineering at M.I.T.

## 1. INTRODUCTION

This paper addresses the problem of the optimal control of dynamical systems subject to random perturbations. It does so in the following way. First, in Section 3, an abstract mathematical model is proposed in which the choice of controller is modelled as choosing a probability measure over the measurable space of state trajectories. This idea was first developed by Benes [1,2] and Duncan and Varaiya [11] in order to prove existence of an optimal control when the perturbations form a Brownian motion. Second, in Section 4, we derive optimality conditions for the abstract model using dynamic programming and elements of martingale theory in the way developed by Davis and Varaiya [9] for the Brownian motion case. Their approach in turn was motivated by the work of Rishel [20]; it also has some resemblance to earlier work by Kushner [16], and Stratonovich [26]. Some of the extensions of their results as given in Section 4 are special cases of recent results of Striebel [25]. While the abstract model does serve to unify previous results, further comprehension of the scope of the model can be gained and an evaluation of its practical import can be made only by working through with more specialized problems with additional structure. Hence, in Sections 5 and 6, the case where the random perturbations constitute a jump process is discussed in detail. Related results using different methods have been reported by Rishel [21] and Stone [24] and we shall compare them later. We note that there are control problems with jump disturbances which must be modelled quite differently from the model of Sections 5 and 6. As examples of these we mention the work of Rishel [22] and Sworder [27].

## 2. CONVENTIONS AND NOTATIONS

Let $(\Omega, \mathcal{F})$ be a measurable space. Let $I = [0,T]$ or $[0,\infty)$ be a fixed time interval with the corresponding final time denoted T. A stochastic process is always a triple $(z_t, \mathcal{F}_t, \mathcal{P})$, $t \in I$, where $\mathcal{P}$ is a probability measure on $(\Omega, \mathcal{F})$, $(\mathcal{F}_t)$ is an increasing family of sub-σ-fields of $\mathcal{F}$ and $(z_t)$ is a family of $(\mathcal{F}_t)$-adapted random variables with values in some unspecified measurable space. When the context makes it clear we write the stochastic process $z = (z_t, \mathcal{F}_t, \mathcal{P})$, $t \in I$, as $(z_t)$ or $(z_t, \mathcal{F}_t)$ or $(z_t, \mathcal{P})$ or $z$; $z_t$, without the parentheses, usually denotes the random variable at time t instead of the process.

All probability spaces are assumed complete, and every increasing family of σ-fields, $(\mathcal{F}_t)$, is assumed right-continuous i.e., $\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s$. An $(\mathcal{F}_t, \mathcal{P})$-martingale is a uniformly integrable martingale $(m_t, \mathcal{F}_t, \mathcal{P})$, $t \in I$, with $m_0 = 0$ a.s. The collection of all such martingales is denoted $\mathcal{M}^1(\mathcal{F}_t, \mathcal{P})$. In a similar way, we define $\mathcal{M}^2(\mathcal{F}_t, \mathcal{P})$, $\mathcal{M}^1_{loc}(\mathcal{F}_t, \mathcal{P})$, $\mathcal{M}^2_{loc}(\mathcal{F}_t, \mathcal{P})$, the classes of $(\mathcal{F}_t, \mathcal{P})$- uniformly square integrable, locally integrable, locally square integrable martingales, and it will be assumed that a version of these processes is chosen such that it has right-continuous sample paths with left-hand limits.

$\mathcal{A}^+(\mathcal{F}_t, \mathcal{P})$ is the class of all processes $(a_t, \mathcal{F}_t, \mathcal{P})$, $t \in I$ which vanish at 0, $a_0 = 0$ a.s., with right-continuous, non-decreasing sample paths, and which are uniformly integrable, $\sup_t E\, a_t < \infty$. $\mathcal{A}(\mathcal{F}_t, \mathcal{P}) = \mathcal{A}^+(\mathcal{F}_t, \mathcal{P}) - \mathcal{A}^+(\mathcal{F}_t, \mathcal{P})$ is then the class of processes with integrable variation. The classes $\mathcal{A}^+_{loc}$, $\mathcal{A}_{loc}$ are defined in the usual way.

A family $(z_t)$ of $(\mathcal{F}_t)$-adapted functions and taking values in a metric space is said to be $(\mathcal{F}_t)$-predictable if there is a sequence of such families $(z_t^n)$, $n=1,2,\ldots$, with left-continuous sample paths such that $\lim_{n\to\infty} z_t^n(\omega) = z_t(\omega)$ for all $(t,\omega) \in I\times\Omega$.

## 3. ABSTRACT MODEL OF THE CONTROL PROBLEM

The model proposed below is similar to the one presented and investigated in [25]. It consists of three interconnected parts: a description of the dynamical system i.e., the way in which it is affected by the control action, a description of the set of allowable control laws, and a description of the cost associated with each control law. The assumptions imposed are given next.

We suppose given measurable spaces $(Z, \mathcal{Z})$, the state space, and $(\Omega, \mathcal{X})$, the trajectory or sample space. Also given is a function $x_t(\omega): I \times \Omega \to Z$ which is measurable with respect to $\mathcal{B}_I \times \mathcal{X}$. Let $\mathcal{X}_t = \sigma\{x_s | s \leq t\}$ and without losing generality we assume that $\mathcal{X} = \sigma\{x_t | t \in I\}$. We now assume

$\underline{S}_1$ The behavior of the system under the action of any (admissible) control law u is completely described by the specification of a probability measure $\mathcal{P}^u$ on $(\Omega, \mathcal{X})$.

Thus for each control law u, $x^u = (x_t, \mathcal{X}_t, \mathcal{P}^u)$, $t \in I$, is a well-defined stochastic process. We are evidently modelling the system as a controlled probability space rather than as a controlled set of trajectories which is more customary. Of course in the deterministic context the latter model is the more natural one. We now describe the set of control laws.

We suppose given a measurable space $(\mathcal{U}, \mathcal{B}_u)$, the control space, where $\mathcal{U}$ is a metric space. Also given is an increasing family of σ-fields, $(\mathcal{Y}_t)$ called the family of observations, such that $\mathcal{Y}_t \subset \mathcal{X}_t$, $t \in I$. A collection $\mathcal{U}$ of functions $u_t(\omega): I \times \Omega \to \mathcal{U}$ is a collection of (admissible) control laws if the following holds:

$\underline{S}_2$ (i) $(u_t)$ is $(\mathcal{Y}_t)$-adapted and $(u_t, \mathcal{Y}_t, \mathcal{P}^u)$, $t \in I$ is a measurable process.

(ii) $\mathcal{U}$ is closed under concatenation i.e., if u, $v \in \mathcal{U}$ then so does $(u,v,t)$ where $(u,v,t)(s) = u(s)$ for $s \leq t$, $= v(s)$ for $s>t$.

(iii) For each $u \in \mathcal{U}$ and $A \in \mathcal{X}_t$, $\mathcal{P}^u(A)$ depends only on $u_s$, $s \leq t$ i.e., if $v \in \mathcal{U}$ is such that $u_s \equiv v_s$, $s \leq t$, then $\mathcal{P}^u(A) = \mathcal{P}^v(A)$; for each $u \in \mathcal{U}$, $A \in \mathcal{X}$ $E^u(1_A|\mathcal{X}_t)$ does not explicitly depend on $u_s$, $s \leq t$ i.e., if $v \in \mathcal{U}$ such that $u_s \equiv v_s$, $s \geq t$, then $E^u(1_A|\mathcal{X}_t) = E^v(1_A|\mathcal{X}_t)$.

In the above, (iii) is a version of a causality condition and also expresses the notion that the past trajectory $x_s$, $s \leq t$ serves as a state at t whereas (ii) is essential for dynamic programming. In (i) the requirement that $u_t$ is $\mathcal{Y}_t$-measurable indicates that $\mathcal{Y}_t$ is the $\sigma$-field of observations available up to t.

We can now describe the cost of control. Associated with each $u \in \mathcal{U}$ is a unique cost $J(u)$ given by

$$J(u) = E^u\left[\int_I r_0^t\, c(t,u(t))\,d\Lambda^u(t) + r_0^T\, J_T\right] \tag{3.1}$$

where $E^u$ denotes expectation with respect to $\mathcal{P}^u$, T denotes the final time of I, and the other terms are described below.

$\underline{C_1}$ The <u>instantaneous cost</u> $c : I \times U \times \Omega \to R$ is a non-negative function which is jointly measurable with respect to $\mathcal{B}_I \times \mathcal{B}_U \times \mathcal{X}$ ($\mathcal{B}_I, \mathcal{B}_U$ are the Borel sets of I,U), continuous with respect to u for fixed $t,\omega$ and measurable with respect to $\mathcal{X}_t$ for fixed t,u.

$\underline{C_2}$ The <u>time rate</u> $\Lambda^u : I \times \Omega \to R$, defined for each $u \in \mathcal{U}$, is adapted to the family $(\mathcal{X}_t)$ and, for each $\omega$, the sample path $t \to \Lambda^u(t,\omega)$ is right-continuous and increasing. Furthermore, $d\Lambda^{(u,v,t)}(s) = d\Lambda^u(s)$ for $s \leq t$, $= d\Lambda^v(s)$ for $s>t$. (See $S_2$ above for a definition of $(u,v,t)$).

Since $\Lambda^u$ can have discontinuous sample paths, the indefinite (Stieltjes) integral $\int_0^t r_0^s\, c(s,u(s))\,d\Lambda^u(s)$ can be discontinuous. The most useful examples of time rates are

a) $\Lambda^u(t) \equiv t$; whenever the sample paths are absolutely continuous with respect to Lebesgue measure on I this case obtains.

b) $\Lambda^u(t,\omega) = \sum 1_{\{t \geq \tau_i(\omega)\}}$ which counts the number of $(\mathcal{X}_t)$-stopping times $\tau_i$, i=1,2,..., which occur before t.

c) $\Lambda^u$ is the predictable increasing process associated with the counting process in b), and which can replace the latter in (1) whenever c(t,u(t)) is a $(\mathcal{X}_t)$-predictable process, since the values of the integrals coincide (see [19]).

$\underline{C}_3$ The discounting rate $r_s^t(\omega)$: is a non-negative function defined for $\omega \in \Omega$, s,t in I with $s \leq t$. For fixed s, $r_s^t(\omega)$ is $(\mathcal{Y}_t)$-adapted, jointly $\mathcal{B}_I \times \mathcal{Y}$ measurable, and uniformly integrable. Furthermore, for each u

$$r_{t_1}^{t_3} = r_{t_1}^{t_2} r_{t_2}^{t_3} \qquad \text{a.s. } \mathcal{P}^u \text{ for } t_1 \leq t_2 \leq t_3$$

$$r_t^t = 1 \qquad \text{a.s. } \mathcal{P}^u \text{ for all } t$$

$\underline{C}_4$ The terminal cost $J_T : \Omega \to R$ is a non-negative $\mathcal{X}$-measurable function. $J_T$ is the cost incurred at or after the final time T. When $T = \infty$ it will be assumed that $J_T \equiv 0$.

$\underline{C}_5$ For all $u \in \mathcal{U}$, $J(u) < \infty$.

The problem of optimal control is to find $u^* \in \mathcal{U}$ such that

$$J(u^*) = \inf_{u \in \mathcal{U}} J(u)$$

Such u* is called an optimal control.

Remark 3.1 (i) The fixed time interval I can be replaced by a random interval $[0,\tau] \subset I$ where $\tau$ is a $(\mathcal{X}_t)$-stopping time. This can be achieved by setting $c(t,u,\omega) = 0$ for $t \geq \tau(\omega)$ or by making $\Lambda^u$ constant after $\tau$. If $\tau$ does not depend on u one can set $r_0^t(\omega) = 0 \quad t \geq \tau(\omega)$.

(ii) The discounting rate $r_s^t(\omega)$ is not allowed to depend explicitly on u. In an economic context this implies that the controller cannot directly influence the interest rate. Of course, since the distribution of $r_s^t$ is dependent upon $\mathcal{P}^u$ there is a possibility of introducing indirect control.

(iii) Except for the special results with complete information (i.e., $\mathcal{Y}_t \equiv \mathcal{X}_t$) or Markovian assumptions, the final cost $J_T$ can depend explicitly on the control law u. Again, except for these special cases, $c(t,u,\omega)$ can be made to depend upon the past $u_s$, $s \leq t$ of the control. These generalizations are not made here since the notational burdens become intolerable.

## 4. OPTIMALITY RESULTS FOR THE ABSTRACT MODEL

Since the proofs of the results are simple modifications of proofs published in [9] we have been content with citing the correspondence. The assumptions made in Section 3 are enforced throughout.

### 4.1 Principle of Optimality

Let u, $v \in \mathcal{U}$ and $t \in I$. We define

$$\psi(u,v,t) = E^{(u,v,t)}\left\{\int_t^T r_t^s c(s,v_s)d\Lambda_s^v + r_t^T J_T \middle| \mathcal{Y}_t\right\}$$

Evidently, from the assumptions made above,

$$\psi(u,v,t) \in L^1(\Omega, \mathcal{Y}_t, \rho^u)$$

The random variable $\psi(u,v,t)$ is the conditional expectation given the observation $\mathcal{Y}_t$ of the future cost beyond time t, evaluated at t, when u is adopted on [0,t] and v is adopted beyond t. To evaluate these costs at time 0 it is only necessary to multiply $\psi(u,v,t)$ by $r_0^t$. Since $L^1$ is a complete lattice under the natural partial ordering for real-valued functions the following infimum exists,

$$W(u,t) = \bigwedge_{v\in\mathcal{U}} \psi(u,v,t) \in L^1(\Omega, \mathcal{Y}_t, \rho^u)$$

Note that $W(u,0) = J^* = \bigwedge_{u\in\mathcal{U}} J(u)$ is the infimum of the acheivable costs. The process $(W(u,t), \mathcal{Y}_t, \rho^u)$ is called the <u>value function</u> corresponding to u. The next definition was introduced by Rishel [20]. It was used in [9].

<u>Definition 4.1</u> $\mathcal{U}$ is said to be <u>relatively complete</u> with respect to W if for each $u \in \mathcal{U}$, $t \in I$, $\varepsilon>0$ there exists $v \in \mathcal{U}$ such that

$$\psi(u,v,t) \leq W(u,t) + \varepsilon \quad \text{a.s. } \rho^u$$

Lemma 4.1 $\mathcal{U}$ is relatively complete with respect to W.

Proof Identical with that of [9, Lemma 3.1] □

Theorem 4.1 For $t_1 \le t_2$ in I and $u \in \mathcal{U}$ we have

$$W(u,t_1) \le E^u\left[\int_{t_1}^{t_2} r_{t_1}^s c(s,u_s)d\Lambda_s^u \middle| \mathcal{Y}_{t_1}\right] + E^u\left[r_{t_1}^{t_2} W(u,t_2) \middle| \mathcal{Y}_{t_1}\right], \tag{4.1}$$

$$W(u,T) = E^u\left[J_T \middle| \mathcal{Y}_T\right] \tag{4.2}$$

Furthermore, u is optimal if and only if equality holds in (4.1).

Proof The proof depends on Lemma 4.1 and follows the same lines as that of [9, Theorem 3.1]. □

Corollary 4.1 For $u \in \mathcal{U}$, the process

$$\hat{J}_t^u = r_0^t W(u,t) + E^u\left[\int_0^t r_0^t c(s,u_s)d\Lambda^u(s) \middle| \mathcal{Y}_t\right]$$

is a $(\mathcal{Y}_t, P^u)$ sub-martingale. u is optimal if and only if this process is a martingale.

Proof Immediate from Theorem 4.1 □

Since the process

$$\left(E^u\left[\int_I r_0^s c(s,u_s)d\Lambda_s^u + r_0^T J_T \middle| \mathcal{Y}_t\right]\right) \in \mathcal{M}^1\left(\mathcal{Y}_t, P^u\right),$$

therefore the process (w(u,t)) is a $(\mathcal{Y}_t, P^u)$-supermartingale, where

$$w(u,t) = E^u\left[\int_t^T r_0^s c(s,u_s)d\Lambda_s^u + r_0^T J_T \middle| \mathcal{Y}_t\right] - r_0^t W(u,t)$$

$$= r_0^t\left[\psi(u,u,t) - W(u,t)\right]$$

Corollary 4.2 For $u \in \mathcal{U}$, the process $(w(u,t), \mathcal{Y}_t, \mathcal{P}^u)$ is a potential. u is optimal if and only if $w(u,t) \equiv 0$.

Remark 4.1 (i) The model proposed above is a special case of the one presented by Striebel [25] and the results obtained above can be obtained from hers. In particular Corollary 4.1 is a version of [25, Theorem 3]. The additional structure that we have imposed will be used to obtain the more detailed results given below. It is possible to replace the "relative completeness" property by the slightly weaker "$\varepsilon$-lattice" property introduced by Striebel.

(ii) Following Samuelson [23] we can give a heuristic interpretation of the submartingale $(\hat{J}_t^u)$. Its value is the expected cost evaluated at t, using the observation $\mathcal{Y}_t$, given that u is adopted up to t and an optimal control is adopted beyond t. This expected value will increase if the non-optimal control is used for a longer time, accounting for the sub-martingale property. If u is optimal, however, then the expected cost remains constant.

(iii) Theorem 4.1 can be rederived from Corollary 4.1. Hence the Optional Sampling Theorem implies that in (4.1) we may replace the deterministic times $t_1$ and $t_2$ by any $(\mathcal{Y}_t)$-stopping times $\tau_1 \leq \tau_2$ with values in I. This observation is often useful.

(iv) Sometimes, as in [2,9,11], there exists a probability measure $\mathcal{P}$ on $(\Omega, \mathcal{X})$ such that $\mathcal{P}^u \ll \mathcal{P}$ for all $u \in \mathcal{U}$. One can then introduce $L(u) = \dfrac{d\mathcal{P}^u}{d\mathcal{P}}$ and

$$\phi(u,v,t) = E\left\{L(u,v,t)\left[\int_t^T r_t^s c(s,v_s)\, d\Lambda_s^v + r_t^T J_T\right] \middle| \mathcal{Y}_t\right\}$$

$$V(u,t) = \bigwedge_{v \in \mathcal{U}} \phi(u,v,t).$$

The previous results can be restated in terms of the unnormalized value function. While in an optimal filtering context working with such unnormalized quantities has certain advantages (see e.g. [6]), we are unable to observe similar advantages in the optimal control context.

(iv) The random variable $w(u,t)$ expresses the loss incurred by using $u$ beyond $t$ as compared with an optimal control.

From the definition of $(w(u,t), \mathcal{Y}_t, \mathcal{P}^u)$ we can verify directly that it is a potential of class (D). Hence by Meyer's decomposition theorem [18], there is a unique predictable process $(A_0^t w(u)) \in \mathcal{A}^+(\mathcal{Y}_t, \mathcal{P}^u)$ and a martingale $(m^w(u,t)) \in \mathcal{M}^1(\mathcal{Y}_t, \mathcal{P}^u)$ such that

decomposition theorem [18], there is a unique predictable process $(A_0^t w(u)) \in \mathcal{A}^+(\mathcal{Y}_t, \mathcal{P}^u)$ and a martingale $(m^w(u,t)) \in \mathcal{M}^1(\mathcal{Y}_t, \mathcal{P}^u)$ such that

$$w(u,t) = \tilde{J}(u) - A_0^t w(u) - m^w(u,t)$$

where $\tilde{J}(u) = w(u,0) = J(u) - J^*$. We know, furthermore, that the following weak limit (in the sense of the $\sigma(L^1, L^\infty)$-topology) exists (see [18]).

$$\begin{aligned} A_0^t w(u) &= \underset{h\to 0}{\text{weak lim}} \int_0^t \frac{1}{h} E^u\left[w(u,s) - w(u,s+h) \,\middle|\, \mathcal{Y}_s\right] ds \\ &= \underset{h\to 0}{\text{weak lim}} \left\{ \int_0^t \frac{1}{h} E^u\left[\int_s^{s+h} r_0^\sigma c(\sigma, u_\sigma)\, d\Lambda_\sigma^u \,\middle|\, \mathcal{Y}_s\right] ds \right. \\ &\qquad \left. - \int_0^t \frac{1}{h} E^u\left[r_0^s W(u,s) - r_0^{s+h} W(u,s+h) \,\middle|\, \mathcal{Y}_s\right] ds \right\} \qquad (4.3) \end{aligned}$$

Now it is easy to see that there exists a predictable process $(\gamma(u)) \in \mathcal{A}^+(\mathcal{Y}_t, \mathcal{P}^u)$ such that

$$\gamma(u) = \text{weak}\lim_{h\to 0} \int_0^t \frac{1}{h} E^u\left[\int_s^{s+h} r_0^\sigma c(\sigma,u_\sigma)\,d\Lambda_\sigma^u \,|\, \mathcal{Y}_s\right] ds \tag{4.4}$$

From (4.3), (4.4) we may conclude that there exists a predictable process $(A_0^t W(u)) \in \mathcal{A}(\mathcal{Y}_t, \mathcal{P})$ viz. $A_0^t W(u) = \gamma(u) - A_0^t w(u)$, such that

$$A_0^t W(u) = \text{weak}\lim_{h\to 0} \int_0^t \frac{1}{h} E^u\left[r_0^s W(u,s) - r_0^{s+h} W(u,s+h) \,|\, \mathcal{Y}_s\right] ds$$

This is sufficient to apply Meyer's decomposition theorem to the process $(r_0^t W(u,t))$ and we may conclude that

$$r_0^t W(u,t) = r_0^0 W(u,0) - A_0^t W(u) + m^W(u,t)$$

$$= J^* - A_0^t W(u) + m^W(u,t), \tag{4.5}$$

where $(m^W(u,t)) \in \mathcal{M}(\mathcal{Y}_t, \mathcal{P}^u)$. Furthermore, since $(W(u,t))$ is evidently of class (D), the decomposition in (4.5) is unique.

In terms of this decomposition we can rewrite (4.1), after multiplying both sides by $r_0^t$, as

$$E^u\left[A_0^{t_2} W(u) - A_0^{t_1} W(u) \,|\, \mathcal{Y}_{t_1}\right] \le E^u\left[\int_{t_1}^{t_2} r_0^s c(s,u_s)\,d\Lambda_s^u \,|\, \mathcal{Y}_{t_1}\right] \text{a.s. } \mathcal{P}^u, \tag{4.6}$$

and we have equality if and only if u is optimal. With these results in hand we can proceed as in the proof of [9, Theorem 4.1] to establish the next proposition.

<u>Theorem 4.2</u> There exists a constant $J^*$ and for every $u \in \mathcal{U}$ there exists a predictable process $(A_0^t(u)) \in \mathcal{A}(\mathcal{Y}_t, \mathcal{P}^u)$ such that

$$E^u A_0^T(u) = J^* - E^u(r_0^T J_T),$$

and such that for $(\mathcal{Y}_t)$-stopping times $\tau_1 \le \tau_2$ with values in I,

$$E^u\left[- A_{\tau_1}^{\tau_2}(u) + \int_{\tau_1}^{\tau_2} r_0^s c(s,u_s) d\Lambda_s^u \mid \mathcal{Y}_{\tau_1}\right] \ge 0 \text{ a.s. } \mathcal{P}^u \tag{4.8}$$

A control law $u = u^*$ is optimal if and only if equality holds in (4.8) for deterministic times $t_1 \le t_2$, and then, furthermore,

$$J(u^*) = J^*,$$

$$r_0^t W(u^*,t) = E^u[A_t^T(u^*) + r_0^T J_T \mid \mathcal{Y}_t] \text{ a.s. } \mathcal{P}^{u^*}.$$

Remark 4.2 This result is a considerable improvement over [9, Theorem 4.1] since there the inequality (4.6) and hence (4.7) is established only for those u which are "value decreasing" i.e., for which $(W(u,t))$ is a supermartingale. The same shortcoming can be noticed in [20]. Of course, if u is value decreasing, then in (4.5) $(A_0^t W(u))$ is an increasing process.

## 4.2 Local Optimality Conditions

One can divide both sides in (4.8) by $\tau_2-\tau_1$ and take limits as $\tau_2-\tau_1 \to 0$. The basic idea is to express $A_0^t(u)$ as an integral with respect to $\Lambda^u$. It appears necessary however to restrict attention to value decreasing controls.

So let $u \in \mathcal{U}$ be such that $(W(u,t))$ is a supermartingale. Then $(A_0^t W(u))$ is an increasing process and (4.6) can be refined as

$$0 \le E^u\left[A_{t_1}^{t_2} W(u) \mid \mathcal{Y}_{t_1}\right] \le E^u\left[\int_{t_1}^{t_2} r_0^s c(s,u_s) d\Lambda_s^u \mid \mathcal{Y}_{t_1}\right] \text{ a.s. } \mathcal{P}^u$$

For any nonnegative, well-measurable process $(\theta_t, \mathcal{Y}_t, \mathcal{P}^u)$ we must therefore have

$$0 \leq E^u \int_0^t \theta_s dA_0^s W(u) \leq E^u \int_0^t \theta_s r_0^s c(s,u_s) d\Lambda_s^u$$

so that whenever the second integral vanishes so does the first. By the Radon-Nikodym theorem there exists a nonnegative process $(\alpha_t(u), \mathcal{Y}_t, \mathcal{P}^u)$ so that

$$A_0^t W(u) = \int_0^t \alpha_s(u) d\Lambda_s^u$$

Using this representation we can restate Theorem 4.2 in a "local" version.

<u>Theorem 4.3</u> There exists a constant $J^*$ and for every value decreasing $u \in \mathcal{U}$ there exists a nonnegative process $(\alpha_t(u), \mathcal{Y}_t, \mathcal{P}^u)$ such that

$$E^u \int_0^T \alpha_t(u) d\Lambda_t^u = J^* - E^u(r_0^T J_T), \text{ and} \tag{4.9}$$

$$-\alpha_t(u)(\omega) + E^u[r_0^t c(t,u_t) | \mathcal{Y}_t](\omega) \geq 0 \tag{4.10}$$

for almost all $(t,\omega)$ with respect to $d\Lambda^u \times d\mathcal{P}^u$ measure. A control law $u = u^*$ is optimal if and only if equality holds in (4.10) and then, furthermore,

$$r_0^t W(u^*,t) = E^u\left[\int_t^T \alpha_s(u^*) d\Lambda_s^{u^*} + r_0^T J_T | \mathcal{Y}_t\right] \text{ a.s. } \mathcal{P}^{u^*}.$$

## 4.3 <u>Complete Information</u>

Suppose $\mathcal{Y}_t \equiv \mathcal{X}_t$ so that at each time t the controller has complete information about the past. Then

$$\psi(u,v,t) = E^{(u,v,t)}\left[\int_t^T r_t^s c(s,v_s)\,d\Lambda_s^v + r_t^T J_T \,\middle|\, \mathcal{Y}_t\right]$$

$$= E^v\left[\int_t^T r_t^s c(s,v_s)\,d\Lambda_s^v + r_t^T J_T \,\middle|\, \mathcal{Y}_t\right]$$

$$= \psi(v,v,t)$$

by assumption $S_2$(iii). Hence $W(u,t)$ does not depend on u, and the preceding results are simpler. Nevertheless the process $A_0^t W(u)$ still depends upon past values of the control law u. Its "derivative" $\alpha_t(u)$ however will often be independent of values of u before to as seen in [9] and in the following sections.

## 5. OPTIMALITY RESULTS FOR JUMP PROCESSES

In this section the abstract model of Section 3 is specialized to the case of a dynamical system whose state process is a (fundamental) jump process as studied in [5,6]. The additional structure gives more content to the formal results established earlier. For a review of the definitions and properties of jump processes see [ 6, §2].

### 5.1 The model and its limitations

The state space $(Z, \mathcal{Z})$ is now also a Blackwell space. $\Omega$ consists of all functions $\omega : I \to Z$ which are piecewise constant, right-continuous and have only a finite number of jumps in a finite time interval. $x_t(\omega) : I \times \Omega \to Z$ is just the evaluation function, $x_t(\omega) = \omega(t)$. $\mathcal{X}_t, \mathcal{X}$ are defined as before.

The observations $(\mathcal{Y}_t)$ are obtained as follows. We suppose given a Blackwell space $(Y, \mathcal{Y})$ and a measurable map $\gamma$: $Z \to \mathcal{Z}$. Let $y_t = \gamma(x_t)$ and $\mathcal{Y}_t = \sigma\{y_s | s \leq t\}$.

With $(x_t)$ and $(y_t)$ we can now associate the following discrete random measures.

$$P^x(B,t)(\omega) = \sum_{s \leq t} 1_{\{x_{s-}(\omega) \neq x_s(\omega) \in B\}} \tag{5.1}$$

= number of jumps of $x(\omega)$ which occur before t and end in $B \in \mathcal{Z}$;

$$P^y(C,t)(\omega) = \sum_{s \leq t} 1_{\{y_{s-}(\omega) \neq y_s(\omega) \in C\}} \tag{5.2}$$

= number of jumps of $y(\omega)$ which occur before t and end in $C \in \mathcal{Y}$

Note that $P^x(B,t)$ is $\mathcal{X}_t$-measurable and $P^y(C,t)$ is $\mathcal{Y}_t$-measurable.

We can now define the collection of admissible control laws $\mathcal{U}$ and the probability measures $\mathcal{P}^u$, $u \in \mathcal{U}$. Let $(\mathcal{U}, \mathcal{B}_u)$ be the control space, where $\mathcal{U}$ is a metric space. $\mathcal{U}$ is the collection of all functions $u_t(\omega)$: $I \times \Omega \to \mathcal{U}$ which are $(\mathcal{Y}_t)$-predictable. It is supposed that for each $u \in \mathcal{U}$ there is given a probability measure $\mathcal{P}^u$ on $(\Omega, \mathcal{X})$ such that the stochastic process $\mathcal{X}^u = (x_t, \mathcal{X}_t, \mathcal{P}^u)$, $t \in I$, is a jump process in the sense of [5,6]. (It is evident that $S_2$(i), $S_2$(ii) are satisfied by these assumptions.) Now from the results of [5,6] or [15] we know that to say that $x_u$ is a jump process it is equivalent to say that there exist <u>continuous</u> processes $(\tilde{P}_u^x(B,t)) \in \mathcal{A}^+_{loc}(\mathcal{X}_t, \mathcal{P}^u)$ for each $B \in \mathcal{Z}$ such that

$$(Q_u^x(B,t)) = (P^x(B,t) - \tilde{P}_u^x(B,t)) \in \mathcal{M}^1_{loc}(\mathcal{X}_t, \mathcal{P}^u) \tag{5.3}$$

Thus the action of $u \in \mathcal{U}$ is completely described by specfying the correspondence

$$u \to \{(\tilde{P}_u^x(B,t), \mathcal{X}_t, \mathcal{P}^u) \mid B \in \mathcal{Z}\}$$

To guarantee assumption $S_2$(iii) and to simplify some notation later on we suppose that for $u \in \mathcal{U}$ and $B \in \mathcal{Z}$ $\tilde{P}_u^x(B,t)(\omega)$ is given by

$$\tilde{P}_u^x(B,t)(\omega) = \iint_{B\,0}^{t} f(z,s,u_s(\omega),\omega)\mu(dz,ds,u_s(\omega),\omega) \tag{5.4}$$

where the integral is an ordinary Stieltjes integral and the prespecified functions f and $\mu$ satisfy these conditions:

(i) $f(z,s,u) = f(z,s,u,\omega)$: $Z \times I \times \mathcal{U} \times \Omega \to R_+$ is jointly measurable, continuous in u for fixed z, s, $\omega$ and for fixed z, u, $(f(z,t,u,\omega))$ is $(\mathcal{X}_t)$-predictable.

(ii) $\mu(B,t,u,\omega) = \mu(B,t,u): \mathcal{Z} \times I \times \mathcal{U} \times \Omega \to R_+$ is jointly measurable and for each fixed B, u, $(\mu(B,t,u))$ is $(\mathcal{Y}_t)$-predictable, continuous and increasing. (In practice $\mu$ is usually a deterministic process.)

Finally the cost J(u) incurred by $u \in \mathcal{U}$ is supposed given by

$$J(u) = E^u\left[\int_Z\int_I r_0^s c(z,s,u_s)\tilde{P}_u^x(dz,ds) + r_0^T J_T\right] \tag{5.5}$$

where c satisfies the same conditions as f does, and $r_s^t$, $J_T$ satisfy the conditions imposed in Section 3. It is assumed that $J(u) < \infty$ for all u. This completes the description of the mathematical model.

Before turning to the analysis of the model we discuss its limits in terms of which empirical control problems can and which cannot be adequately reflected in the model. First of all, as far as the behavior of the state trajectories is concerned the most serious limitation is the requirement that $(\tilde{P}_u^x(B,t))$ have continuous sample paths. It is known (see e.g. [5]) that this restriction is equivalent to saying that the stopping times at which the state jumps i.e. the times of discontinuity of $(x_t(\omega))$, are totally inaccessible. In intuitive terms this means that if the controller observes the first n jumps, then the probability with which it can predict the (n+1)st jump <u>exactly</u> is zero, for each n=0,1,2...(see [5, Lemma 2.4]). Now most problems of queueing, inventory control, machine failures etc. indeed have this property. But there are some problems which do not. For example suppose that in an inventory control problem there is a fixed (deterministic) delay between the time an order is placed and the time that the corresponding delivery is made; evidently the total inventory jumps when the delivery is made and this time of jump can be predicted exactly, and so the model proposed here is inadequate for this example. Now the only reason why we have insisted on the total inaccessibility of the jump times is so that we can use the martingale representation theorems derived in [5]. More recently, such theorems have been obtained without the restriction on the jump times (see [7,8,13,15]) and therefore the results announced below should be extendable to arbitrary jump processes.

The second limitation of the model appears to the requirement that controls have to be predictable processes. One reason for this is based on empirical considerations. Since, the time when the state jumps cannot be anticipated with positive probability, and since in empirical situations there is an infinitesimal delay before the controller can observe and react to a change in state, therefore the predictability requirement seems appropriate to us. In any event since $\mu$ is continuous in t, therefore $\tilde{P}^x_u$ defined by (5.4) is always continuous in t even if u is measurable and not predictable. Hence the results below remain unchanged whether we permit u to be any measurable process so long as we always take the predictable projection of f (as well as of c in (5.5)), or whether one insists at the outset that u be predictable.

Finally, the cost functional (5.5) may appear too limiting since in many situations one may wish to have the cost increase only when a jump occurs. Thus one would prefer to have as cost the amount

$$E^u\left[\int_Z\int_I r_0^s c(z,s,u_s)P_u^x(dz,ds) + r_0^T J_T\right]$$

$$= E^u\left[\sum_{\substack{s\in I\\ x_{s-}\neq x_s}} r_0^s c(x_s,s,u_s) + r_0^T J_T\right]$$

But since $P^x - \tilde{P}^x$ is a martingale and since the integrand above is predictable the quantity above is equal to J(u) given by (5.3) and so there is no loss in generality. (This equality does not obtain if u is not predictable.)

## 5.2 Preliminary Analysis

To simplify notation we write $1_C(z) = 1_{\{\gamma(z) \in C\}}$. Then, from (5.2),

$$P^y(C,t) = \iint_{Z\,0}^{\;\;t} 1_C(z)P^x(dz,ds)$$

We calculate the unique processes $(\tilde{P}^y_u(C,t), \mathcal{Y}_t, \mathcal{P}^u)$ so that

$$(Q^y_u(C,t)) = (P^y(C,t) - \tilde{P}^y_u(C,t)) \in \mathcal{M}^1_{loc}(\mathcal{Y}_t, \mathcal{P}^u) \tag{5.7}$$

For an arbitrary process $(g_t)$ let $\hat{g}_t = E^u(g_t|\mathcal{Y}_t)$ (the appropriate u will always be clear from the context.) Then from (5.4) and (5.6)

$$P^y(C,t) - \iint_{Z\,0}^{\;\;t} \widehat{1_C f(z,s,u_s)}\; \mu(dz,ds,u_s)$$

$$= \iint_{Z\,0}^{\;\;t} 1_C(z)[P^x(dz,ds) - f(z,s,u_s)]\mu(dz,ds,u_s)$$

$$\iint_{Z\,0}^{\;\;t} [1_C(z)f(z,s,u_s) - \widehat{1_C f(z,s,u_s)}]\; \mu(dz,ds,u_s)$$

which is evidently a member of $\mathcal{M}^1_{loc}(\mathcal{Y}_t, \mathcal{P}^u)$. Hence

$$\tilde{P}^y_u(C,t) = \iint_{Z\,0}^{\;\;t} \widehat{1_C(z)f(z,s,u_s)}\; \mu(dz,ds,u_s) \tag{5.8}$$

A calculation similar to that of $\tilde{P}^y_u$ above gives us the expected value, given $\mathcal{Y}_t$, of the increment of the instantaneous cost on the right hand side of (4.6). In terms of the cost functional (5.5),

$$E^u\left[\iint_{Z\ t_1}^{t_2} r_0^s c(z,s,u_s)\tilde{P}_u^x(dz,ds)\,|\,\mathcal{Y}_{t_1}\right]$$

$$= E^u\left[\iint_{Z\ t_1}^{t_2} r_0^s c(z,s,u_s)f(z,s,u_s)\mu(dz,ds,u_s)\,|\,\mathcal{Y}_{t_1}\right]$$

$$= E^u\left[\iint_{Z\ t_1}^{t_2} \widehat{r_0^s c(z,s,u_s)f(z,s,u_s)}\ \mu(dz,ds,u_s)\,|\,\mathcal{Y}_{t_1}\right]$$

so that, by (4.6),

$$0 \le E^u\left[- A_{t_1}^{t_2}W(u) + \iint_{Z\ t_1}^{t_2} \widehat{r_0^s c(z,s,u_s)f(z,s,u_s)}\ \mu(dz,ds,u_s)\,|\,\mathcal{Y}_{t_1}\right]$$

which means that the process

$$a_t = - A_0^t W(u) + \iint_{Z\ 0}^{t} \widehat{r_0^s c(z,s,u_s)f(z,s,u_s)}\ \mu(dz,ds,u_s) \tag{5.9}$$

is a $(\mathcal{Y}_t,\mathcal{P}^u)$-submartingale. It is evidently of class (D) and so by Meyer's decomposition theorem there is a unique predictable process $(b_t) \in \mathcal{A}^+(\mathcal{Y}_t,\mathcal{P}^u)$ and $(m_t) \in \mathcal{M}^1(\mathcal{Y}_t,\mathcal{P}^u)$ so that

$$a_t = b_t + m_t$$

But from (5.9) we know that $(a_t)$ is also $(\mathcal{Y}_t)$-predictable. Hence $(m_t)$ is a predictable process with integrable variation. It must therefore vanish so that $a_t = b_t$. Hence $a_t$ itself is increasing so that (5.9) can be expressed as

$$0 \leq - A_{t_1}^{t_2} W(u) + \int_Z \int_{t_1}^{t_2} \widehat{r_0^s c(z,s,u_s) f(z,s,u_s)} \mu(dz,ds,u_s) \text{ a.s. } \mathcal{P}^u \tag{5.10}$$

for every admissible control u. Furthermore we have equality if and only if u is optimal.

## 5.3 Optimality Condition for Partial Information

Recall the following definition from [5]. A measurable function $\beta : Y \times I \times \Omega \to R$ is said to be in $L^1(\tilde{P}_u^y)$ if for fixed y, $\beta(y,\cdot)$ is a $(\mathcal{Y}_t)$-predictable process, and $E^u\left[\int_I \int_Z |\beta(y,s)| \tilde{P}_u^y(dy,ds)\right] < \infty$. $\beta$ is said to be in $L^1_{loc}(\tilde{P}_u^y)$ if there is a sequence of $(\mathcal{Y}_t)$-stopping times $T_k \uparrow T$ a.s. $\mathcal{P}^u$ such that $(\beta 1_{\{t \leq T_k\}}) \in L^1(\tilde{P}_u^y)$ for each k.

We have the following version of Theorem 4.2.

Theorem 5.1 There exists a constant $J^*$ and for every $u \in \mathcal{U}$ there exist a predictable process $(\bar{A}_0^t(u)) \in \mathcal{A}(\mathcal{Y}_t, \mathcal{P}^u)$ and a process $\beta^u \in L^1_{loc}(\tilde{P}_u^y$ so that

$$\bar{A}_0^T(u) - \int_Y \int_0^T \beta^u(y,s) P^y(dy,ds) = J^* - E^u[r_0^T J_T | \mathcal{Y}_T], \tag{5.11}$$

$$- \bar{A}_{\tau_1}^{\tau_2}(u) + \int_Z \int_{\tau_1}^{\tau_2} \overline{[\beta^u(\gamma(z),s) + r_0^s c(z,s,u_s)] f(z,s,u_s)} \mu(dz,ds,u_s) \geq 0 \tag{5.12}$$

a.s. $\mathcal{P}^u$ for $(\mathcal{Y}_t)$-stopping times $\tau_1 \leq \tau_2$ with values in I. A control law $u = u^*$ is optimal if and only if equality holds in (5.11) for deterministic times $t_1 \leq t_2$, and then, furthermore,

$$J(u^*) = J^*,$$

$$r_0^t W(u^*,t) = J^* - \bar{A}_0^t(u^*) + \int_Y \int_0^t \beta^{u^*}(y,s) P^y(dy,ds) \tag{5.13}$$

<u>Proof</u> (Necessity) Let $u \in \mathcal{U}$. We have the representation (4.5),

$$r_0^t W(u,t) = J^* - A_0^t W(u) + m^W(u,t) \tag{5.14}$$

where $A_{t_1}^{t_2} W(u)$ satisfies (5.10),

$$0 \leq - A_{t_1}^{t_2} W(u) + \iint_{Z\ t_1}^{t_2} \widehat{r_0^s c(z,s,u_s) f(z,s,u_s)} \mu(dz,ds,u_s) \tag{5.15}$$

with equality holding for $u^*$. By the martingale representation theorem [5, Theorem 3.4] there exists $\beta^u \in L^1_{loc}(\tilde{P}_u^y)$ such that

$$\begin{aligned} m^W(u,t) &= \iint_{Y\ 0}^{t} \beta^u(y,s) Q_u^y(dy,ds) \\ &= \iint_{Y\ 0}^{t} \beta^u(y,s) P^y(dy,ds) - \iint_{Y\ 0}^{t} \beta^u(y,s) \tilde{P}_u^y(dy,ds) \\ &= \iint_{Y\ 0}^{t} \beta^u(y,s) P^y(dy,ds) - \iint_{Z\ 0}^{t} \widehat{\beta^u(\gamma(z),s) f(z,s,u_s)} \mu(dz,ds,u_s) \end{aligned} \tag{5.16}$$

by (5.8). Define

$$\bar{A}_0^t(u) = A_0^t W(u) + \iint_{Z\ 0}^{t} \widehat{\beta^u(\gamma(z),s) f(z,s,u_s)} \mu(dz,ds,u_s) \tag{5.17}$$

Substitution for $m^W(u,t)$ and $A_0^t W(u)$ from (5.16), (5.17) into (5.14) and (5.15) yields (5.11), (5.12) and (5.13).

(Sufficienty) Now suppose (5.11), (5.12) holds. If we define $m^W(u,t)$ and $\bar{A}_0^t W(u)$ via (5.16) and (5.17), then (5.14) and (5.15) hold and the optimality of $u^*$ follows from Theorem 4.2.

□

Remark 5.1 (i) If we define $\Lambda^u(t) = \int_0^t \mu(Z,ds,u_s)$ then there exists a kernel $n(dz,t,u_t)$ such that $\mu(dz,dt,u_t) = n(dz,t,u_t)\Lambda^u(dt)$. Then $\Lambda^u$ can act as a time rate and so exactly as in Section 4.2 we can derive a local version of condition (5.12).

(ii) In many applications it is reasonable to suppose the existence of a probability measure $\mathcal{P}$ on $(\Omega,\mathcal{X})$ such that $\mathcal{P}^u << \mathcal{P}$ for all u. Then $\mathcal{P}^u$ can be described by specifying $\mathcal{P}$ and $L(u) = E[\frac{d\mathcal{P}^u}{d\mathcal{P}} |\mathcal{X}]$. Suppose further that $(x_t,\mathcal{P})$ is a jump process with compensating processes $(\tilde{P}^x(B,t),\mathcal{P})$, $B \in \mathcal{Z}$ given by

$$\tilde{P}^x(B,t) = \int_B\int_0^t f(z,s)\mu(dz,ds)$$

where $f(z,\cdot)$ is $(\mathcal{X}_t)$-predictable for each z and $(\mu(B,t))$ is a $(\mathcal{Y}_t)$-predictable increasing process. It can be shown then (see [6]) that for each u there is a process $\phi^u$: $Z \times I \times \Omega \to R$ such that

$$L_t(u) = E[\frac{d\mathcal{P}^u}{d\mathcal{P}} |\mathcal{X}_t]$$

is given by

$$L_t(u) = \prod_{\substack{x_{s-}\neq x_s \\ s\leq t}} [1 + \phi^u(x_s,s)]\exp[-\int_Z\int_0^t \phi^u(z,s)f(z,s)\mu(dz,ds)]$$

As a model (which satisfies the various assumptions of Section 2) we can propose that the effect of a control u is determined by the process $(L_t(u))$ above in which

$$\phi^u(z,t,\omega) = \Phi(z,t,u_t,\omega)$$

where $\Phi$: $Z \times I \times U \times \Omega \to R$ is a fixed function. The processes $(\tilde{P}^x_u(B,t), \mathcal{X}_t, \mathcal{P}^u)$ are then given by (see [6])

$$\tilde{P}^x_u(B,t) = \iint_{B\,0}^{t} [1 + \Phi(z,s,u_s)]\tilde{P}^x(dz,ds)$$

The function $\Phi$ can be interpreted as the change in the rate at which jumps occur for $\mathcal{P}^u$ as compared with $\mathcal{P}$. In terms of this special model condition (5.12) reads as

$$-\bar{A}_{\tau_1}^{\tau_2}(\omega) + \iint_{Z\ \tau_1}^{\tau_2} \overline{[\beta^u(\gamma(z),s) + r_0^s c(z,s,u_s)][1 + \Phi(z,s,u_s)]f(z,s)}\mu(dz,ds) \geq 0$$

5.4 Complete Information

We assume that $y_t \equiv x_t$. Then, as observed in Section 4.3, $W(u,t) = W(t)$ does not depend on u. However, it may appear that in the representation for $W(u,t)$ obtained in (5.13), (5.16) and (5.17), the processes $\bar{A}_0^t(u)$ and $\beta^u$ still depend on u. To see that this is not the case, consider any two controls u,v. Then

$$r_0^t\, W(t) = J^* - \bar{A}_0^t(u) + \int_Z \int_0^t \beta^u(z,s)\ P^x(dz,ds)$$

$$= J^* - \bar{A}_0^t(v) + \int_Z \int_0^t \beta^v(z,s)\ P^x(dz,ds)$$

Now $(J^* - \bar{A}_0^t(u))$ and $(J^* - \bar{A}_0^t(v))$ are $(\mathcal{X}_t)$-predictable processes whereas the integrals in the equations above are piecewise constant with discontinuities at the jump times of the $(x_t)$ process. It follows that

$$\int_Z \int_0^t \beta^u(z,s)\ P^x(dz,ds) = \int_Z \int_0^t \beta^v(z,s)\ P^x(dz,ds), \tag{5.19}$$

$$\bar{A}_0^t(u) = \bar{A}_0^t(v), \tag{5.20}$$

and so we have a considerably simpler version of Theorem 5.1.

Theorem 5.2 Suppose $y_t \equiv x_t$. Then $u^*$ is optimal if and only if there exist a constant $J^*$, a predictable process $(\bar{A}_0^t) \in \mathcal{A}(\mathcal{X}_t, \mathcal{P}^{u^*})$ and a process $\beta \in L^1_{loc}(\tilde{P}^x_{u^*})$ so that

$$\bar{A}_0^T - \iint_{Z\,0}^{T} \beta(z,s)P^x(dz,ds) = J^* - r_0^T J_T, \tag{5.21}$$

$$-\bar{A}_{t_1}^{t_2} + \iint_{Z\,t_1}^{t_2} [\beta(z,s) + r_0^s c(z,s,u_s)]f(z,s,u_s)\mu(dz,ds,u_s) \geq 0 \tag{5.22}$$

a.s. $\mathcal{P}^u$ for all $u \in \mathcal{U}$ with equality holding for $u = u^*$. Then, furthermore,

$$J^* = J(u^*),$$

$$r_0^t W(t) = J^* - \bar{A}_0^t + \iint_{Z\,0}^{t} \beta(z,s)P^x(dz,ds) \tag{5.23}$$

Suppose henceforth (see Remark 5.1) that $\tilde{P}^x_u(dz,ds)$ has the form $\tilde{P}^x_u(dz,ds) = n(dz,s,u_s)\Lambda(ds)$ for some kernel n and continuous increasing process $\Lambda(t)$ independent of u. Then, as shown in Section 4.2, we can represent

$$\bar{A}_0^t = \int_0^t \alpha_s \Lambda(ds) \tag{5.24}$$

for some $(\mathcal{X}_t)$-predictable $(\alpha_t)$. The local version of (5.21) now becomes

$$-\alpha_t + \int_Z [\beta(z,t) + r_0^t c(z,t,u_t)]n(dz,t,u_t) \geq 0 \tag{5.25}$$

for all $(t,\omega)$ with respect to $d\Lambda \times d\mathcal{P}^u$ measure, with equality when $u = u^*$.

This gives us a version of the Dynamic Programming Programming equation,

$$-\alpha_t + \operatorname*{Min}_{u\in U} \int_Z [\beta(z,t) + r_0^t c(z,t,u)] n(dz,t,u) = 0, \tag{5.26}$$

and the minimum is achieved at $u^*(t,\omega)$ for almost all $(t,\omega)$ with respect to $d\Lambda \times d\mathcal{P}^{u^*}$ measure.

We shall now use (5.23) and (5.24) to directly relate $\bar{A}$ (or equivalently $\alpha$ and $\beta$) to the process $(r_0^t W(t))$. The basic idea is to note that $\bar{A}_0^t$ on the right hand side in (5.23) is continuous whereas the integral term is piecewise constant with discontinuities occuring only at the jump times of the $(x_t)$ process. Thus the discontinuous changes in $rW$ account for $\beta$ and the continuous changes account for $\alpha$. To identify these changes we need a more detailed representation of $rW$. Set $T_0 \equiv 0$ and let $T_1 < T_2 < \cdots$ be the jump times of $x$ defined by

$$T_{k+1}(\omega) = \inf\{t > T_k(\omega) \mid x_t(\omega) \neq x_{T_k}(\omega)\},\ k = 0,1,..$$

It is shown in [5] that $\mathcal{X}_t = \sigma(x_{T_k \wedge t}, T_k;\ 0 \le k < \infty)$, $\mathcal{X}_{T_n+} = \mathcal{X}_{T_n} = \sigma(x_{T_k}, T_k;\ 0 \le k \le n)$. Since $(r_0^t W(t))$ is adapted to $(\mathcal{X}_t)$, therefore there exist functions $w_k(t,t_0,z_0,\ldots,t_k,z_k)$, measurable in their arguments, so that

$$r_0^t W(t) = \sum_{k\ge 0} 1_{\{T_k \le t < T_{k+1}\}} w_k(t,T_0,x_{T_0},\ldots,T_k,x_{T_k}) \tag{5.27}$$

The discontinuities of $rW$, which occur only at the $T_k$'s, can now be identified as

$$r_0^{T_k} W(T_k) - r_0^{T_k-} W(T_k-) = w_k(T_k,T_0,x_{T_0},\ldots,T_k,x_{T_k}) - w_{k-1}(T_k,T_0,x_{T_0},\ldots,T_{k-1},x_{T_{k-1}}) \tag{5.28}$$

Hence the function $\beta$ can be related to rW by

$$\beta(z,t) = \sum_{k\geq 0} 1_{\{T_k<t\leq T_{k+1}\}}\, b_k(z,t) \tag{5.29}$$

where, from (5.28),

$$b_{k-1}(z,t) = w_k(t,T_0,x_{T_0},\ldots,T_{k-1},x_{T_{k-1}},t,z) - w_{k-1}(t,T_0,x_{T_0},\ldots,T_{k-1},x_{T_{k-1}}) \tag{5.30}$$

To obtain the relation between $\alpha$ and W recall that we have supposed $\tilde{P}^x_u(dz,ds) = n(dz,s,u_s)\Lambda(ds)$ where $\Lambda(t)$ is increasing and continuous. Now suppose further that $\Lambda$ is absolutely continuous i.e.,

$$\Lambda(ds) = \lambda(s)\,ds \tag{5.31}$$

for some non-negative process $\lambda$. Then, using [18, VI Theorem 21], $\alpha(t)\lambda(t)$ can be identified by

$$\mathcal{L}^u_t[rW] = -\alpha(t)\lambda(t) + \int_Z \beta(z,t)n(dz,t,u_t) \tag{5.32}$$

where,

$$\mathcal{L}^u_t[rW] = \underset{h\to 0}{\text{weak lim}}\ \frac{1}{h}\{E^u[r_0^{t+h}W(t+h)\,|\,\mathcal{Y}_t] - r_0^t W(t)\} \tag{5.33}$$

(Here weak lim means limit in the $\sigma(L^1,L^\infty)$ topology). Now, by (5.27), (5.23) we get

$$\sum_{k\geq 0} 1_{\{T_k\leq t<T_{k+1}\}} w_k(t,T_0,\ldots,x_{T_k}) = J^* - \bar{A}_0^t + \iint_{Z\,0}^{t} \beta(z,s)P^x(dz,ds),$$

and since the integral term jumps only at the $\{T_k\}$ whereas $\bar{A}_0^t$ is absolutely continuous by (5.24) and (5.31), therefore $w_k$ is absolutely continuous in t. We can now obtain a formula for $\mathcal{L}^u_t[rW]$ as follows. We observe that in the stochastic interval $T_k \leq t < T_{k+1}$,

$$E^u[r_0^{t+h}W(t+h)|\mathcal{X}_t] - r_0^t W(t)$$

$$= [w_k(t+h,T_0,\ldots,x_{T_k}) - w_k(t,T_0,\ldots,x_{T_k})]$$

$$\times \mathcal{P}^u[x \text{ does not jump in } [t,t+h]|\mathcal{X}_{T_k},T_{k+1} > t]$$

$$+ \int_Z\int_0^h [w_{k+1}(t+s,T_0,\ldots,x_{T_k},t+s,z) - w_k(t,T_0,\ldots,T_k,x_{T_k})]$$

$$\times \mathcal{P}^u[T_{k+1}-T_k \in ds, x_{T_{k+1}} \in dz|\mathcal{X}_{T_k},T_{k+1}>t] + o(h) \tag{5.34}$$

Now,

$$\lim_{h\to 0} \mathcal{P}^u[x \text{ does not jump in } [t,t+h]|\mathcal{X}_{T_k},T_{k+1}>t] = 1, \tag{5.35}$$

$$\mathcal{P}^u[T_{k+1}-T_k \in ds, x_{T_{k+1}} \in dz|\mathcal{X}_{T_k},T_{k+1}>t] = \frac{F_k^u(dz,ds)}{1 - F_k^u(Z,t-T_k)}, \tag{5.36}$$

where, by definition,

$$F_k^u(B,t) = \mathcal{P}^u[T_{k+1}-T_k \leq t, x_{T_{k+1}} \in B|\mathcal{X}_{T_k}] \tag{5.37}$$

Finally, as shown first in [7], and subsequently in [8] and [15],

$$\tilde{p}^u(B,t) = \int_B\int_0^t n(dz,s,u_s)\lambda_s ds$$

$$= \sum_{i=0}^{k-1} \left[ \int_0^{T_{i+1}-T_i} \frac{F_i^u(B,ds)}{1-F_i(Z,s)} \right] + \int_0^{t-T_k} \frac{F_k(B,ds)}{1-F_k(Z,s)} \text{ for } T_k \leq t < T_{k+1} \tag{5.38}$$

From (5.33) - (5.38) we obtain

$$\mathcal{L}_t^u[rW] = \frac{\partial w_k}{\partial t}(t,T_0,..,x_{T_k}) + \int_Z [w_{k+1}(t,T_0,..,x_{T_k},t,z)$$

$$- w_k(t,T_0,..,x_{T_k})]n(dz,t,u_t)\lambda_t \text{ for } T_k \leq t < T_{k+1} \tag{5.39}$$

Substituting from (5.32) and (5.39) into (5.26) gives the next result.

<u>Theorem 5.3</u> Suppose $y_t \equiv x_t$ and suppose that for $u \in \mathcal{U}$

$$\tilde{P}_u^x(dz,ds) = n(dz,s,u_s)\lambda_s ds \tag{5.40}$$

Then $u^*$ is optimal if and only if there exist measurable functions $w_k$ $(t,t_0,..,t_k,z_k)$, which are absolutely continuous in t, so that

$$\frac{\partial w_k}{\partial t}(t,T_0,..,x_{T_k}) + \underset{u \in \mathcal{U}}{\text{Min}}\ \lambda_t \int_Z [w_{k+1}(t,T_0,..,x_{T_k},t,z) - w_k(t,T_0,..,x_{T_k})$$

$$+ r_0^t c(z,t,u)]n(dz,t,u) = 0, \text{ for } T_k \leq t < T_{k+1}, \tag{5.41}$$

$$w_k(T,T_0,..,x_{T_k}) = J_T, \text{ for } T_k \leq T < T_{k+1} \tag{5.42}$$

and the minimum in (5.40) is achieved at $u = u^*(t,\omega)$ a.s. $\mathcal{P}^{u^*}$. Then, furthermore

$$r_0^t W(t) = \sum_{k \geq 0} 1_{\{T_k \leq t < T_{k+1}\}} w_k(t,T_0,..,x_{T_k}). \tag{5.43}$$

We are now in a position to compare our results with those of Rishel [21]. First of all his model of the dynamics of the jump processes is a special case of the one used in Theorem 5.3. Secondly, the observation σ-fields, $\mathcal{Y}_t$, that he permits are much more general even than those of Theorem 5.1. For he only requires that $(\mathcal{Y}_t)$ be "locally increasing" i.e., for each t there is $\eta>0$ so that $\mathcal{Y}_t \subset \mathcal{Y}_s$ for $s \in [t,t+\eta]$. Thirdly, the structure of the cost functional is the same as the one used here. For an admissible control u let $J_t(u) = E\{\text{cost incurred in } [t,T] \text{ using } u | \mathcal{X}_t\}$. The process $(J_t(u),\mathcal{X}_t,\mathcal{P}^u)$ can be expressed as

$$J_t(u) = \sum_{k\geq 0} 1_{\{T_k\leq t<T_{k+1}\}}\, j_k^u(t,T_0,..,T_k,x_{T_k})$$

as in (5.27). Rishel derives differential equations for the $j_k^u$ similar to our equation (5.39). Finally he compares $J_t(u^*)$, for an optimal control $u^*$ with $J_t(v)$ where v is a control obtained from $u^*$ by a local perturbation. The necessary condition $E[J_t(u^*) - J_t(v)|\mathcal{Y}_t] \leq 0$ is translated into a necessary condition on the $j_k^u$ (see [21], Theorem 6). Since $u^*$ is compared with controls obtained by a local perturbation, therefore these necessary conditions are fairly weak as compared say with Theorem 5.3 above.

### 5.5 Markovian Case

To simplify the discussion in this section we suppose $T < \infty$ and $r \equiv 1$. Now suppose $y_t \equiv x_t$ and suppose as in (5.40) that

$$\tilde{P}_u^x(dz,ds) = n(dz,s,u_s)\lambda_s ds,$$

where n and λ have the form

$$n(dz,s,u,\omega) = n(dz,s,u,x_{s-}(\omega)), \tag{5.44}$$

$$\lambda(s,\omega) = \lambda(s,x_{s-}(\omega)) \tag{5.45}$$

Similarly, suppose that in the cost functional (5.5) we have

$$c(z,s,u,\omega) = c(z,s,u,x_{s-}(\omega)), \tag{5.46}$$

$$J_T(\omega) = J_T(x_T(\omega)) \tag{5.47}$$

Next, call $u \in \mathcal{U}$ a <u>Markovian control</u> if $u_t$ is of the form $u_t(\omega) = u_t(x_{t-}(\omega))$. Let $\mathcal{U}^M$ be the set of Markovian controls. We assume <u>M</u>

<u>M</u>. If $u \in \mathcal{U}^M$, then $(x_t, \mathcal{X}_t, \mathcal{P}^u)$ is a Markov process.

With these assumptions it is reasonable to expect that a Markovian control is optimal in the class $\mathcal{U}$ i.e.,

$$\bigwedge_{u \in \mathcal{U}^M} J(u) = \bigwedge_{u \in \mathcal{U}} J(u), \tag{5.48}$$

and it will then follow that the (complete information) value function $W(t)$ has a representation $W(t,\omega) = w(t,x_t(\omega))$.

To prove this assertion we begin by defining the Markovian value function. For u, v in $\mathcal{U}^M$, as before let

$$\begin{aligned}\psi(u,v,t) &= E^{(u,v,t)}\left\{\int_Z\int_t^T c(z,s,v_s)\tilde{P}_v^x(dz,ds) + J_T \middle| \mathcal{F}_t^x\right\} \\ &= E^v\left\{\int_Z\int_t^T c(z,s,v_s)\tilde{P}_v^x(dz,ds) + J_T \middle| x_t\right\} \text{ by } \underline{M}, \\ &= \eta(v,t,x_t) \text{ say,}\end{aligned} \tag{5.49}$$

$$V(t,x_t) = \bigwedge_{v\in \mathcal{U}^M} \eta(v,t,x_t)$$

To show that $V(t,x_t) = W(t)$ it is enough, as we will see, to prove a version of the optimality principle, Theorem 4.1, for the function V and $u \in \mathcal{U}^M$. It is here that we face a difficulty because the proof of Theorem 4.1 relies on Lemma 4.1 and in the proof of the latter critical use is made of the fact that $u_t$ can depend arbitrarily on $\mathcal{Y}_t$ and that these are increasing; whereas here $u_t$ can depend arbitrarily only on $\sigma(x_{t-})$ and these are certainly not increasing.

We shall circumvent this difficulty by assuming that it is possible to approximate the time-continuous optimal control problem by a time-discrete problem. Since for the latter an optimality principle is available, we will be able to obtain such a result for the original problem.

For each $t \in I$ and integer N let $t = t_1 < t_2 <\cdots< t_{2^N} = T$ be equally spaced instances of time and let $\mathcal{U}_t^N$ be the set of all $(u_s)$, $s \geq t$ of the form

$$u_s(\omega) = u_s(x_{t_k}(\omega)) \quad \text{for} \quad t_k<s\leq t_{k+1}$$

We impose the following assumption of approximation.

$\underline{A}_1$ For all $\varepsilon>0$, $t \in I$, $u \in \mathcal{U}^M$ there exists K such that for all $N\geq K$ there exists $v \in \mathcal{U}_t^N$ with $\eta(v,t,x_t) \leq \eta(u,t,x_t) + \varepsilon$

$\underline{A}_2$ For all $\varepsilon>0$, $t \in I$ there exists K such that for all $N\geq K$, $v \in \mathcal{U}_t^N$ there exists $u \in \mathcal{U}^M$ with $\eta(u,t,x_t) \leq \eta(v,t,x_t) + \varepsilon$

Theorem 5.4 Suppose (5.44) - (5.47) and $\underline{M}$ hold. Then for $t_1\leq t_2$ in I and $u \in \mathcal{U}^M$ we have

$$V(t_1,x_{t_1}) \leq E^u\left[\iint_Z\int_{t_1}^{t_2} c(z,s,u_s)\tilde{P}_u^x(dz,ds)\,|\,x_{t_1}\right] + E^u\left[V(t_2,x_{t_2})\,|\,x_{t_1}\right], \tag{5.50}$$

$$V(T,x_T) = J_T(x_T) \tag{5.51}$$

If equality holds in (5.50) for $u = u^*$ then $u^*$ is optimal in $\mathcal{U}^M$. Finally if $\underset{\sim}{A}_1$, $\underset{\sim}{A}_2$ hold, then this condition is necessary for optimality.

Proof. For $u \in \mathcal{U}^M$ we have

$$V(t_1,x_{t_1}) \leq E^u\left[\iint_Z\int_{t_1}^{t_2} c(z,s,u_s)\tilde{P}_u^x(dz,ds)\,|\,x_{t_1}\right] + \bigwedge_{v\in\mathcal{U}^M} E^u\left[\eta(v,t_2,x_{t_2})\,|\,x_{t_1}\right]$$

with equality if and only if u is optimum. Since obviously

$$\bigwedge_{v\in\mathcal{U}^M} E^u\left[\eta(v,t_2,x_{t_2})\,|\,x_{t_1}\right] \geq E^u\left[\bigwedge_{v\in\mathcal{U}^M} \eta(v,t_2,x_{t_2})\,|\,x_{t_1}\right], \tag{5.52}$$

therefore the sufficiency part of the assertion follows. Now suppose $\underset{\sim}{A}_1$, $\underset{\sim}{A}_2$ hold. To prove the final assertion it is enough to show that the reverse inequality holds in (5.52). Fix $\varepsilon>0$. We must show that there is $v \in \mathcal{U}^M$ so that

$$E^u\left[\eta(v,t_2,x_{t_2})\,|\,x_{t_1}\right] \leq E^u\left[\bigwedge_{vt\mathcal{U}^M} \eta(v,t_2,x_{t_2})\,|\,x_{t_1}\right] + \varepsilon \tag{5.53}$$

Choose K so large that the inequality in $\underset{\sim}{A}_2$ holds for $N \geq K$ and for $\frac{\varepsilon}{3}$ and $t_2$. Then, using $\underset{\sim}{A}_1$, choose $N \geq K$ so that

$$E^u\left[\bigwedge_{v\in\mathcal{U}_{t_2}^N} \eta(v,t_2,x_{t_2})\,|\,x_{t_1}\right] \leq E^u\left[\bigwedge_{v\in\mathcal{U}^M} \eta(v,t_2,x_{t_2})\,|\,x_{t_1}\right] + \frac{\varepsilon}{3} \tag{5.54}$$

Next, we apply discrete backwards dynamic programming to obtain $v' \in \mathcal{U}^N_{t_2}$ so that

$$\eta(v',t_2,x_{t_2}) \leq \bigwedge_{v\in \mathcal{U}^N_{t_2}} \eta(v,t_2,x_{t_2}) + \frac{\varepsilon}{3} \tag{5.55}$$

Finally, using $\underline{A}_2$, we may find $v \in \mathcal{U}^M$ so that

$$\eta(v,t_2,x_{t_2}) \leq \eta(v',t_2,x_{t_2}) + \frac{\varepsilon}{3} \tag{5.56}$$

From (5.54) - (5.56) we see that v satisfies (5.53). The assertion is proved.

Now let $V_t = V(t,x_t)$. Fix $u \in \mathcal{U}^M$ and consider the process $(V_t, \mathcal{X}_t, \mathcal{P}^u)$. Then, using the same argument which led to (4.5), we obtain the representation

$$V_t = J_M - A_0^t(u) + m^V(u,t), \tag{5.57}$$

where $J_M = \inf\{J(u)\,|\,u \in \mathcal{U}^M\}$, $m^V(u) \in \mathcal{M}(\mathcal{X}_t,\mathcal{P}^u)$, and for $t_1 \leq t_2$

$$A_{t_1}^{t_2} = \text{weak}\lim_{h\to 0} \int_{t_1}^{t_2} \frac{1}{h} E^u[V_s - V_{s+h}|\mathcal{X}_s]ds$$

By $\underline{M}$ $E^u[V_s-V_{s+h}|\mathcal{X}_s] = E^u[V_s-V_{s+h}|x_s]$ so that $A_{t_1}^{t_2}$ is measurable with respect to $\mathcal{X}_{t_1}^{t_2} = \sigma(x_s;\ t_1\leq s \leq t_2)$. (This implies also that $m(u,t_2) - m(u,t_1)$ is $\mathcal{X}_{t_1}^{t_2}$-measurable i.e., $m(u)$ is an additive functional of the Markov process $(x_t,\mathcal{X}_t,\mathcal{P}^u)$.) We therefore obtain the following version of Theorem 4.2.

<u>Theorem 5.5</u> Suppose (5.44) - (5.47) and $\underline{M}$ hold. There exists a constant $J_M$ and for every $u \in \mathcal{U}^M$ there exists a predictable process $(A_0^t(u)) \in \mathcal{A}(\mathcal{X}_t,\mathcal{P}^u)$ such that

$$E^u A_0^T(u) = J_M - E^u J_T, \tag{5.58}$$

and such that for $t_1 \leq t_2$ $A_{t_1}^{t_2}(u)$ is $\mathcal{X}_{t_1}^{t_2}$-measurable and

$$E^u\left[- A_{t_1}^{t_2}(u) + \int_Z\int_{t_1}^{t_2} c(z,s,u_s)\tilde{P}_u^x(dz,ds)\,|x_{t_1}\right] \geq 0 \text{ a.s. } \mathcal{P}^u \tag{5.59}$$

Suppose equality holds in (5.59) for some $u = u^*$ in $\mathcal{U}^M$. Then $u^*$ is optimal in $\mathcal{U}^M$, $J(u^*) = J_M$ and

$$V_t = E^{u^*}\left[A_t^T(u^*) + J_T|x_t\right] \text{ a.s. } \mathcal{P}^{u^*} \tag{5.60}$$

Finally if $\underline{A}_1$ and $\underline{A}_2$ hold, then this condition is necessary for optmality.

We return to the representation (5.57). Since $m^V(u)$ is an additive functional it can be represented as

$$m^V(u,t) = \int_Z\int_0^t \beta^u(z,s)\left[P^x(dz,ds) - \tilde{P}^x(dz,ds)\right]$$

where $\beta^u \in L^1_{loc}(\tilde{P}_u^x)$ is of the form

$$\beta^u(z,s,\omega) = \beta^u(z,s,x_{s-}(\omega)).$$

As before (cf.(5.17)) let

$$\bar{A}_0^t(u) = A_0^t(u) + \int_Z\int_0^t \beta^u(z,s)\tilde{P}_u^x(dz,ds)$$

$$= A_0^t(u) + \int_Z\int_0^t \beta^u(z,s)n(dz,s,u_s)\lambda_s ds$$

and we may conclude again (see (5.19), (5.20)) that for u, v in $\mathcal{U}^M$

$$\bar{A}_0^t(u) = \bar{A}_0^t(v) = \bar{A}_0^t \text{ say}$$

$$\int_Z\int_0^t \beta^u(z,s)P^x(dz,ds) = \int_Z\int_0^t \beta^v(z,s)P^x(dz,ds)$$

Furthermore there exists a predictable process $(\alpha_t)$ such that

$$\bar{A}_0^t = \int_0^t \alpha_s\lambda_s ds$$

But $\bar{A}_{t_1}^{t_2}$ is $\mathcal{X}_{t_1}^{t_2}$-measurable and $\lambda_t(\omega) = \lambda(t,x_{t-}(\omega))$ by (5.45). Hence $\alpha_t$ is of the form $\alpha_t(\omega) = \alpha(t,x_{t-}(\omega))$. The local version of (5.59) now becomes (cf. (5.26))

$$-\alpha(t,x_{t-}(\omega)) + \underset{u\in\mathcal{U}}{\text{Min}} \int_Z [\beta(z,t) + c(z,t,u)]n(dz,t,u) = 0 \tag{5.61}$$

and the minimum is achieved at $u^*(t,x_{t-}(\omega))$ for almost all $(t,\omega)$ with respect to $d\Lambda \times d\mathcal{P}^{u^*}$ measure. But from (5.61) it is evident that $u^*$ is now an optimal control in the class $\mathcal{U}$ of all controls and not just Markovian controls. It follows then that $V(t,x_t) = W(t)$. Theorem 5.3 simplifies as follows.

<u>Theorem 5.6</u> Suppose (5.44) - (5.47) and <u>M</u> hold. Suppose there exist $u^* \in \mathcal{U}^M$ and a measurable function V(t,x) which is absolutely continuous in t, so that

$$\frac{\partial V}{\partial t}(t,x_{T_k}) + \underset{u\in\mathcal{U}}{\text{Min}}\ \lambda(t,x_{T_{k-}}) \int_Z [V(t,z) - V(t,x_{T_k}) + c(t,z,u)]n(dz,t,u) = 0,$$

$$\text{for } T_k \le t < T_{k+1}, \tag{5.62}$$

$$V(T,x_{T_k}) = J_T(x_T(\omega)), \quad \text{for} \quad T_k \leq T < T_{k+1}, \tag{5.63}$$

and the minimum is achieved at $u^*(t,x_{t-}(\omega))$ a.s. with respect to $\mathcal{P}^{u^*}$ measure. Then $u^*$ is optimal in the class of all control laws, and furthermore $V(t,x_t) = W(t)$. Finally if $\underline{A}_1$ and $\underline{A}_2$ hold, then this condition is also necessary for optimality.

We can compare the result above with the main result of Stone [24, Theorem 4.5]. Essentially our result is a special case of his result since the latter applies to semi-Markov processes and not just to Markov processes as we have insisted. Of course it is possible to obtain his result from ours by imbedding the semi-Markov process into a Markov process (see [24, Theorem 2.1]). One difference may be worth noting. Stone only considers controls which give rise to Markov processes with stationary transition probabilities; he is then able to use the infinitesimal generator of the process as the main tool of analysis. The martingale theoretic approach followed here permits us to dispense with the stationarity restriction.

## 6. EXAMPLES

We use the results derived above to solve some simple optimal control problems.

### 6.1 Queues

(i) The simplest case imaginable is that of controlling the rate of a counting process over the interval $I = [0,T]$, $T < \infty$. Z is then the set of non-negative integers. Let $U = [a,b]$ where $b>a\geq 0$. Let $P^x(t)(\omega)$ = number of jumps of $x_s(\omega)$ in the interval $[0,t]$. Suppose $y_t \equiv x_t$, and for $u \in \mathcal{U}$ let

$$\tilde{p}^x_u = u_t$$

so that the controller can vary the rate of the process $(x_t)$ to any desired value in $[a,b]$. Suppose $r_0^t \equiv 1$ and $c(t) = c(t,u_t,x_{t-})$, $J_T = J(x_{T-})$. Then the optimal control must be Markovian. The optimality criterion becomes

$$0 = \min_{a\leq u\leq b} \{ \frac{\partial W}{\partial t}(t,z) + [W(t,z+1) - W(t,z) + c(t,z,u)]u\}, \quad z=0,1,.. \tag{6.1}$$

with the boundary condition

$$W(T,z) = J(z) \tag{6.2}$$

One possible problem of this type, suggested by Professor D. Snyder, related to minimizing the damage to a sample in electron microscopy, is to seek u to maximize $\mathcal{P}^u(x_T = k)$ where k is a fixed integer. Since

$$\mathcal{P}^u(x_T = k) = E^u(1_{\{x_T=k\}}),$$

and since we are maximizing the optimality criterion can be rewritten (setting $\hat{W} = -W$) as

$$0 = \max_{a \leq u \leq b} \{\frac{\partial \hat{W}}{\partial t}(t,z) + [\hat{W}(t,z+1) - \hat{W}(t,z)]u\} \tag{6.3}$$

$$\hat{W}(T,z) = 1_{\{z=k\}} \tag{6.4}$$

(6.3) gives the optimal Markovian control,

$$\begin{aligned} u^*(t,z) &= b \quad \text{if} \quad \hat{W}(t,z+1) - \hat{W}(t,z) > 0, \\ &= a \quad \text{if} \quad \hat{W}(t,z+1) - \hat{W}(t,z) < 0, \end{aligned}$$

which upon substitution in (6.3) yields,

$$\begin{aligned} 0 = \frac{\partial \hat{W}}{\partial t}(t,z) &+ b \max [\hat{W}(t,z+1) - \hat{W}(t,z), 0] \\ &+ a \min [\hat{W}(t,z+1) - \hat{W}(t,z), 0] \end{aligned}$$

for $0 \leq t \leq T$, and $z = 0,1,2,\ldots,$ and which can be solved recursively.

Remark 6.1. Suppose there were a second, independent Poisson process $(N_t)$ and suppose the objective was to maximize

$$\mathcal{P}^u(x_T + N_T = k)$$

Suppose $(N_t)$ cannot be observed or controlled, whereas $(x_t)$ can, just as before. This is now a problem with partial information. Nevertheless, it is easy to see the optimality equation (6.3) is still valid here, with the boundary condition (6.4) replaced by

$$W(T,z) = e^{-T} \frac{T^{k-i}}{(k-i)!} \qquad \text{for} \quad z = i,\ i = 0,.,\ldots,k$$

$$= 0 \qquad \text{for} \quad z > k$$

This follows from the fact that

$$\mathcal{P}^u(x_t + N_t = k) = \sum_{i=0}^{k} \mathcal{P}^u(x_t = i)\, \mathcal{P}(N_t = k - i).$$

(ii) Consider the simplest problem of controlling a queue length by varying the service rate (or number of servers). The $(x_t)$ process is now the queue length $(Q_t)$ defined as follows. Let $(A_t)$, $(D_t)$ respectively represent the arrivals and departures. Then $Q_t$ is defined by

$$Q_t = A_t - \int_0^t 1_{\{Q_{t-}>0\}}\, dD_s,$$

where the integrand manifests the fact that no departure can occur when the queue is empty. Now suppose that the arrival rate is a constant $\lambda$ which cannot be controlled, but that the departure rate can be set to any $u \in U = \{0,1,\ldots,N\}$. Then, in the notation of Section 5.5,

$$\hat{p}_u^x(dz,dt,Q_{t-}) = 1_{\{Q_{t-}+1\in dz\}}\lambda dt + 1_{\{Q_{t-}-1\in dz\}}\, 1_{\{Q_{t-}>0\}}\, u_t dt$$

where the first term on the right corresponds to a jump of $+1$ in $Q$ and the second term corresponds to a jump of $-1$.

Suppose the cost function is of the form

$$J(u) = E^u\left[\int_0^T c(s,u_s,Q_{s-})ds + f(Q_{T-})\right]$$

Then there is a Markovian optimal control and the value function W(t,Q) satisfies

$$0 = \min_{u \in U} \left\{\frac{\partial W}{\partial t}(t,Q) + c(t,u,Q) + [W(t,Q+1) - W(t,Q)]\lambda + \right.$$

$$\left. + [W(t,Q-1) - W(t,Q)]\, 1_{\{Q>0\}}\, u\right\}, \qquad (6.5)$$

with boundary condition

$$W(T,Q) = f(Q) \qquad (6.6)$$

Next, suppose that the cost is a linear function of the total waiting time and the total service time, i.e.,

$$c(t,u,Q) = au + Q, \quad \underline{f = 0}$$

where $a > 0$ is a constant. Hence from (6.5) the optimal control is "bang-band". It can be exactly specified as

$$u^*(t,Q_{t-}) = N\, 1_{\{Q_{t-}>0\}} \qquad \text{for} \quad t \in [0,T-a]$$

$$= 0 \qquad \text{for} \quad t \in [T-a,T]$$

This follows because in the interval [T-a,T]

$$W(t,Q) = (T-t)Q + \frac{\lambda}{2}(T-t)^2$$

so that

$$W(t,Q-1) - W(t,Q) = -(T-t)$$

$$< a, \qquad \text{for } t \in (T-a,T)$$

and the fact that $W(t,Q-1) - W(t,Q)$ must increase with t.

(iii) A somewhat more complicated problem is that in which only one of two queues can be served at any given time (e.g. traffic light at an intersection). Each of the two queues, $Q_t^1$, $Q_t^2$ say, are described as above, and the possible values of the pair of service rates $u = (u^1,u^2) \in U = \{(0,1),(1,0)\}$.

$$\begin{aligned} 0 = \min_{u \in U} \{ & c(t,u^1,u^2,Q^1,Q^2) + \frac{\partial W}{\partial t}(t,Q^1,Q^2) + \\ & + [W(t,Q^1+1,Q^2) - W(t,Q^1,Q^2)]\lambda^1 \\ & + [W(t,Q^1,Q^2+1) - W(t,Q^1,Q^2)]\lambda^2 \\ & + [W(t,Q^1-1,Q^2) - W(t,Q^1,Q^2)]\, 1_{\{Q^1>0\}}\, u^1 \\ & + [W(t,Q^1,Q^2-1) - W(t,Q^1,Q^2)]\, 1_{\{Q^2>0\}}\, u^2 \} \end{aligned} \tag{6.7}$$

with the boundary condition

$$W(T,Q^1,Q^2) \equiv 0 \tag{6.8}$$

We are unable to derive an explicit form for the optimal control.

6.2 Investment. An example of a jump process with an infinite number of sizes of jump is the following. Assume that there are N stocks with $\pi_i(t)$ as the price of the i-th stock. The i-th price changes at random times with a rate $\lambda_i$ and at these times the price changes from $\pi_i(t-)$ to $\pi_i(t) = \pi_i(t-) + \alpha_i(t)\,\pi_i(t-)$ where $\alpha_i(t) \geq -1$ is a random variable with distribution function $n_i(d\alpha_i,t)$. Then an investor, with wealth $K(t)$, who has invested a fraction $k_i(t)$ in the stock i, faces the accounting equation

$$dK_t = \sum_{i=1}^{N} k_i(t)\ K_{t-} \frac{d\pi_i(t)}{\pi_i(t-)}\ ,\quad \sum_{i=1}^{N} k_i(t) = 1 \tag{6.9}$$

$(K_t)$ is therefore a jump process which has jumps of size $k_i\ K\alpha_i$ occuring at rates $\lambda_i$. Here, as before, the probability measure of the "state" process $(K_t)$ depends on the "control" $(k_i(t))$, $i = 1,\ldots N$. In a simpler setting it has been shown [30] that the problem of choosing $k = \{(k_i(t))\}$ to maximize $E^k(J(K_T))$, where $J$ is the utility of wealth, can be reduced to a static optimization problem. We solve here a more general problem. Suppose the investor also has a wage income $y_t dt$ in $[t,t+dt]$, beyond his control, and can consume an amount $c_t dt$ of his wealth in the interval $[t,t+dt]$, where $c_t \geq 0$ can be chosen freely and is therefore an additional control. Then (6.9) is replaced by

$$dK_t = (y_t - c_t)dt + \sum_{i=1}^{N} k_i(t)\ K_{t-} \frac{d\pi_i(t)}{\pi_i(t-)} \tag{6.10}$$

The investor's objective is to maximize

$$E^u[\int_0^T J(t,c_t)dt + J_T(K_T)] \tag{6.11}$$

where $u = \{(k_1(t)), \ldots, (k_N(t)),(c(t))\}$ is the control, $J$ denotes utility from consumption and $J_T$ denotes utility from terminal wealth. In this formulation $(K_t)$ is no longer a jump process, because of the first term in the right-hand side of (6.10). Nevertheless, if we assume that the rate process $(\lambda_i(t))$ and the distributions $(n_i)$ depend only upon $K_{t-}$ then $(K_t, \rho^u)$ is still a Markov process for a Markov control $u$ and the infinitesimal generator $\mathcal{L}^u[\hat{W}]$ of the value function

$$\hat{W}(t,K_t) = \sup_u E^u[\int_t^T J(t,c_t)dt + J_T(K_T)]$$

is, from (6.10)

$$\underset{h \to 0}{\text{w.lim}} \frac{1}{h} \{E^u[\hat{W}(t+h,K_{t+h}) \mid K_t] - \hat{W}(t,K_t)\}$$

$$= \frac{\partial \hat{W}}{\partial t}(K_t,t) + \frac{\partial \hat{W}}{\partial K}(K_t,t)\ (y_t - c_t) + \sum_{i=1}^{N} \lambda_i(t,K_t) \int_{-1}^{\infty} [\hat{W}(t,[1 + \alpha_i k_i(t)]K_t) - \hat{W}(t,K_t)]n_i(d\alpha_i,t,K_t) \tag{6.12}$$

(We could have permitted a Brownian motion component in (6.10) as studied in [17]).

The optimality criterion is

$$0 = \max_{u \in U} \{J(t,c_t) + \frac{\partial \hat{W}}{\partial t}(t,K) + (y_t - c_t)\frac{\partial \hat{W}}{\partial K}(t,K) + \sum_{i=1}^{N} \lambda_i(t,K) \int_{-1}^{\infty} [\hat{W}(t,[1 + \alpha_i k_i]K) - \hat{W}(t,K)]\ n_i(d\alpha_i,t,K)\} \tag{6.13}$$

with the boundary condition

$$\hat{W}(T,K) = J_T(K) \tag{6.14}$$

We can solve (6.13), (6.14) for the following special case. Assume $y_t \equiv 0$, $J(t,c) = \frac{c^\gamma}{\gamma}$, $J_T(K) = a\frac{K^\gamma}{\gamma}$, where $a > 0$ and $0 < \gamma \leq 1$ are constants, and $\lambda_i, n_i$ independent of K and t. Then (6.13), (6.14) have the following solution

$$\hat{W}(t,K) = f(t)\frac{K^\gamma}{\gamma}, \quad 0 \leq t \leq T,\ K \geq 0,$$

where,

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

The constant $A$ and the optimal control are given by

$$c_t^* = \frac{K_t}{f(t)^{1/1-\gamma}} ,$$

$(k_i^*(t))$ are optimal solutions of the static problem

$$\operatorname*{Max}_{k_i \geq 0, \Sigma k_i = 1} \{\sum_{i=1}^{N} \lambda_i \int_{-1}^{\infty} [(1 + k_i \alpha_i)^{\gamma} - 1] \, n_i(d\alpha_i)\} = A$$

<u>Remark 6.2</u>. As the referee has pointed out to us it is possible to regard the process K of this example as a jump process by taking Z to be a space of functions. The value between two successive jump times would then be the trajectory of K during these two instants of time.

<u>Acknowledgement</u>. This paper was extensively rewritten in the light of a thorough critique of the previous version by Dr. Pierre Brémaud. We are very grateful to him for his interest.

## 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. J. M. Bismut: Integrales Convexes et Probabililes, Institut de Recherche d'Informatique et d'Antomatique, INF. 7201/72001, Rocquencourt, France, 1972.

4. R. Boel: Optimal Control of Jump Processes, Electronics Research Laboratory, University of California, Berkeley, Memo ERL M-448, September 1974.

5. R. Boel, P. Varaiya and E. Wong: Martingales on Jump Processes I: Representation Results, Electronics Research Laboratory, University of California, Berkeley, Memo ERL M-407, September 1973, submitted to SIAM J. Control.

6. R. Boel, P. Varaiya and E. Wong: Martingales on Jump Processes II: Applications, Electronics Research Laboratory, University of California, Berkeley, Memo ERL M-409, December 1973, submitted to SIAM J. Control.

7. C.S. Chou and P.A. Meyer: Sur la Représentation des Martingales comme Intégrales Stochastiques dans les Processus Ponctuels, 1974, to appear in: Séminaire de Probabilités VIII, Lecture Notes in Mathematics, Springer Verlag.

8. M.H.A. Davis: The Representation of Martingales of Jump Processes, unpublished, 1974.

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

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

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

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

13. R.J. Elliott: Martingales of a Jump Process and Absolutely Continuous Changes of Measure, presented at Symposium on Stochastic Systems, University of Kentucky, Lexington, 1975.

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

15. J. Jacod: Multivariate Point Processes: Predictable Projection, Radon-Nikodym Derivatives, Representation of Martingales, Z. Wahr. verw. Geb., 31, 235-253, 1975.

16. H. Kushner: Stochastic Stability and Control, Academic Press, 1967.

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

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

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

20. R. Rishel: Necessary and Sufficient Dynamic Programming Conditions for Continuous-Time Stochastic Optimal Control, SIAM J. Control, 559-571, 1970.

21. R. Rishel: A Minimum Principle for Controlled Jump Processes, in A. Bensonssan and J.L. Lions (eds.), Control Theory, Numerical Methods and Computer Systems Modelling, Lecture Notes in Economics and Mathematical Systems, 107, 493-508, 1975.

22. R. Rishel: Dynamic Programming and Minimum Principles for Systems with Jump Markov Disturbances, SIAM J. Control 13, 338-371, 1975.

23. P.A. Samuelson: Proof that properly discounted present values of assets vibrate randomly, Bell J. of Economics and Management Science 4, 2, 369-374, 1973

24. L. Stone: Necessary and Sufficient Conditions for Optimal Control of Semi-Markov Jump Processes, SIAM J. Control 11, 187-201, 1973.

25. C. Striebel: Martingale Conditions for the Optimal Control of Continuous Time Stochastic Systems, to appear, 1974.

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

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

28. J. van Schuppen: Estimation Theory for Continuous Time Processes, a Martingale Approach, Electronics Research Laboratory, University of California, Berkeley, Memo ERL M-405, September 1973.

29. J. van Schuppen and E. Wong: Transformation of Local Martingales under a Change of Law, Electronics Research Laboratory, University of California, Berkeley, Memo ERL M-385, May 1973.

30. P. Varaiya: The Martingale Theory of Jump Processes, IEEE Trans. Auto. Contr., AC-20, 34-41, 1975.