IMPULSE CONTROL OF STOCHASTIC PROCESSES

by

Hemant Vasant Desai

Memorandum No. ERL-M572

December 1975

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Research sponsored by the National Science Foundation Grant
ENG74-06651-A01.

TABLE OF CONTENTS

CHAPTER 0

INTRODUCTION

This thesis examines a class of Stochastic Control Problems where the control laws are restricted to having piecewise constant sample paths. Thus a control law can be looked upon as a choice of random times when the value of the control is changed. The system is modelled as a controlled probability space rather than as a controlled space of trajectories. This is in keeping with the recent results of Boel-Varaiya [8]. This technique of modelling the effect of control laws was first used by Benes [2,3] to prove the existence of solutions of systems driven by a Wiener Process. Subsequently it was used by Davis-Varaiya [10] to derive optimality criteria for such systems and by Boel [7] to derive optimality criteria for jump processes. The restricted class of control laws gives rise to the problem of Impulse Control. This problem has been considered for systems driven by a Wiener process by Bensoussan and Lions [4,5,6]. The results presented in the first chapter can be looked upon as the abstract versions of their results. The differences between the optimality criteria for the restricted class of control laws and for the class where no such restriction exists are most clearly brought out in the local criteria presented in Chapter one (see 1.5). These are to be compared with the local criteria of Boel-Varaiya [8].

The problem of Inventory Control is a typical example of the above type of problem. There is an extensive literature on this type of problem some of which has been referenced in Chapter 2. The basic problem which arises here is that in the general case it is very difficult to characterize the jump times of the optimal control. The use of (s,S) inventory policies circumvents this difficulty by restricting the jump times to ones that can be easily characterized (viz. exit times from very simple sets in state space). It is clear however that an (s,S) policy cannot be optimal for a finite horizon problem. In chapter two we show that the abstract optimality criteria lead to a complete characterization of the optimal policy.

Chapter 3 is devoted to showing differentiability properties of the value function obtained for the Impulse Control problem of Chapter 2. There we see that under suitable assumptions the value function is continuously differentiable in its spatial variable.

CHAPTER 1

ABSTRACT MODEL: OPTIMALITY CRITERIA

1.1. Model, Terminology, Preliminary Definitions

We fix a non-empty set Ω called the sample space. Let $I = [0,1]$ be the time interval of interest. Let $x : I \times \Omega \to \mathbb{R}^p$ be a fixed map. For each $\omega \in \Omega$, the map $t \to x(t,\omega)$ will be called the path ω of x. The collection of all paths of x manifest all possible evolutions in time of the process under consideration.

For each $t \in I$, let $\mathcal{F}_t$ be the σ-field of subsets of Ω generated by the maps $\{x_s\}_{s \leq t}$. This will be written as $\mathcal{F}'_t = \sigma\{x_s;\ s \leq t\}$. It is clear that the family of σ-fields $\{\mathcal{F}'_t\}_{t \in I}$ is an increasing family of σ-fields. i.e. $t \leq s$ implies $\mathcal{F}'_t \subseteq \mathcal{F}'_s$. The family $\{\mathcal{F}'_t\}_{t \in I}$ is called the family of complete information σ-fields. Let $\mathcal{F}'$ denote the σ-field generated by the union of the σ-fields $\mathcal{F}'_t$, $t \in I$. This will be written as $\mathcal{F}' = \vee_t \mathcal{F}'_t$.

A stopping time T of $\mathcal{F}'_t$ is an $\mathcal{F}'$-measurable map $T : \Omega \to \mathbb{R}_+$ s.t. for all $t \in I$, $\{\omega / T(\omega) \leq t\} \in \mathcal{F}'_t$. If T is a stopping time of $\mathcal{F}'_t$, let $\mathcal{F}'_T$ be the σ-field of all sets $A \in \mathcal{F}'$ s.t. $A \cap \{T \leq t\} \in \mathcal{F}'_t$ for all $t \in I$. It can be verified that the collection of all such sets do in fact form a σ-field. (See IV-D35 of Meyer [13].)

The family of σ-fields $\{\mathcal{F}'_t\}$ is said to be free of times of discontinuity if for every increasing sequence T_n of stopping times of $\{\mathcal{F}'_t\}$ we have

$$\mathcal{F}'_{\{\lim T_n\}} = \vee_n \mathcal{F}'_{T_n} .$$

For a definition and discussion on times of discontinuity of the family $\{\mathcal{F}'_t\}$ refer to VII-D39, D40, No. 54 of Meyer [13].

Let $\mathcal{Y}_t$ be a fixed (but arbitrary) increasing family of sub σ-fields of $\mathcal{F}_t$. i.e. $\mathcal{Y}_t \subseteq \mathcal{F}_t$ for all $t \in I$. We call $\{\mathcal{Y}_t\}$ the <u>family of partial information σ-fields.</u> Each $\mathcal{Y}_t$ is the observation σ-field at time t.

Throughout this thesis we <u>assume</u> that $\mathcal{F}_t$ and $\mathcal{Y}_t$ are free of times of discontinuity. Note that this is an assumption on the map x.

A map $u : I \times \Omega \to \mathbb{R}^p$ is said to be <u>$\mathcal{Y}_t$-predictable</u>. if for each t, u_t is $\mathcal{Y}_t$-measurable and there is a sequence of maps $u^n : I \times \Omega \to \mathbb{R}^p$ such that each u^n has left continuous paths and $\lim_{n\to\infty} u^n(t,\omega) = u(t,\omega)$ for all $(t,\omega) \in I \times \Omega$.

<u>A stopping time T of $\mathcal{Y}_t$ is Algebraically Predictable</u> if the indicator function of the stochastic interval [T,1] is predictable in the sense of the above definition. (See Meyer [14].)

If $\mathcal{P}$ is a probability measure on $(\Omega,\mathcal{F})$, a <u>stopping time T is predictable</u> if there exists an increasing sequence of stopping times $S_1 \leq S_2 \leq \ldots$ such that

$$\mathcal{P}\{T = 0, \text{ or } S_k < T \text{ for all } k \text{ and } \lim_{k\to\infty} S_k = T\} = 1.$$

$\mathcal{A}^+(\mathcal{Y}_t,\mathcal{F})$ is the class of all Stochastic Processes $(a_t,\mathcal{Y}_t,\mathcal{P})$, $t \in I$ such that $a_o = 0$ a.s., a_t has right continuous, nondecreasing paths and which are uniformly integrable: $\sup_t Ea_t < \infty$ · $\mathcal{A}(\mathcal{Y}_t,\mathcal{P}) = \mathcal{A}^+(\mathcal{Y}_t,\mathcal{P}) - \mathcal{A}^+(\mathcal{Y}_t,\mathcal{P})$.

The definitions and terminology introduced above are quite standard. They have been recorded for the sake of completeness.

1.2. <u>Control Laws, Cost</u>

Fix a set $U \subseteq \mathbb{R}^p$. U is the set of admissible control values.

Definition: A control law is a map $u : I \times \Omega \to U$.

Definition: If u and v are control laws and $t \in I$, the concatenation of u and v at time t (written (utv)) is the control law defined by

$$(utv) = \begin{cases} u \text{ on } [0,t] \times \Omega \\ v \text{ on } (t,1] \times \Omega. \end{cases}$$

Definition: The class $\mathcal{U}$ of admissible control laws is any class of control laws which satisfies the following conditions.

1) $\mathcal{U}$ is closed under concatenation.
2) The paths of every $u \in \mathcal{U}$ are piecewise constant, right continuous with left limits and have a finite number of jumps in a finite time interval.
3) Every $u \in \mathcal{U}$ is $\mathcal{Y}_t$-Predictable.

The assumption 1) is necessary for Dynamic Programming. The assumption 2) rules out continuous control laws. This class of control laws is a subclass of the class considered by Boel-Varaiya [8]. The assumption 2) gives rise to the problem of Impulse control. It is particularly useful for capacity programming where there is indivisibility or set up costs. It is also useful in an economic context when prices are modelled as control laws. Then institutional constraints require that prices cannot be changed in a continuous fashion.

Action of Control Laws:

The action of a control law $u \in \mathcal{U}$ is characterized by a probability measure $\mathcal{P}^u$ on $(\Omega, \mathcal{F})$ such that for each $t \in I$, $\mathcal{P}^u$ restricted to $\mathcal{F}_t$ depends only on the values of u on $[0,t]$.

Thus each control law determines a state process $(x_t, \mathcal{F}_t, \mathcal{P}^u)$ and a control process $(u_t, \mathcal{Y}_t, \mathcal{P}^u)$. Changes in control laws affect the Probability Measures on $(\Omega, \mathcal{F})$ and not the paths of the state process. This technique of Modelling the effect of control laws is implicit in the methods of proof used by Davis-Varaiya [10] for the control of Systems driven by a Wiener process. Boel [7] has used it in modelling the control problem for Jump processes. In Chapter 2 we shall encounter a problem where it is better to model the effect of changes in control laws by changing the state process. However the martingale techniques developed in this chapter still apply. For the control of Jump processes the results of Boel-Varaiya show that this model is suitable to tackle a wide class of problems.

If $u \in \mathcal{U}$, we define the jump times of u inductively by:

$$T_o^u = 0, \; T_{k+1}^u(\omega) = \text{Inf}\{t/t \geq T_k(\omega) \text{ and } u_t(\omega) \neq u_{T_k(\omega)}(\omega)\}$$

where the Infinum over the empty set is taken to be $+\infty$. Since we are only dealing with a finite time interval, this is merely a device to say that: $T_k^u(\omega) = +\infty$ means the path ω of u has no more than k-1 jumps. We note that the jump times $T_k^u(\omega)$ are defined since each path of u has only a finite number of jumps in a finite time interval.

Since each control law u is $\mathcal{Y}_t$-predictable, it follows from the results of Meyer [14] that each T_k^u is an algebraically predictable stopping time of $\mathcal{Y}_t$. From now on we assume that the σ-fields $\mathcal{Y}_t$ are computed with respect to $\mathcal{P}^u$. Then by T52 of Meyer [13] we have each T_k^u is a predictable stopping time of the σ-field $\mathcal{Y}_t$ completed with respect to $\mathcal{P}^u$.

With each control law $u \in \mathcal{U}$ we associate the process $N^u(s) = \sum_k 1_{\{s \geq T_k^u\}}$. Where $1_{\{s \geq T_k^u\}}$ is the indicator function of the set

$\{s \geq T_k^u\} = \{\omega / s \geq T_k^u(\omega)\}$.

Thus the process $N^u(s)$ counts the number of jumps of u up to time s. It is clear that for each u, $N^u(s)$ is an increasing $\mathcal{Y}_t$-adapted predictable counting process.

We are now in a position to describe the cost associated with each control law. In order to do this we suppose given the instantaneous cost function c which is a <u>bounded</u> map $c : I \times U \times \Omega \to \mathbb{R}_+$ which for each fixed $u \in U$ is $\mathcal{F}_t$-adapted and $\mathcal{B}(I) \times \mathcal{F}$-measurable on $I \times \Omega$. (where $\mathcal{B}(I)$ is the Borel σ-field on I.)

If u is a control law E^u denotes the expectation operator on $(\Omega, \mathcal{F})$ induced by the probability measure $\mathcal{P}^u$. The cost incurred by $u \in \mathcal{U}$ is

$$J(u) = E^u\{\int_0^1 c(s,u(s))\, ds + \int_0^1 dN^u(s)\}.$$

(In the interests of notational simplification we have overlooked the inclusion of a terminal cost and of a discounting rate. The equations which follow can easily be written down for these cases.)

The inclusion of the term $E^u\{\int_0^1 dN^u(s)\}$ expresses the fact that each change in the control value costs one unit. This justifies the terminology Impulse Control. In an economic context this cost could be interpreted as the cost for broadcasting changes in prices. On the other hand from a welfare point of view it says that very frequent price variations are detrimental to consumer welfare.

We are interested in discovering necessary and sufficient conditions for a control law u^* to be optimal in the sense that $J(u^*) \leq J(u)\ \forall u \in \mathcal{U}$. This end will be achieved by the methods of Dynamic Programming in the interest of which we make the following definitions.

1.3. Principle of Optimality

We now assume that the class of admissible control laws is reduced to include only control laws u such that $E^u\{\int_0^1 c(s,u(s)\ ds + \int_0^1 dN^u(s)\}$ $< \infty$. If this new class is empty then $J(u) = \infty$ for all u and there is no problem. Thus we assume that this new class is non-empty. Because of this assumption the following family of processes (Indexed by control laws u, v) is well defined and each member of the family is Integrable.

$$J(t,(utv)) = E^{utv}\{\int_t^1 c(s,v(s))\ ds + \int_t^1 dN^v(s)/\mathcal{Y}_t\}.$$

J(t,(utv)) is the cost to go from time t onwards given the information available at time t and that control law u is used up to time t and v thereafter. Because of the assumption made at the beginning of this section we have $J(t,(utv)) \in L^1(\Omega,\mathcal{Y}_t,\mathcal{P}^u)$ for each t. Since $L^1(\Omega,\mathcal{Y}_t,\mathcal{P}^u)$ is a complete lattice with respect to the natural partial ordering for real-valued functions (see IV-8-22 of Meyer [13].) the following infinum exists for each t:

$$W(u,t) = \inf_v J(t,(utv)) \text{ and } W(u,t) \in L^1(\Omega,\mathcal{Y}_t,\mathcal{P}^u).$$

W(u,t) is called the <u>value process</u> associated with u. It is clear from the definition that W(u,t) is $\mathcal{Y}_t$-adapted. If u is a control law, W(u,t) is the minimum cost to go from time t, given that control law u is used upto time t. In order to state our first Theorem we need the following definition which was first introduced by Rishel [15] and used subsequently by Davis-Varaiya [10].

<u>Definition</u>: The class $\mathcal{U}$ is $\mathcal{Y}_t$ <u>relatively complete</u> with respect to W(u,t)

if for all $u \in \mathcal{U}$, for all $T \in I$, for all $\varepsilon > 0$, there exists $v \in \mathcal{U}_t$ such that

$$J(t,(utv)) \leq W(u,t) + \varepsilon \quad \text{a.s.} \, P^u.$$

We note that $\mathcal{U}_t$ is obtained from $\mathcal{U}$ by restricting the domains of all $u \in \mathcal{U}$ to $[t,1]$.

By lemma 3.1 of Davis-Varaiya [10] we have that $\mathcal{U}$ is $\mathcal{Y}_t$-relatively complete with respect to $W(u,t)$. We can now state our first theorem.

<u>Theorem 1.3.1</u>. (Principle of Optimality)

For all $u \in \mathcal{U}$, for all $0 \leq t \leq t + h \leq 1$

$$W(u,t) \leq E^u\{\int_t^{t+h} c(s,u(s))\, ds + \int_t^{t+h} dN^u(s)/\mathcal{Y}_t\} + E^u\{W(u,t+h)/\mathcal{Y}_t\}$$

with equality if and only if u is optimal.

<u>Proof</u>: The proof is as in Theorem 2.1 of Boel [7] except for minor modifications and will not be repeated here. ⌑

We now introduce the following notation: If u is a control law, let $\hat{c}(s,u(s))$ be the process defined by $\hat{c}(s,u(s)) = E^u(c(s,u(s))/\mathcal{Y}_s)$. Then $\hat{c}(s,u(s))$ is adapted to $\mathcal{Y}_s$ for all s.

<u>Corollary 1.3.1</u>.

$$(W(u,t) + \int_0^t \hat{c}(s,u(s))\, ds + \int_0^t dN^u(s), \mathcal{Y}_t, P^u)$$ is a submartingale adapted to $\mathcal{Y}_t$ and a martingale if and only if u is optimal.

<u>Proof</u>: By definition, $W(u,t)$ is adapted to $\mathcal{Y}_t$ $\int_0^t \hat{c}(s,u(s))\, ds$ is adapted to $\mathcal{Y}_t$ since $c(s,u(s))$ is and $\mathcal{Y}_t$ is an increasing family of σ-fields. Similarly $\int_0^t dN^u(s)$ is adapted to $\mathcal{Y}_t$. It remains to show

the submartingale part of the statement. This is shown as follows:

$$E^u\{W(u,t+h) + \int_0^{t+h} \hat{c}(s,u(s))\, ds + \int_0^{t+h} dN^u(s)/\mathcal{Y}_t\} - [W(u,t) + \int_0^t \hat{c}(s,u(s))\, ds + \int_0^t dN^u(s)]$$

$$= [E^u\{W(u,t+h)/\mathcal{Y}_t\} - W(u,t)] + [E^u(\int_0^{t+h} \hat{c}(s,u(s))\, ds/\mathcal{Y}_t) - \int_0^t \hat{c}(s,u(s))\, ds] + E^u[(\int_0^{t+h} dN^u(s)/\mathcal{Y}_t) - (\int_0^t dN^u(s))] \qquad (a)$$

The second term in the above sum is:

$$E^u(\int_0^{t+h} \hat{c}(s,u(s))\, ds/\mathcal{Y}_t) - \int_0^t \hat{c}(s,u(s))\, ds$$

$$= E^u(\int_0^{t+h} E^u(c_s^u/\mathcal{Y}_s)\, ds/\mathcal{Y}_t) - \int_0^t E^u(c_s^u/\mathcal{Y}_s)\, ds$$

$$= E^u(\int_0^{t+h} c(s,u(s))\, ds/\mathcal{Y}_t) - E^u(\int_0^t c(s,u(s))/\mathcal{Y}_t)$$

$$= E^u(\int_t^{t+h} c(s,u(s))\, ds/\mathcal{Y}_t).$$

Similarly the third term is seen to be equal to $E^u(\int_t^{t+h} dN^u(s)/\mathcal{Y}_t)$

Thus the right hand side of (a) becomes

$$[E^u(W(u,t+h)/\mathcal{Y}_t) - W(u,t)] + E^u(\int_t^{t+h} c(s,u(s))\, ds/\mathcal{Y}_t) + E^u[\int_t^{t+h} dN^u(s)/\mathcal{Y}_t]$$

which is non-negative for all control laws u and zero if and only if u is optimal (by the principle of optimality). This proves the statement of the corollary. ⌑

Corollary 1.3.2.

The principle of optimality holds if t and t+h are replaced by $\mathcal{Y}_t$-stopping times T and S such that $T \leq S$.

Proof: Theorem 1.3.1 can be rederived from Corollary 1.3.1. Thus the result follows from the Optimal Sampling Theorem. ¤

We remark that the principle of optimality says that if u^* is an optimal control law, $(W(u^*,t),\mathcal{Y}_t,\mathcal{P}^{u^*})$ is a supermartingale. This prompts the following definition.

Definition: A control law u is value decreasing if $(W(u,t),\mathcal{Y}_t,\mathcal{P}^u)$ is a supermartingale. This definition has been used by Davis-Varaiya [10] Rishel [15] and Boel [7]. Far value decreasing controls we have the following result which follows from the Optimality Principle.

Corollary 1.3.3.

For all value decreasing controls u, the process $(W(u,t),\mathcal{Y}_t,\mathcal{P}^u)$ is a right Continuous Potential of class D. (upto right continuous modification).

Proof: By the Optimality Principle we have

$$|E^u[W(u,t+h)-W(u,t)]| \leq E^u\int_t^{t+h} c(s,u(s))\,ds + E^u\int_t^{t+h} dN^u(s)$$

as $h \to 0$ we have the first term on the right goes to zero since c is bounded. The second term goes to zero since u has right continuous paths. Thus we have that the map $t \to E^u W(u,t)$ is right continuous. Since u is value decreasing $W(u,t)$ is a supermartingale. Thus by VI-T4 of Meyer [13], the supermartingale $W(u,t)$ admits a right continuous modification. Also since the σ-fields $\mathcal{Y}_t$ are free of times of discontinuity every path of this modification is free of oscillitory discontinuities. Note also that $W(u,s) \geq 0$ and $W(u,t) \to 0$ as $t \to 1$ a.s. $\mathcal{P}^u$ and in $L^1(\mathcal{P}^u)$. Thus $W(u,t)$ is a class D potential. This concludes the proof of the corollary.

¤

The next corollary provides a relationship between the values of the value function W just before and just after the jump times of a value decreasing control.

Corollary 1.3.4.

1) For all value decreasing control laws u we have $W(u,T_k^u-) \leq 1 + W(u,T_k^u)$ for all $k = 0,1,\ldots$ a.s. $\mathcal{P}^u$ with equality if u is optimal.

2) For all predictable stopping times T of $\mathcal{Y}_t$ such that $T_{k-1}^u < T < T_k^u$ for some u for some k, we have $W(u,T-) < 1 + W(u,T)$.

Proof: 1) Fix a value decreasing control u. Let $\{T_k^u\}$ be its jump times. Fix k. Since T_k^u is $\mathcal{Y}_t$-predictable, there is an increasing sequence of $\mathcal{Y}_t$-stopping times $\{S_n\}$ such that $S_n \uparrow T_k^u$ a.s. $\mathcal{P}^u$. By considering $S_n \vee T_{k-1}^u$ if necessary we can assume that $S_n \geq T_{k-1}^u$. Since the optimality principle holds for stopping times of $\mathcal{Y}_t$, we have:

$$\text{For all } n,\ W(u,S_n) \leq E^u\{\int_{S_n}^{T_k^u} c(s,u(s))\,ds + \int_{S_n}^{T_k^u} dN^u(s)/\mathcal{Y}_{S_n}\}$$

$$+ E^u\{W(u,T_k^u)/\mathcal{Y}_{S_n}\}. \qquad \text{(a)}$$

Thus the inequality is preserved in the limit as $n \to \infty$. The limit of the left hand side of (a) exists and equals $W(u,T_k^u-)$ a.s. $\mathcal{P}^u$ since the paths of W are free of oscillitory discontinuities and $S_n \uparrow T_k^u$. As for the right hand side we have:

$$E^u\{\int_{S_n}^{T_k^u} c(s,u(s))\, ds/\mathcal{Y}_{S_n}\} \to 0 \quad \text{a.s. } \mathcal{P}^u (\text{Since } S_n \uparrow T_k^u)$$

and $$E^u\{\int_{S_n}^{T_k^u} dN^u(s)/\mathcal{Y}_{S_n}\} \to 1 \quad \text{a.s. } \mathcal{P}^u (\text{Since } T_k^u \text{ is a jump time of u.})$$

and $$E^u\{W(u,T_k^u)/\mathcal{Y}_{S_n}\} \to E^u\{W(u,T_k^u)/\mathcal{Y}_{T_k^u-}\}.$$

But the σ-fields $\mathcal{Y}_t$ are free of times of discontinuity and T_k^u is a predictable stopping time. Thus $\mathcal{Y}_{T_k^u-} = \mathcal{Y}_{T_k^u}$. Thus

$$E^u\{W(u,T_k^u)/\mathcal{Y}_{T_k^u-}\} = E^u\{W(u,T_k^u)/\mathcal{Y}_{T_k^u}\} = W(u,T_k^u)$$

since W is $\mathcal{Y}_t$-adapted. Putting it all together we obtain $W(u,T_k^u-) \leq 1 + W(u,T_k^u)$ a.s. $\mathcal{P}^u$. The equality for an optimal u is proved similarly noting that the optimality principle jolds with equality for an optimal control.

(2) Fix a stopping time T, a control law u and an index k satisfying the conditions of the corollary. Then there is a sequence $\{S_n\}$ of stopping times of $\mathcal{Y}_t$ such that $S_n \geq T_{k-1}^u$ and $S_n \uparrow T$. As in 1) we obtain

$$W(u,S_n) \leq E^u\{\int_{S_n}^{T} c(s,u(s))\, ds + \int_{S_n}^{T} dN^u(s)/\mathcal{Y}_{S_n}\} + E^u\{W(u,T)/\mathcal{Y}_{S_n}\}.$$

Taking limits as $n \to \infty$ we see that

$$W(u,T-) \leq W(u,T), \text{ Since } \lim_n E^u\{\int_{S_n}^{T} dN^u(s)/\mathcal{Y}_{S_n}\} = 0$$

in view of the fact that T is strictly between jump times of u. Thus certainly we have

$$W(u,T-) < 1 + W(u.T) \quad \text{a.s. } \mathcal{P}^u. \qquad \square$$

1.4. Conditions for Optimality

In this section we develop conditions which will eventually lead us to a local characterization of the value function and of the optimal control law. The argument used in the next Theorem is the same as that used in Theorem 4.2 of Boel-Varaiya [8] we repeat it here because it is used later.

Since the process $E^u\{\int_0^1 c(s,u(s))\,ds + \int_0^1 dN^u(s)/\mathcal{Y}_t\}$ is a $(\mathcal{Y}_t,\mathcal{P}^u)$ Martingale, the process $w(u,t)$ is a $(\mathcal{Y}_t,\mathcal{P}^u)$ supermartingale where

$$w(u,t) = E^u[\int_t^1 c(s,u(s))\,ds + \int_t^1 dN^u(s)/\mathcal{Y}_t] - W(u,t)$$

Now exactly as we did in Corollary 1.3.3 we can verify that $w(u,t)$ is a class D potential. Thus by The Supermartingale Decomposition Theorem of Meyer [13] (VII-T29) there is a unique predictable increasing process $A_t(u) \in \mathcal{A}^+(\mathcal{Y}_t,\mathcal{P}^u)$ which generates $w(u,t)$. That is there is a $(\mathcal{Y}_t,\mathcal{P}^u)$ martingale $m_t(u)$ such that $w(u,t)$ admits the decomposition $w(u,t) = \tilde{J}(u) - A_t(a) + m_t(u)$. Where $\tilde{J}(u) = w(u,0) = J(u) - J^*$. Furthermore the Decomposition Theorem says that $A_t(u)$ is the following weak limit in the $\sigma(L^1,L^\infty)$ topology on L^1:

$$A_t(u) = \underset{h\to 0}{\text{weak lim}} \int_0^t \frac{1}{h} E^u[w(u,s) - w(u,s+h)/\mathcal{Y}_s]\,ds$$

$$= \underset{h\to 0}{\text{weak lim}}\{\int_0^t \frac{1}{h} E^u[\int_s^{s+h} c(s,u(s))\,ds + \int_s^{s+h} dN^u(s)/\mathcal{Y}_s]\,ds$$

$$-\int_0^t \frac{1}{h} E^u[W(u,s) - W(u,s+h)/\mathcal{Y}_s]\, ds \qquad \text{(a)}$$

It follows that there is a predictable process $\gamma_t(u) \in \mathcal{A}^+(\mathcal{Y}_t,\mathcal{P}^u)$ such that

$$\gamma_t(u) = \underset{h\to 0}{\text{weak lim}} \int_0^t \frac{1}{h} E^u[\int_s^{s+h} c(s,u(s))\, ds + \int_s^{s+h} dN^u(s)/\mathcal{Y}_s] \quad \text{(b)}$$

From (a) and (b) we can conclude that there exists a predictable process $\bar{A}_t(u) \in \mathcal{A}(\mathcal{Y}_t,\mathcal{P}^u)$ viz. $\gamma_t(u) - A_t(u)$, such that

$$\bar{A}_t(u) = \underset{h\to 0}{\text{weak lim}} \int_0^t \frac{1}{h} E^u[W(u,s) - W(u,s+h)/\mathcal{Y}_s]\, ds.$$

This is sufficient to apply Meyer's Decomposition Theorem to $W(u,t)$ and we may conclude that

$$W(u,t) = W(u,0) - \bar{A}_t(u) + \bar{m}_t(u) = J^* - \bar{A}_t(u) + \bar{m}_t(u)$$

where $\bar{m}_t(u)$ is a $(\mathcal{Y}_t,\mathcal{P}^u)$ martingale and $\bar{A}_t(u) \in \mathcal{A}(\mathcal{Y}_t,\mathcal{P}^u)$. We remark that if u is value decreasing we can immediately use the Decomposition Theorem on $W(u,t)$. In that case we would obtain $\bar{A}_t(u) \in \mathcal{A}^+(\mathcal{Y}_t,\mathcal{P}^u)$. When we develop local conditions, it will be necessary to restrict attention to value decreasing controls viz. controls such that $W(u,t)$ is a Supermartingale. We can now state the first Theorem of this section.

Theorem 1.4.1.

There exists a constant J^* and for each $u \in \mathcal{U}$, there exists a predictable process $\bar{A}_t(u) \in \mathcal{A}(\mathcal{Y}_t,\mathcal{P}^u)$ such that

1) $E^u\bar{A}_1(u) = J^*$ for all $u \in \mathcal{U}$.

2) $E^u\{-\bar{A}_t^{t+h}(u) + \int_t^{t+h} c(s,u(s))\, ds + \int_t^{t+h} dN^u(s)/\mathcal{Y}_t\} \geq 0$ a.s. $\mathcal{P}^u$.

for all $0 \leq t \leq t+h \leq 1$.

A control law u^* is optimal if and only if equality holds in 2) for all $0 \leq t \leq t+h \leq 1$. Then furthermore, $J(u^*) = J^*$ and

$$W(u^*,t) = E^{u^*}\{\bar{A}_t^1(u^*)/\mathcal{Y}_t\} \quad \text{a.s. } \mathcal{P}^{u^*}.$$

where $\bar{A}_t^1(u) = \bar{A}_1(u) - \bar{A}_t(u)$.

Proof: as in Theorem 2.2 of Boel [7]. ¤

We can now use the fact that the σ-fields are free of times of discontinuity to obtain the following sharper version of the above Theorem.

Theorem 1.4.2.

There exists a constant J^* and for each $u \in \mathcal{U}$ there exists a predictable process $\bar{A}_t(u) \in \mathcal{A}(\mathcal{Y}_t, \mathcal{P}^u)$ such that

1) $E^u\bar{A}_t(u) = J^*$ for all $u \in \mathcal{U}$.

2) $-\bar{A}_t^{t+h} + \int_t^{t+h} \hat{c}(s,u(s))\, ds + \int_t^{t+h} dN^u(s) \geq 0$ a.s. $\mathcal{P}^u$

for all $0 \leq t \leq t+h \leq 1$. A control law u^* is optimal if and only if equality holds in 2) for all $0 \leq t \leq t+h \leq 1$. Then furthermore $J^* = J(u^*)$ and $W(u^*,t) = E^{u^*}\{A_t^1(u^*)/\mathcal{Y}_t\}$.

Proof: For each $u \in \mathcal{U}$, let $\bar{A}_t(u)$ and $\bar{m}_t(u)$ be the processes introduced just before Theorem 1.4.1. Then we have $W(u,t) = J^* - \bar{A}_t(u) + \bar{m}_t(u)$.

Thus from the Optimality Principle we may conclude that $-\bar{A}_t(u) + \int_0^t \hat{c}(s,u(s))\, ds + \int_0^t dN^u(s)$ is a $(\mathcal{Y}_t, \mathcal{P}^u)$ Submartingale. Thus by the Doob decomposition Theorem (Meyer [13], VII, T31) it admits a decompoition

$$-\bar{A}_t(u) + \int_0^t \hat{c}(s,u(s))\, ds + \int_0^t dN^u(s) = A + B_t(u) + m_t(u)$$

where $B_t(u)$ is a predictable increasing process and $m_t(u)$ is a martingale and A is a constant. The expression on the left is predictable and of integrable variation. Thus $m_t(u)$ must be predictable and of integrable variation. But a predictable martingale of integrable variation on a family of σ-fields which are free of times of discontinuity vanishes. Thus $m_t(u) = 0$. Thus the process on the left is increasing. This shows that condition 2) is satisfied. On the other hand $E^u\bar{A}_1(u) = J^*$ by Construction. Let now u^* be optimal. Then the argument used above applies noting that $B_t(u^*) = 0$ since $-\bar{A}_t(u^*) + \int_0^t \hat{c}(s,u^*(s))\, ds + \int_0^t dN^{u^*}(s)$ is a Martingale by the optimality Principle. The sufficiency part follows from Theorem 1.4.1 since a process $\bar{A}_t(u)$ which satisfies 2) of Theorem 1.4.2 clearly satisfies 2) of Theorem 1.4.1. This completes the proof of the Theorem. ◻

Theorem 1.4.2 is stronger than Theorem 1.4.1 since it is a statement regarding the behaviour of the paths whereas Theorem 1.4.1 is a statement regarding their expected value.

In the next section we use the above theorem to obtain local conditions for optimality. This is accomplished by showing that $\bar{A}_t(u)$ is absolutely continuous in t. However we shall need that $\bar{A}_t(u)$ is an increasing process. Thus we will henceforth restrict attention to value decreasing controls.

1.5. Local Conditions for Optimality

We begin with the following lemma which yields a suitable representation for the process $A_t(u)$. Further we restrict attention to value decreasing controls. Thus for the statement of the next lemma $A_t(u)$ is the generator of the potential $W(u,t)$.

Lemma 1.5.1.

For all value decreasing controls u, there exist non-negative, $\mathcal{Y}_t$-adapted predictable processes

α_s^u, β_s^u satisfying: $A_t(u) = \int_0^t \alpha_s^u \, ds + \int_0^t \beta_s^u \, dN^u(s)$.

Proof: By the Optimality Principle we have

$$E^u[A_{t+h}(u) - A_t(u)/\mathcal{Y}_t] \leq E^u[\int_t^{t+h} c(s,u(s)) \, ds + \int_t^{t+h} dN^u(s)/\mathcal{Y}_t]$$

This implies that for any non-negative well measurable process $(\phi_s, \mathcal{Y}_t, P^u)$ we have

$$0 \leq E^u \int_0^t \phi_s dA_s(u) \leq E^u[\int_0^t \phi_s c(s,u(s)) \, ds + \int_0^t \phi_s dN^u(s)]$$

Thus whenever the second integral vanishes so does the first. Thus by the Radon-Nikodym Theorem, there exist non-negative $\mathcal{Y}_t$-predictable processes α_s^u, β_s^u satisfying

$$A_t(u) = \int_0^t \alpha_s^u ds + \int_0^t \beta_s^u \, dN^u(s).$$

This completes the proof of the lemma. ◻

Next we prove a lemma which relates the process β_s^u to the potential $W(u,t)$.

<u>Lemma 1.5.2</u>

Let A_t be an increasing integrable process and let W_t be the potential generated by it. Assume that there exists an increasing sequence $\{T_k\}$ of $\mathcal{Y}_t$-predictable stopping times and non-negative $\mathcal{Y}_t$-predictable processes α_s, β_s such that $A_t = \int_0^t \alpha_s ds + \sum_k \beta_{T_k} 1_{t \geq T_k}$. Then for each $k = 0,1,\ldots$ we have $\beta_{T_k} = W_{T_k-} - W_{T_k}$.

<u>Proof</u>: Fix k. Since T_k is a $\mathcal{Y}_t$-predictable stopping time, there exists an increasing sequence of stopping times T^n such that $T^n \uparrow T_k$. a.s. By considering $T^n \vee T_{k-1}$ if necessary we can assume that $T^n \geq T_{k-1}$ for all n. Now fix n. Fix a set $B^n \in \mathcal{Y}_{T^n}$. Since A_t generates W_t, we have $W_t = -A_t + m_t$ where m_t is a Martingale. Thus

$$E\int_0^1 1_{B^n} 1_{(T^n,T_k]} dW_t = -E\int_0^1 1_{B^n} \cdot 1_{(T^n,T_k]} dA_t + E\int_0^1 1_{B^n} 1_{(T^n,T_k]} dm_t$$

The second term on the right is zero since m_t is a martingale. Thus substituting for A_t, we obtain:

$$E\int_0^1 1_{B^n} 1_{(T^n,T_k]} dW_t = -E\int_0^1 1_{B^n} 1_{(T^n,T_k]} \alpha_s ds - E\, 1_{B^n} \cdot \beta_{T_k}$$

Since $B^n \in \mathcal{Y}_{T^n}$ is arbitrary it follows that

$$E[W_{T_k} - W_{T^n} / \mathcal{Y}_{T^n}] = -E[\int_{T^n}^{T_k} \alpha_s ds / \mathcal{Y}_{T^n}] - E[\beta_{T_k} / \mathcal{Y}_{T^n}]$$

or

$$W_{T^n} - E[W_{T_k} / \mathcal{Y}_{T^n}] = E[\int_{T^n}^{T_k} \alpha_s ds / \mathcal{Y}_{T^n}] + E[\beta_{T_k} / \mathcal{Y}_{T^n}]$$

Taking the limit as $n \to \infty$ on both sides of the above equality, we obtain:

$$W_{T^-_k} - E[W_{T_k} / \mathcal{Y}_{T^-_k}] = E[\beta_{T_k} / \mathcal{Y}_{T^-_k}].$$

Since the σ-fields are free of times of discontinuity and T_k is predictable we have $\mathcal{Y}_{T^-_k} = \mathcal{Y}_{T_k}$. Thus we obtain $W_{T^-_k} - W_{T_k} = \beta_{T_k}$ since W is $\mathcal{Y}_t$-adapted. This concludes the proof of the lemma. ⌑

We now come to the main Theorem of this chapter. It provides local necessary and sufficient conditions for optimality.

Theorem 1.5.1.

There exists a constant J^* and for all value decreasing controls $u \in \mathcal{U}$ processes α^u_s and β^u_s which are $\mathcal{Y}_t$-adapted and predictable satisfying

1) $E^u[\int_0^1 \alpha^u_s ds + \int_0^1 \beta^u_s dN^u(s)] = J^*$ for all u.

2) $\hat{c}(s,u(s)) - \alpha^u_s \geq 0$ a.s. $ds \times d\mathcal{P}^u$.

3) $\beta^u_s \leq 1$ a.s. $dN^u(s) \times d\mathcal{P}^u$.

A control law u^* is optimal if and only if 2) and 3) above are satisfied with equality for u^*. Then $J^* = J(u^*)$ the cost of the optimal control law and $W(u^*,t) = E^{u^*}[A^1_t(u^*) / \mathcal{Y}_t]$ where $A_t(u^*) = \int_0^t \alpha^{u^*}_s ds + \int_0^t \beta^{u^*}_s dN^{u^*}(s)$.
Furthermore we have: $\beta^{u^*}_{T^{u^*}_k} = W(u^*, T^{u^*}_k -) - W(u^*, T^{u^*}_k)$.

Proof: Fix a value decreasing control u. Then W(u,t) is a Potential of class D. Let then $A_t(u)$ be the unique increasing predictable process which generates W(u,t). Thus there exists a constant J^* and a martingale $m_t(u)$ s.t. $W(u,t) = J^* - A_t(u) + m_t(u)$. Since $W(u,1) = 0$ and $m_t(u)$ is a martingale, J^* clearly satisfies $E^u A_1(u) = J^*$ for all u. Let α^u_s and β^u_s

be the processes of Lemma 1.5.1. Then it is clear that α_s^u and β_s^u satisfy the condition 1) of the Theorem. From Theorem 1.4.2 we have

$$\int_t^{t+h} (\hat{c}_s^u - \alpha_s^u)\, ds + \int_t^{t+h} (1-\beta_s^u)\, dN^u(s) \geq 0 \quad \text{for all } h,t\ldots \tag{b}$$

Identifying the jumps in the 2nd integral we see that $\hat{c}_s^u - \alpha_s^u \geq 0$ a.s. $ds \times d\mathcal{P}^u$ and $\beta_s^u \leq 1$ a.s. $dN^u(s) \times d\mathcal{P}^u$. Thus α_s^u and β_s^u satisfy conditions 2) and 3) of the Theorem. Let now u^* be an optimal control. Then again from Theorem 1.4.2, (b) above holds with equality. Thus it follows that conditions 2) and 3) hold with equality for u^*. This proves the necessary part of the Theorem.

<u>Sufficiency</u>: Assume there exists u^* satisfying 2) and 3) with equality. Let

$$B_t(u) = \int_0^t \alpha_s^u ds + \int_0^t \beta_s^u dN^u(s)$$

$$V_t(u) = E^u\{B_1(u)/\mathcal{F}_t\} - B_t(u)$$

$$C_t(u) = \int_0^t \hat{c}_s^u\, ds + \int_0^t dN^u(s)$$

$$\psi_t(u) = E^u\{C_1(u)/\mathcal{F}_t\} - C_t(u)$$

Consider

$$\psi_t(u) - V_t(u) = E^u\Big[\int_0^1 \hat{c}_s^u ds + \int_0^1 dN^u(s)/\mathcal{F}_t\Big] - \Big[\int_0^t \hat{c}_s^u ds + \int_0^t dN^u(s)\Big]$$

$$- E^u\Big[\int_0^1 \alpha_s^u ds + \int_0^1 \beta_s^u dN^u(s)/\mathcal{F}_t\Big] + \Big[\int_0^t \alpha_s^u ds + \int_0^t \beta_s^u dN^u(s)\Big]$$

$$= E^u\left[\int_0^1 (\hat{c}^u_s - \alpha^u_s)ds + \int_0^1 (1-\beta^u_s)dN^u(s)/\mathcal{Y}_t\right]$$

$$- \left[\int_0^t (\hat{c}^u_s - \alpha^u_s)ds + \int_0^t (1-\beta^u_s)dN^u(s)\right]$$

$$= E^u\left[\int_t^1 (\hat{c}^u_s - {}^u_s)ds + \int_t^1 (1-\beta^u_s)dN^u(s)/\mathcal{Y}_t\right] \geq 0$$

by conditions 2) and 3). It follows that

$$\psi_t(u) \geq V_t(u) \quad \text{a.s. } ds \times d\mathcal{P}^u \text{ and } \quad \psi_t(u^*) = V_t(u^*) \quad \text{a.s. } ds \times d\mathcal{P}^{u^*}$$

Thus $E^u \ \psi_o(u) \geq E^u V_o(u)$ and

$$E^{u^*} \ \psi_o(u^*) = E^{u^*} V_o(u^*)$$

But $E^u \psi_o(u) = J(u)$, $E^{u^*} \ \psi_o(u^*) = J(u^*)$ and $E^u V_o(u) = E^{u^*} V_o(u^*) = J^*$ by Condition 1). Thus we have $J(u) \geq J(u^*) = J^*$. This shows that u^* is optimal and has cost J^*. To show that $W_t(u^*) = E^{u^*} \{A^1_t(u^*)/\mathcal{Y}_t\}$ we apply the sufficiency part of Theorem 1.4.1. Finally the stated relation between $\beta^{u^*}_{T^u_k{}^*}$ and $W(u^*, T^{u^*}_k)$ follows from Lemma 1.5.2. This concludes the proof of the Theorem. ◻

We note that the above theorem is the version which holds in our case of the optimality criterion derived for Markov processes by Kushner [11], for Conditional Markov Processes by Stratonovich [20], for processes on a Wiener Space by Davis-Varaiya [10]. Finally it is the abstract version of Theroem 2 of Bensoussan and Lions [4].

1.6. Optimality Conditions with Complete Information.

We now consider the case where the Information σ-fields are $\mathcal{F}_t$ rather than $\mathcal{Y}_t$. In order to make the problem meaningful we impose the following restriction on the probability measures $\mathcal{P}^u$.

Assumption: for all $t \in [0,1]$, for all control laws u, v, w

$$E^{utv}\left[\int_t^1 c(s,v_s)\,ds + \int_t^1 dN^v(s)/\mathcal{F}_t\right] = E^{wtv}\left[\int_t^1 c(s,v_s)ds + \int_t^1 dN^v(s)/\mathcal{F}_t\right]$$

This assumption implies that the dynamics defined by the probability measures $\mathcal{P}^u$ in fact define a dynamical system. For a version of this assumption which is stated for the probability measures instead of for the expectation operators defined by them see Boel-Varaiya [8]. Due to this assumption the value function now no longer depends on u, the control law used up to time t, since

$$W(u,t) = \inf_v J(t,utv) = \inf_v J(t,(wtv)) \text{ and}$$

therefore $W(u,t) = W(w,t)$ or W is independent of u. However the processes in the decomposition of W can still depend on u since the decomposition holds a.s. with respect to $\mathcal{P}^u$ measure. Some simplification does occur as can be seen by the following:

$$W(t) = J^* - A_t(u) + m_t(u) = J^* - A_t(v) + m_t(v) \text{ or}$$

$$J^* - A_t(u) + m_t^c(u) + m_t^d(u) = J^* - A_t(v) + m_t^c(v) + m_t^d(v) \qquad \text{(a0}$$

where $m_t^c(u)$ and $m_t^d(u)$ are the purely continuous and purely discontinuous parts of the martingale $m_t(u)$. Similarly for $m_t^c(v)$ and $m_t^d(v)$. Now we can identify the continuous and discontinuous parts in equality (a) to obtain

$$-A_t(u) + m_t^c(u) = -A_t(v) + m_t^c(v); \; m_t^d(u) = m_t^d(v).$$

Unfortunately this is as far as we can go. In the presence of a Martingale representation Theorem we can do better since each martingale above can be suitably represented as a Stochastic Integral with respect to the "basis" Martingales. A Martingale representation Theorem for martingales on σ-fields generated by fundamental jump processes has been proved by Boel-Varaiya-Wong in [9]. Boel [7] has shown how this can be used to simplify the optimality criteria for jump process. Davis-Varaiya [10] have shown how the Martingale representation Theorem for Martingales on a Wiener Space can be used to simplify the optimality criteria for the control of systems driven by a Wiener Process. We shall encounter the case of a system driven by a Wiener Process in the next chapter, where no such simplification occurs. This is because for the system considered, it is not possible to model the effect of control laws as changing the Probability Measures on a fixed Sample Space of paths.

We conclude that for the model considered above the simplification which results in the complete information case is that all the results of the previous sections hold with W(u,t) replaced by W(t), $\mathcal{Y}_t$ by $\mathcal{F}_t$ and $\hat{c}$ by c. We can replace $\hat{c}$ by c since c is $\mathcal{F}_t$-adapted.

CHAPTER 2

AN APPLICATION TO THE PROBLEM OF CONTINUOUS TIME INVENTORY CONTROL

2.1 Model.

Fix a time interval [0,1].

Let B_s denote a zero mean Brownian Motion Process on [0,1] with values in $\mathbb{R}^p$ and $B_0 = 0$. We assume that the components of B_s are independent. Let $\mu: [0,1] \to \mathbb{R}^p$ be a bounded integrable function and σ a $p \times p$ diagonal matrix. The accumulated demand between times t and s is

$$D(t,s) = \int_t^s \mu(z)\, dz + \sigma(B_s - B_t)$$

where $0 \leq t \leq s \leq 1$.

Control Laws:

Fix a set $U \subseteq \mathbb{R}_+^p$. Let $\mathcal{F}_t = \sigma\{B_s;\ s \leq t\}$. An admissible control law is an $\mathcal{F}_t$-predictable map $u: [0,1] \times \Omega \to \mathbb{R}^p$ satisfying the following conditions

1) u has piecewise constant sample paths which are right continuous with left limits.

2) Each sample path of u has a finite number of jumps in a finite time interval.

3) The jump of u at time $t = \Delta u = u_t - u_{t-} \in U$ for all t and for all paths of u.

Let $\mathcal{U}$ denote the set of admissible control laws. We can inductively define the jump times of u by:

$T_0^k(\omega) = 0$, $T_{k+1}^u(\omega) = \text{Inf}\{t \mid t \geq T_k^u(\omega),\ u_t(\omega) \neq u_{T_k^u(\omega)}(\omega)\}$ where the infinuum over the empty set is understood to be $+\infty$. The same comment as made in Chapter 1 applies here. It is clear that each jump time of u is a stopping time of the family $\{\mathcal{F}_t'\}$. Similarly we can define the jump heights of u by $h_k^u(\omega) = \Delta u_{T_k^u(\omega)}(\omega)$.

The jump times of u represent times at which orders for inventories are places, the jump heights are the quantities ordered. Both are random variables since we are interested in feedback control laws. It is clear that the pair of sequences $\{T_k^u, h_k^u\}_k$ is an equivalent description of the control law. The two descriptions are related by

$$u_t = \sum_k 1_{t \geq T_k^u} \cdot h_k^u.$$

Thus if u is an admissible control law, u_t represents the total supply ordered up to time t.

We now define the <u>trajectory generated by a control law u</u> to be

$$x_t^u = \int_0^t \mu(s)\, ds + \sigma B_t - u_t$$

$$= \int_0^t \mu(s)\, ds + \sigma B_t - \sum_k 1_{t \geq T_k^u} h_k^u.$$

If $x_t^u > 0$ we have that the accumulated demand between 0 and t is bigger than the total supply ordered up to time t. A similar interpretation holds for the reverse inequaltiy and for equality.

Cost Associated with a Control Law

Let c be a bounded continuous map $c: [0,1] \times \mathbb{R}^p \to \mathbb{R}_+$. The cost associated with the control law u is then

$$J(u) = E\{\int_0^1 c(s,x_s^u)\, ds + \int_0^1 d\, N^u(s)\}$$

where $N^u(t) = \sum_k 1_{t \geq T_k^u}$ is the counting process which counts the number of jumps of u up to time t. Thus each placement of an order costs one unit. The role of a fixed cost of ordering is essential. In effect it rules out continuous control laws which would lead to infinite costs.

We are interested in determining necessary and sufficient conditions for a control law to be optimal in the sense that

$J(u^*) \leq J(u)$ for all admissible control laws u.

For a general survey of inventory control models see Scarf [18]. For a discrete time approach see Scarf [19], Vienott [21]. For another approach in continuous time with infinite horizon see Bather [1]. The model we have adopted is similar to the one used by Bensoussan and Lions [4].

We shall solve this problem by the methods of Dynamic Programming. We thus need a definition of concatenation of control laws.

Concatenation of control laws:

Let u and v be admissible control laws and let $0 \leq t \leq 1$. The concatenation of u and v at time t, denoted by utv is the control law defined by:

$$(utv)_s = \begin{cases} u_s & s \leq t \\ u_t + (v_s - v_t) & s > t. \end{cases}$$

It is easily verified that utv is an admissible control law.

2.2 Value Function and conditions for optimality.

$$\text{Let } J(t,(utv)) = E\{\int_t^1 c(s,x_s^{utv})\, ds + \int_t^1 dN^v(s) | \mathcal{F}_t\}.$$

Now exactly as in the abstract case we can define:

$$W(u,t) = \inf_v J(t,(utv))$$

We note that the σ-fields $\mathcal{F}_t$ form an increasing family in t and that they are not affected by changes in control laws. Therefore the relative completeness lemma of Davis-Varaiya [10] holds. Exactly as in the abstract case we can establish the Principle of Optimality.

Theorem 2.2.1.

For all $u \in \mathcal{U}$, for all t, t+h such that $0 \le t \le t+h \le 1$, we have:

$$W(u,t) \le E\{\int_t^{t+h} c(s,x_s^u)\, ds + \int_t^{t+h} dN^u(s) | \mathcal{F}_t\} + E\{W(u,t+h | \mathcal{F}_t\}$$

with equality if and only if u is optimal. The Theorem holds for any $\mathcal{F}_t$ stopping times T and S such that $T \le S$.

Proof: As in Theorem 1.3.1. □

The reasoning from now on is the same as in the abstract case. We thus have the following analog of Theorem 1.4.1.

Theorem 2.2.2. There exists a constant J* and for all controls u a process $\bar{A}_t(u) \in \mathcal{A}(\mathcal{F}_t, \mathcal{P})$ satisfying:

1) $E\, \bar{A}_t(u) = J^*$ for all u.

2) For all $0 \leq t \leq t+h \leq 1$

$$E\{-A_t^{t+h}(u) + \int_t^{t+h} c(s,x_s^u)\, ds + \int_t^{t+h} dN^u(s) \,|\, \mathcal{F}_t\} \geq 0 \quad \text{a.s.}$$

A control law u* is optimal if and only if 2) hold with equality for u*. Then J* = J(u*) the cost of the optimal control and

$$W(u^*,t) = E\{A_t^1(u^*) \,|\, \mathcal{F}_t\}.$$

Proof: As in Theorem 1.4.1. ◻

We now arrive at the main Theorem of this section. It provides a characterization of the value process W(u,t). It is similar to Theorem 4.3 of Davis-Varaiya [10] and to Theorem 1.5.2. It is better than Theorem 1.5.2 because of the presence of a martingale representation theorem.

Theorem 2.2.3.

There exists a constant J* and for all value decreasing controls u, processes α_s^u, β_s^u, γ_s^u satisfying the following conditions.

1(a) $E\int_0^1 \gamma_s^u \, dB_s = 0$ (b) $\int_0^1 |\gamma_s^u|^2 \, ds < \infty$ a.s.

2) $c(s,x_s^u) - \alpha_s^u \geq 0$ a.s.

3) $\beta_s^u \leq 1$ a.s. $dN^u(s)$

4) $V_1(u) = 0$ for all u

where $V_t(u) = J^* - \int_0^t \alpha_s^u\, ds - \int_0^t \beta_s^u\, dN^u(s) + \int_0^t \gamma_s^u dB_s$. A control law u* is optimal if and only if 2) and 3) hold with equality for u*. Then $J^* = J(u^*)$ the cost of the optimal control and $V_t(u^*) = W(t,u^*)$ the value process evaluated along the optimal control law.

Proof: Let u* be an optimal control law. Fix a value decreasing control u. Let $A_t(u)$ be the generator of the potential $W(u,t)$. Then we have $W(u,t) = J^* - A_t(u) + m_t(u)$, where $m_t(u)$ is a square integrable $\mathcal{F}_t$-martingale. Thus by the martingale representation theorem [10,22], there exists a process γ_s^u satisfying 1(a) and 1(b) such that $m_t(u) = \int_0^t \gamma_s^u\, dB_s$. Let α_s^u and β_s^u be the process of Theorem 1.5.1. Then α_s^u and β_s^u satisfy 2) and 3) by Theorem 1.5.1. Clearly $W_1(u) = 0$ for all u and

$$W_t(u) = J^* - \int_0^t \alpha_s^u\, ds - \int_0^t \beta_s^u\, dN^u(s) + \int_0^t \gamma_s^u\, dB_s \text{ by definition.}$$

Conversely, assume there exist processes α_s^u, β_s^u, γ_s^u satisfying the condition of the theorem. The same argument which was used in the sufficiency part of theorem 1.5.1 yields the result. This completes the proof of the theorem. ◻

2.3 Markov Controls, Markovian Value Function V.

The set $\mathcal{M}$ of admissible Markov controls is defined by

$$\mathcal{M} = \{u \in \mathcal{U} \mid \text{for all } 0 \le t-h \le t \le 1,\ u_t - u_{t-h} \text{ is}$$
$$\sigma\{B_s;\ t-h \le s \le t\}\text{-Measurable}\ .\}$$

For each $t \in [0,1]$, we obtain $\mathcal{M}_t$ from $\mathcal{M}$ by restricting the domain [0,1] of control laws in $\mathcal{M}$ to [t,1]. If $u \in \mathcal{M}_t$ let

$$J(t,x,u) = E\{\int_t^1 c(s,x_s^u)\,ds + \int_t^1 dN^u(s)\,|\,x_t = x\}.$$

The Markovian value function $V: [0,1] \times \mathbb{R}^p \to \mathbb{R}_+$ is defined by:

$$V(t,x) = \inf_{u\in\mathcal{M}_t} E\{\int_t^1 c(s,x_s^u)\,ds + \int_t^1 dN^u(s)/x_t = x\}.$$

We wish to show that V satisfies a principle of optimality. In order to do this we use discrete backward dynamic programming. This method is quite similar to the one used by Boel [7].

We fix $t \in [0,1]$ and an integer $N > 0$. For each $n = 0,1,\ldots\, 2^N$ we let $t_n = t + (1-t)\frac{n}{2^N}$. Thus $t_0 = t$ and $t_{2^N} = 1$. This partitions the interval $[t,1]$ into 2^N equal divisions. Let the graph of t_n, $[t_n]$, be defined by

$[t_n] = \{(\omega,t_n)/\omega\in\Omega\}$. Similarly if $u \in \mathcal{M}_t$, define

$$[T_k^u] = \{(\omega,T_k^u(\omega))/\omega\in\Omega\}.$$

Now let $\mathcal{M}_t^N$ be the set of all $u \in \mathcal{M}_t$ such that for all k, $[T_k^u] \subseteq \bigcup_{n=0}^{2^N} [t_n]$. Then by definition of $\mathcal{M}_t$, it is clear that $u \in \mathcal{M}_t^N$ satisfies

$$u_s(\omega) - u_{s-}(\omega) = 0 \text{ for } t_n < s < t_{n+1},\ n=0,\ldots\, 2^N - 1$$

$$u_{t_n}(\omega) - u_{t_n-}(\omega) = h(t_n, x_{t_n-}^u(\omega)) \text{ for some function h.}$$

Thus the decision times for $u \in \mathcal{M}_t^N$ are restricted to be at times t_n and the value of the jump at time t_n depends only on the state at that time. Let $\mathcal{U}_t$ be obtained from $\mathcal{U}$ in the same way that $\mathcal{M}_t$ was obtained from $\mathcal{M}$.

Let $\mathcal{U}_t^N = \{u \in \mathcal{U}_t \mid \text{for all } k\ [T_k^u] \subseteq \bigcup_{n=0}^{2^N} [t_n]\}$. We define $V^N(t,x)$ by backward induction as follows:

Let $f(t_{2^N},x) = 0$ for all $x \in \mathbb{R}^P$.

Then define $f(t_n,x)$ for $x \in \mathbb{R}^P$ as follows:

$$f(t_n,x) = \inf_{h(t_n,x)} E\{\int_{t_n}^{t_{n+1}} c(s,x-h(t_n,x) + D(t_n,s))\, ds + N^h(t_n,x)$$

$$+ f(t_{n+1},x-h(t_n,x) + D(t_n,t_{n+1}))\}$$

where $N^h(t_n,x) = 1$ if $h(t_n,x) > 0$

0 otherwise.

And we define $V^N(t,x) = f(t_o,x)$, $x \in \mathbb{R}^P$. We are now ready to state our first lemma.

<u>Lemma 2.3.1.</u>

Fix $N > 0$ and $\varepsilon > 0$ and $x \in \mathbb{R}^P$. Then there exists $u_N^\varepsilon \in \mathcal{M}_t^N$ such that $J(t,x,u_N^\varepsilon) \leq V^N(t,x) + \varepsilon$.

<u>Proof:</u> By definition of $f(t_n,x)$, there exists a function $h^\varepsilon(t_n,\cdot)$ such that for all $x \in \mathbb{R}^P$,

$$f(t_n,x) + \frac{\varepsilon}{2_N} \geq E\{\int_{t_n}^{t_{n+1}} c(s,x-h^\varepsilon(t_n,x) + D(t_n,s))\, ds + N^h(t_n,s)$$

$$+ f(t_{n+1},x-h^\varepsilon(t_n,s) + D(t_n,t_{n+1})\}.$$

Define $u_N^\varepsilon \in \mathcal{M}_t^N$ by $u_N^\varepsilon(t_n,\omega) - u_N^\varepsilon(t_n^-,\omega) = h^\varepsilon(t_n, x_{t_n^-}^{u_N^\varepsilon}(\omega))$. Then we have

$$J(t,x,u_N^\varepsilon) = \sum_0^{2^N-1} E\{\int_{t_n}^{t_{n+1}} c(s,x_s^{u_N^\varepsilon})\, ds + \int_{t_n}^{t_{n+1}} dN^{u_N^\varepsilon}(s)\}$$

$$\leq f(t_o,x) + \sum_0^{2^N-1} \frac{\varepsilon}{2^N} + \sum_0^{2^N-1} E[-f(t_n,x_{t_n}^{u_N^\varepsilon}) + f(t_n,x_{t_n}^{u_N^\varepsilon})] + f(t_{2^N},x_{t_{2^N}}^{u_N^\varepsilon})$$

$$= f(t_o,x) + \varepsilon \text{ (since } f(t_{2^N},z) = 0 \text{ for all } z.$$

Since $f(t_o,x) = V^N(t,x)$ we have

$$J(t,x,u_N^\varepsilon) \leq V^N(t,x) + \varepsilon.$$ ¤

Lemma 2.3.2.

Fix $x \in \mathbb{R}^P$ and $\varepsilon > 0$. Then for all $u \in \mathcal{M}_t$, there exists N and $u_N^\varepsilon \in \mathcal{M}_t^N$ such that

$$J(t,x,u_N) \leq J(t,x,u) + \varepsilon.$$

Proof: The proof is in two steps. We first show that $u \in \mathcal{M}_t$ can be approximated by a $v_N \in \mathcal{U}_t^N$ for N sufficiently large and then show that we can find $u_N^\varepsilon \in \mathcal{M}_t^N$ which approximates $v_N \in \mathcal{U}_t^N$.

Step 1: Fix $u \in \mathcal{M}_t$. For each positive integer N define $v_N \in \mathcal{U}_t^N$ as follows: Let $\{T_k^u\}$ be the jump times of u. Define the jump times $T_k^{v_N}$ of v_N by $T_k^{v_N}(\omega) = t_j$ where j is the smallest integer between 0 and 2^N satisfying $t_j \geq T_k^u(\omega)$. Define $h_k^{v_N}(\omega) = h_k^u(\omega)$. Then by definition of $T_k^{v_N}$

we have $T_k^{v_N} \geq T_k^u$. Thus for each k, $T_k^{v_N}$ is a stopping time of $\{\mathcal{F}\}_t$ and v_N is an admissible control. Since $u \in \mathcal{M}_t \subseteq \mathcal{U}$ we have $v_N \in \mathcal{U}_t^N$. It is now clear from the definition of v_N that

$T_k^{v_N} \downarrow T_k^u$ and $h_k^{v_N} = h_k^u$. Thus

$$\int_t^1 dN^{v_N}(s) \xrightarrow[N\to\infty]{} \int_t^1 dN^u(s)$$

and $x_s^{v_N} \xrightarrow[N]{} x_s^u$ a.s. for all $s \in [t,1]$. Thus

$$J(t,x,v_N) = E\{\int_t^1 c(s,x_s^{v_N})\, ds + \int_t^1 dN^{v_N}(s)/x_t = x\}$$

$$\xrightarrow[N]{} \{E \int_t^1 c(s,x_s^u)\, ds + \int_t^1 dN^v(s)/x_t = x\} = J(t,x,u)$$

where we have used the boundedness of c to justify the interchange of limit operations. This shows that given $\varepsilon > 0$, there exists N and $v_N \in \mathcal{U}_t^N$ such that

$$J(t,x,v_N) \leq J(t,x,u) + \varepsilon.$$

Step 2: Fix N. Now we note that

$$\operatorname*{Inf}_h E\{\int_{t_{2^N-1}}^1 c(s,x_{t_{2^N-1}} - h + D(t_{2^N-1},s)\, ds + N^h/\mathcal{F}_{t_{2^N-1}}\}$$

where the infinuum is taken over all h which are $\mathcal{F}_{t_{2^N-1}}$-Measurable, is equal to

$$\operatorname*{Inf}_{h} E\{\int_{t_{2^N-1}}^{1} c(s,x_{t_{2^N-1}} - h + D(t_{2^N-1},s)\, ds + N^h/x_{t_{2^N-1}}\}$$

This is because D(t,s) is a Markov Process. We thus obtain that for all $\varepsilon > 0$, there exists a function $h^{\varepsilon}(t_{2^N-1},x)$ such that for all $x \in \mathbb{R}^P$, for all $\mathcal{F}_{t_{2^N-1}}$-Measurable functions h

$$E\{\int_{t}^{1} c(s,x-h^{\varepsilon}(t_{2^N-1},x) + D(t,s)\, ds + N^h\}$$

$$\leq E\{\int_{t_{2^N-1}}^{1} c(s,x-h + D(t,s))\, ds + N^h/\mathcal{F}_{t_{2^N-1}}\} + \frac{\varepsilon}{2^N}$$

$$= E\{\int_{t_{2^N-1}}^{1} c(s,x-h + D(t,s))\, ds + N^h\} + \frac{\varepsilon}{2^N}$$

We can carry this process down to n = 0 to obtain

$$E\{\int_{t}^{1} c(s,x_s^{u_N^{\varepsilon}})\, ds + \int_{t}^{1} dN^{u_N^{\varepsilon}}(s)/x_t\} \leq E\{\int_{t}^{1} c(s,x_s^{u})\, ds + \int_{t}^{1} dN^{u}(s)/\mathcal{F}_t + \varepsilon$$

$$= E\{\int_{t}^{1} c(s,x_s^{u})\, ds + \int_{t}^{1} dN^{u}(s)/x^{t}\} + \varepsilon \quad \text{for all } u \in \mathcal{U}_t.$$

Thus combining steps 1 and 2 we obtain the statement of the lemma. This concludes the proof of the lemma. ◻

Now by lemma 2.3.1 we have

$$V^N(t,x) = \operatorname*{Inf}_{u \in \mathcal{M}_t^N} J(t,x,u)$$

Since by definition, N > k implies $\mathcal{M}_t^N \supseteq \mathcal{M}_t^K$, we have that $V^N(t,x)$ is a decreasing sequence in N. On the other hand it is bounded from below by 0. Thus $\lim_{N\to\infty} V^N(t,x)$ exists and equals $\operatorname{Inf}_N V^N(t,x)$. Let

$U(t,x) = \lim_{N\to\infty} V^N(t,x) = \operatorname{Inf}_N V^N(t,x)$. In the next lemma, we show that

$U(t,x) = V(t,x)$ the Markovian value function.

<u>Lemma 2.3.3.</u>

$$\operatorname{Inf}_N V^N(t,x) = \lim_{N\to\infty} V^N(t,x) = U(t,x) = V(t,x).$$

<u>Proof:</u>

<u>Step 1</u>: We show that $U(t,x) \leq J(t,x,u)$ for all $u \in \mathcal{M}_t$. Towards this end fix $\varepsilon > 0$ and $u \in \mathcal{M}_t$. Then since $U(t,x) = \operatorname{Inf}_N V^N(t,x)$ we have $U(t,x) \leq V^N(t,x)$ for all N. On the other hand since $V^N(t,x) = \operatorname{Inf}_{u\in\mathcal{M}_t^N} J(t,x,u)$ we have $V^N(t,x) \leq J(t,x,u)$ for all $u \in \mathcal{M}_t^N$. Now choose N sufficiently large that

$$J(t,x,u_N^\varepsilon) \leq J(t,x,u) + \varepsilon.$$

We thus obtain:

$$U(t,x) \leq V^N(t,x) \leq J(t,x,u_N^\varepsilon) \leq J(t,x,u) + \varepsilon.$$

Thus $U(t,x) \leq J(t,x,u) + \varepsilon$. Since $\varepsilon > 0$ is arbitrary and u is arbitrary it follows that $U(t,x) \leq J(t,x,u)$ for all $u \in \mathcal{M}_t$.

<u>Step 2.</u> We show that for all $\varepsilon > 0$, there exists $u \in \mathcal{M}_t$ such that

$$J(t,x,u) \leq U(t,x) + \varepsilon.$$

Since $U(t,x) = \operatorname{Inf}_N V^N(t,x)$ we have

(a) $U(t,x) + 2\varepsilon \geq V^N(t,x) + \varepsilon$ for N sufficiently large. Since $V^N(t,x) = \operatorname{Inf}_{u \in \mathcal{M}_t^N} J(t,x,u)$ we have, there exists $u_N^\varepsilon \in \mathcal{M}_t^N$ such that

(b) $V^N(t,x) + \varepsilon \geq J(t,x,u_N^\varepsilon)$. Thus combining (a) and (b) we have

$U(t,x) + 2\varepsilon \geq J(t,x,u_N^\varepsilon)$. But $u_N^\varepsilon \in \mathcal{M}_t^N \subseteq \mathcal{M}_t$. Thus we have established the lemma. ¤

We now impose a Lipschitz condition on c and show that each V^N is then uniformly Lipschitz in x for fixed t. This will enable us to conclude that $V^N(t,x)$ converges uniformly in x to $V(t,x)$ and that for each fixed t, $V(t,x)$ is uniformly Lipschitz in x.

Assumption: We assume that the cost function c satisfies a uniform Lipschitz condition in x. i.e. there exists a constant K such that for all $s \in [0,1]$, for all x' and $x \in \mathbb{R}^P$, $|c(s,x) - c(s,x')| \leq K\,|x-x'|$.

Lemma 2.3.4. For each fixed $t \in [0,1]$, $V^N(t,x)$ converges uniformly to $V(t,x)$ in x as $N \to \infty$

Proof: We show that $V^N(t,x)$ is uniformly Lipschitz in x with Lipschitz constant independent of N. Towards this end we fix N and let

$t_n = t + (1-t)\frac{n}{2^N}$, $n=0,\dots 2^N$ as before. We first show that each function $f(t_n,x)$ is uniformly Lipschitz in x. Fix x and $x' \in \mathbb{R}^P$. Fix $\varepsilon > 0$. Then by definition of $f(t_{2^N-1},x)$ we have there exists $u \in \mathcal{M}_{t_{2^N-1}}^N$ such that

$$f(t_{2^N-1},x) \le E\{\int_{t_{2^N-1}}^{1} c(s,x-h(t_{2^N-1},x) + D(t_{2^N-1},s)\; ds + N^h\} + \varepsilon$$

where $h(t_{2^N-1},x)$ is the jump of u at (t_{2^N-1},x) and $N^h = \begin{cases} 1 & \text{if } h(t_{2^N-1},x) > 0 \\ 0 & \text{Otherwise.} \end{cases}$

Using the same h at x' we have

$$f(t_{2^N-1},x') \le E\{\int_{t_{2^N-1}}^{1} c(s,x'-h(t_{2^N-1},x) + D(t_{2^N-1},s))\; ds + N^h\}$$

Thus we obtain

$$f(t_{2^N-1},x') - f(t_{2^N-1},x) \le E\left[\int_{t_{2^N-1}}^{1} |c(s,x-h(t_{2^N-1},x) + D(t_{2^N-1},s)) - c(s,x'-h(t_{2^N-1},x) + D(t_{2^N-1},s))|ds\right] + \varepsilon$$

since the N^h's cancel each other out. Using the uniform Lipschitz condition on c, we see that

$$f(t_{2^N-1},x') - f(t_{2^N-1},x) \le \frac{K(1-t)}{2^N} |x-x'| + \varepsilon.$$

Similarly by considering an ε-effective $h(t_n,x')$ at x' we obtain

$$f(t_{2^N-1},x) - f(t_{2^N-1},x') \le \frac{K(1-t)}{2^N} |x-x'| + \varepsilon.$$

Combining the last two inequalities, we obtain:

$$|f(t_{2^N-1},x) - f(t_{2^N-1},x')| \leq \frac{K(1-t)}{2^N} |x-x'| + \varepsilon$$

Since $\varepsilon > 0$ is arbitrary we obtain

$$|f(t_{2^N-1},x) - f(t_{2^N-1},x')| \leq \frac{K(1-t)}{2^N} |x-x'|$$

Using this argument backward, we see that

$$|f(t_n,x) - f(t_n,x')| \leq \frac{K(1-t)(2^N-n)}{2^N} |x-x'|, \; n=0\ldots 2^N-1.$$

Thus for n=0 we see that

$$|f(t_o,x) - f(t_o,x')| \leq K(1-t) \; |x-x'|.$$

Since $f(t_o,x) = V^N(t,x)$ we have shown that $V^N(t,x)$ is uniformly Lipschitz in x with coefficient independent of N. It follows that $V^N(t,x)$ decreases uniformly to $V(t,x)$ in x and that $V(t,x)$ satisfies a uniform Lipschitz condition in x. This completes the proof of the lemma. ⌑

We are now ready to prove the Markovian Principle of Optimality.

<u>Theorem 2.3.1.</u>

For all $(t,x) \in [0,1] \times \mathbb{R}^P$, for all h s.t. $t + h \leq 1$, for all $u \in \mathcal{M}_t$, the Markovian value function $V(t,x)$ satisfies:

$$V(t,x) \leq E\left[\int_t^{t+h} c(s,x_s^u)\, ds + \int_t^{t+h} dN^u(s)/x_t = x\right] + E\left[V(t+h,x_{t+h}^u)/x_t = x\right]$$

with equality if and only if $u \in \mathcal{U}_t$ is optimal.

Proof: Step 1

By the definition of V(t,x) we have,

$$V(t,x) \leq E\left[\int_t^{t+h} c(s,x_s^u)ds + \int_t^{t+h} dN^u(s)/x_t = x\right] + \inf_{v \in \mathcal{U}_{t+h}} E\left[E\left\{\int_{t+h}^1 c(s,x_s^v)ds + \int_{t+h}^1 dN^v(s)/x_{t+h}^u\right\}/x_t = x\right]$$

Consider the second term on the right in the above expression. The first statement of the Theorem will be proved if we show that the Inf and the first conditional expectation operator can be interchanged. Towards this end we fix $\varepsilon > 0$. Then since $V^N(t+h,x) \downarrow V(t+h,x)$ uniformly in x we have: There exists N s.t.

$$V(t+h,x_{t+h}^u) + 2\varepsilon \geq V^N(t,x_{t+h}^u) + \varepsilon \quad \text{a.s.}$$

Now by Lemma 2.3.1, there exists $v_N^\varepsilon \in \mathcal{U}_{t+h}^N$ such that

$$V^N(t,x_{t+h}^u) + \varepsilon \geq J(t+h,x_{t+h}^u, v_N^\varepsilon) \quad \text{a.s.}$$

Combining this with the above inequality and taking expectations we obtain:

$$E[V(t+h,x_{t+h}^u)/x_t = x] + 2\varepsilon \geq E[J(t+h,x_{t+h}^u,v_N^\varepsilon)/x_t = x]$$

Taking the Inf over $v \in \mathcal{U}_{t+h}$ on the right, we get:

$$E[V(t+h,x_{t+h}^u)/x_t = x] + 2\varepsilon \geq \inf_{v \in \mathcal{U}_{t+h}} E[J(t+h,x_{t+h}^u,v)/x_t = x]$$

Since $\varepsilon > 0$ is arbitrary it follows that

$$E[V(t+h,x_{t+h}^u)/x_t = x] \geq \inf_{v \in \mathcal{U}_{t+h}} E[J(t+h,x_{t+h}^u,v)/x_t = x]$$

The reverse inequality is clear from the definition of $V(t+h,x^u_{t+h})$.

We have thus proved that:

$$\operatorname*{Inf}_{v\in\mathcal{M}_{t+h}} E\left[E\left\{\int_{t+h}^{1} c(s,x^v_s)ds + \int_{t+h}^{1} dN^v(s)/x^u_{t+h} \right\} /x_t = x\right] = E\left[V(t+h,x^u_{t+h})/x_t = x\right]$$

This establishes the first statement of the Theorem.

<u>Step 2:</u> We now show that $u \in \mathcal{M}_t$ is optimal in $\mathcal{M}_t$ if and only if

$$V(t+h,x^u_{t+h}) = E\left[\int_{t+h}^{1} c(s,x^u_s)\, ds + \int_{t+h}^{1} dN^u(s)/x^u_{t+h}\right] \text{ a.s.}$$

for all $t \leq t+h \leq 1$.

If equality holds above, we immediately obtain that u is optimal by taking $h = 0$. On the other hand suppose u is optimal. Then by definition of $V(t,x)$ we have:

$$V(t,x) = E\left[\int_{t}^{t+h} c(s,x^u_s)ds + \int_{t}^{t+h} dN^u(s)/x_t = x\right] + E\left[\int_{t+h}^{1} c(s,x^u_s)ds + \int_{t+h}^{1} dN^u(s)/x_t = x\right]$$

$$= E\left[\int_{t}^{t+h} c(s,x^u_s)ds + \int_{t}^{t+h} dN^u(s)/x_t = x\right] + E\left[E\left\{\int_{t+h}^{1} c(s,x^u_s)ds + \int_{t+h}^{1} dN^u(s)/x^u_{t+h}\right\} /x_t = x\right]$$

on the other hand by the first statement of the Theorem we have

$$V(t,x) \leq E\left[\int_t^{t+h} c(s,x_s^u)ds + \int_t^{t+h} dN^u(s)/x_t = x\right] + E\left[V(t+h,x_{t+h}^u)/x_t = x\right].$$

Subtracting gives:

$$E\left[E\left\{\int_{t+h}^1 c(s,x_s^u)ds + \int_{t+h}^1 dN^u(s)/x_{t+h}^u\right\} - V(t+h,x_{t+h}^u)/x_t = x\right] \leq 0$$

But by definition

$$E\left\{\int_{t+h}^1 c(s,x_s^u)ds + \int_{t+h}^1 dN^u(s)/x_{t+h}^u\right\} - V(t+h,x_{t+h}^u)$$

is a non-negative random variable. We thus obtain

$$V(t+h,x_{t+h}^u) = E\left\{\int_{t+h}^1 c(s,x_s^u)ds + \int_{t+h}^1 dN^u(s)/x_{t+h}^u\right\} \text{ a.s.}$$

Step 3. We now use step 2 to show that u is optimal iff equality holds in the first statement of the Theorem. If u is optimal then,

$$V(t,x) = E\left[\int_t^{t+h} c(s,x_s^u)ds + \int_t^{t+h} dN^u(s)/x_t = x\right] + E\left[E\left\{\int_{t+h}^1 c(s,x_s^u)ds + \int_{t+h}^1 dN^u(s)/x_{t+h}^u\right\}/x_t = x\right]$$

$$= E\left[\int_t^{t+h} c(s,x_s^u)ds + \int_t^{t+h} dN^u(s)/x_t = x\right] + E\left[V(t+h,x_{t+h}^u)/x_t = x\right]$$

(by step 2).

On the other hand if equality holds in the first statement of the Theorem, we take $h = 1 - t$ to obtain

$$V(t,x) = E\left\{\int_t^1 c(s,x_s^u)ds + \int_t^1 dN^u(s)/x_t = x\right\}.$$

Thus by definition of $V(t,x)$, u is optimal. This concludes the proof of the Theorem. ◻

We can restate the above Theorem in the form of the following equivalent corollary which is in the form of the Abstract Optimality Principle.

Corollary 2.3.1.

For all $u \in \mathcal{M}$, for all $0 \leq t \leq t+h \leq 1$, the function V satisfies:

$$V(t,x_t^u) \leq E\left\{\int_t^{t+h} c(s,x_s^u)ds + \int_t^{t+h} dN^u(s)/x_t^u\right\} + E\left\{V(t+h,x_{t+h}^u)/x_t^u\right\}$$

a.s. with equality iff $u \in \mathcal{M}$ is optimal.

Proof: Follows immediately from the above Theorem. ◻

All the results of section 2.2 now apply. In particular we have the following corollary.

Corollary 2.3.2.

For all $u \in \mathcal{M}$, the process $V(t,x_t^u) + \int_0^t c(s,x_s^u)ds + \int_0^t dN^u(s)$ is an $\mathcal{F}_t$-submartingale and a martingale iff u is optimal.

Proof: Follows immediately from Corollary 2.3.1. ◻

We now proceed as in Section 2.2. However we make the following observation: We construct the process $w(u,t)$ as we did in Section 1.4 and associate the processes $\bar{A}_t(u)$ and $\bar{m}_t(u)$. But now, just as in Theorem 2.5 of Boel [7] we can show that $\bar{A}_t^{t+h}(u)$ and $\bar{m}_t^{t+h}(u)$ depend only on $\mathcal{F}_t^{t+h} = \sigma\{B_s;\ t \le s \le t+h\}$. Thus Theorem 2.2.2 takes on the following version.

Theorem 2.3.2

There exists a constant J_M and for all controls $u \in \mathcal{M}$ a process $\bar{A}_t(u) \in \mathcal{A}(\mathcal{F}_t, \mathcal{P})$ such that $\bar{A}_t^{t+h}(u)$ is $\mathcal{F}_t^{t+h}$-Meas. satisfying

1) $E\,\bar{A}_1(u) = J_M$

2) For all $0 \le t \le t+h \le 1$

$$E\left[-\bar{A}_t^{t+h}(u) + \int_t^{t+h} c(s,x_s^u)ds + \int_t^{t+h} dN^u(s)/x_t^u\right] \ge 0 \text{ a.s.}$$

A control law $u = u^*$ is optimal iff 2 holds with equality for u^*.
Then $J_M = J(u^*) = \inf_{u \in \mathcal{M}} J(u)$ and

$$V(t,x_t^{u^*}) = E\,[\bar{A}_t^1(u^*)/x_t^{u^*}].$$

Proof: as in Theorem 2.2.2. ◻

We can now derive a local version of this Theorem which is the analog of Theorem 2.2.3. However it is necessary to restrict attention to value decreasing controls.

Theorem 2.3.3.

There exists a constant J_M and for all value decreasing controls $u \in \mathcal{U}$, processes α_s^u, β_s^u, γ_s^u satisfying the following conditions:

1(a) $E\int_0^1 \gamma_s^u d\, B_s = 0$ (b) $\int_0^1 |\gamma_s^u|^2\, ds < \infty$ a.s.

2) $c(s,x_s^u) - \alpha_s^u \geq 0$ a.s.

3) $\beta_s^u \leq 1$ a.s. $dN^u(s)$

4) $V_1(u) = 0$ for all u where

$$V_t(u) = J_M - \int_0^t \alpha_s^u\, ds - \int_0^t \beta_s^u dN^u(s) + \int_0^t \gamma_s^u\, dB_s.$$

A control law $u^* \in \mathcal{U}$ is optimal if and only if 2) and 3) hold with equality for u^*. Then $J_M = J(u^*)$ the cost of the optimal Markov control and $V_t(u^*) = V(t,x_t^{u^*})$ the value function evaluated along u^*.

Proof: As in Theorem 2.2.3. ¤

We note that in the above theorem we can say that the processes α_s^u, β_s^u, γ_s^u are in fact obtained from measurable functions $\alpha, \beta, \gamma: [0,1] \times \mathbb{R}^P \to \mathbb{R}$ by

$$\alpha_s^u = \alpha(s,x_s^u),\ \beta_s^u = \beta_1(s,x_s^u), \gamma_s^u = \gamma(s,x_s^u).$$

This is because of the observation made just prior to Theorem 2.3.2. We have not done this because it makes the notation intolerable.

2.4 Characterization of V under Differentiability Hypothesis.

In this section we use Ito's rule to show that the value function can be characterized as the solution of a partial differential equation with inequality constraints. In order to state the differentiability hypotheses we need some terminology from distribution theory. For definitions of all the terms used below refer to Rudin [17] or to Yosida [23].

By condition D_β on a function $V: [0,1] \times \mathbb{R}^P \to \mathbb{R}_+$ which is continuous and bounded, we mean D_β: The first and second partial distribution derivatives of V in the x variable are in $L^\beta([0,1] \times \mathbb{R}^P)$ and the first partial distribution derivative with respect to t is in $L^\beta([0,1] \times \mathbb{R}^P)$.

The imposition of condition D_β on V is weaker than imposing continuous differentiability. A Theorem of Rishel [16] states that a version of the Ito rule is valid if V satisfies the weak differentiability hypothesis D_β for some $\beta > 1$. If V satisfies D_β for $\beta > 1$ we define the action of the partial differential operator $\Lambda(t,x)$ on V by:

$$\Lambda(t,x)\ [V] = \frac{\partial V}{\partial t}(t,x) + \sum_i \mu_i(t) \frac{\partial V}{\partial x_i}(t,x) + \sum_{i,j} \sigma_i \sigma_j \frac{\partial^2 V}{\partial x_i \partial x_j}(t,x)$$

Using this notation, we have the following Theorem.

<u>Theorem 2.4.1.</u>

Assume there exists a bounded continuous function $V:[0,1] \times \mathbb{R}^P \to \mathbb{R}_+$ and a control law u* satisfying the following conditions:

1) $V(1,x) = 0$ for all $x \in \mathbb{R}^P$

2) V satisfies condition D_β for some $\beta > 1$

3) $c(t,x_t^u) + \Lambda(t,x_t^u)[V] \geq 0 = c(t,x_t^{u*}) + \Lambda(t,x_t^{u*})[V]$ a.s.

4) $V(T_k^u, x_{T_k^u}) - V(T_k^u, x^u_{T_k^u}) \leq 1 = V(T_k^{u*}, x^{u*}_{T_k^{u*}}) - V(T_k^{u*}, x^{u*}_{T_k^{u*}})$ a.s.

where 3) and 4) hold for all controls u.

Then u* is optimal, V is the Markovian value function and $V(0,0) = J_M$ the cost of the optimal control law.

<u>Proof:</u> Fix $t \in [0,1]$. Fix a control law u. Let T_k^u be the jump times of u. Define the sequence of stopping times S_k by $S_k = t \wedge T_k^u$. Then S_k is an increasing sequence of stopping times and $S_k \uparrow t$. Now we have:

$$\text{(a)}\quad V(t,x_t^u) - V(0,0) = \sum_{k=0}^{\infty} [V(S_{k+1}^-, x_{S_{k+1}^-}^u) - V(S_k, x_{S_k}^u)] - \sum_{k=0}^{\infty} [V(S_{k+1}^-, x_{S_{k+1}^-}^u) - V(S_k, x_{S_k}^u)]$$

But the 2^{nd} term on the right hand side of the equality (a) is

$$\int_0^t [V(s-, x_{s-}^u) - V(s, x_s^u)]\, dN^u(s)$$

$$= \int_0^t [V(s, x_{s-}^u) - V(s, x_s^u)]\, dN^u(s)$$

in view of the continuity of V in the t variable. Thus equality (a) becomes

$$\text{(b)}\quad V(t,x_t^u) = V(0,0) + \sum_k [V(S_{k+1}, x_{S_{k+1}^-}^u) - V(S_k, x_{S_k}^u)]$$

$$- \int_0^t [V(s, x_{s-}^u) - V(s, x_s^u)]\, dN^u(s)$$

From Theorem 2 of Rishel [16], it follows that

$$E\int_{S_k}^{S_{k+1}} \Lambda(s, x_s^u)\, [V]\, ds = E[V(S_{k+1}, x_{S_{k+1}^-}^u) - V(S_k, x_{S_k}^u)]$$

Thus taking expectations in (b) we obtain

$$\text{(c)}\quad E\,V(t,x_t^u) = V(0,0) + E\sum_k \int_{S_k}^{S_{k+1}} \Lambda(s,x_s^u)\,[V]\,ds - E\int_0^t [V(s,x_{s-}^u) - V(s,x_s^u)]\,dN^u(s)$$

or carrying out the summation in the 2^{nd} term on the right, we obtain

$$E\,V(t,x_t^u) = V(0,0) + E\int_0^t \Lambda(s,x_s^u)[V]\,ds - E\int_0^t [V(s,x_{s-}^u) - V(s,x_s^u)]\,dN^u(s)$$

Now put $V(0,0) = J^*$, $-\Lambda(s,x_s^u)[V] = \alpha_s^u$ and $[V(s,x_{s-}^u) - V(s,x_s^u)] = \beta_s^u$. Then it is seen that the conditions of the sufficiency part of Theorem 1.5.1. are satisfied. This proves the Theorem. ¤

Corollary 2.4.1.

The V appearing in the preceeding Theorem is unique..

Proof: Clear since the V is the Markovian value function. ¤

A Partial Converse of the above Theorem is provided by the next Theorem. Here we assume that the Markovian value function V satisfies condition D_β. Towards this end we have already shown that V is Lipshitz in x and absolutely continuous in t (due to the integral representation for it in Theorem 2.3.3). In the next chapter we deduce some differentiability properties of V under alternate hypotheses.

Theorem 2.4.2.

Assume that the Markovian value function V satisfies condition D_β for some $\beta > 1$. Then u* is an optimal control implies for all value decreasing controls u, the Markovian value function satisfies.

1) $c(t,x_t^u) + \Lambda(t,x_t^u)[V] \geq 0 = c(t,x_t^{u*}) + \Lambda(t,x_t^{u*})\ [V]$ a.s.

2) $V(T_k^u, x^u_{T_k^u -}) - V(T_k^u, x^u_{T_k^u}) \leq 1 = V(T_k^{u*}, x^{u*}_{T_k^{u*} -}) - V(T_k^{u*}, x^{u*}_{T_k^{u*}})$ a.s.

3) $V(1,x) = 0$ for all $x \in \mathbb{R}^P$

Proof:

Condition 3) follows immediately. Let now u* be optimal and fix a value decreasing control u. By the same argument as in the previous Theorem we have

$$E\,V(t,x_t^u) = V(0,0) + E\int_0^t \Lambda(s,x_s^u)[V]\,ds - E\int_0^t [V(s,x_{s-}^u) - V(s,x_s^u)]dN^u(s)$$

We can now use the optimality principle to conclude that

$$-A_t(u) + \int_0^t c(s,x_s^u)\,ds + \int_0^t dN^u(s) \text{ is a submartingale where}$$

$$A_t(u) = \int_0^t -\Lambda(s,x_s^u)\,ds + \int_0^t [V(s,x_{s-}^u) - V(s,x_s^u)]\,dN^u(s)$$

The technique which was used in the necessary part of Theorem 1.4.2 yields the result. This concludes the proof of the Theorem. ◻

Theorem 2.4.1. is to be compared with Theorem 2 of Bensoussan and Lions [4]. It is a slightly stronger version of their Theorem since we only need their "Differential Inequalities" to be satisfied along the admissible trajectories generated by control laws.

CHAPTER 3

DIFFERENTIABILITY PROPERTIES OF V UNDER ADDITIONAL HYPOTHESES

In this chapter we show that the Markovian Value function V of the Impulse Control problem introduced in Chapter 2, is continuously differentiable in x under suitable assumptions on the cost function c and the set U of admissible jump heights. We will further assume that V satisfies a condition which makes it into a "good" function. For ease of exposition we assume that $\mu(t) = 0$.

We know that V is a bounded continuous function $V : [0,1] \times \mathbb{R}^p \to \mathbb{R}_+$. From 1) of Corollary 1.3.4 it follows that

$$V(t,x) \leq 1 + \inf_{u \in U} V(t,x-u) \quad \text{for all } (t,x) \in [0,1] \times \mathbb{R}^p$$

Let $C = \{(t,x) \in [0,1] \times \mathbb{R}^p / V(t,x) < 1 + \inf_{u \in U} V(t,x-u)\}$.

Let $S = \tilde{C}$ (viz. the complement of C). Note that C is the set of all points in (t,x)-space where a change in control value is not advisable. Thus an optimal control if it exists cannot jump in C. The region C may therefore be called the continuation region. On the other hand in the region S it is imperative that we change the control value. Intuitively an optimal control should jump as soon as the trajectory reached the boundary of C and the height of the jump should be the u which minimizes V(t,x-u). For a Theorem regarding the optimality of this policy refer to Bensoussan and Lions [4]. We now make the following assumptions:

<u>A1</u> The set C is non-empty, its complement S has non-empty interior and $\partial\bar{C} = \partial S$ is regular (where ∂ denotes boundary and the bar over C denotes closure).

A2 For all $(t,x) \in S$, there exists a unique $u^*(t,x) \in U$ such that $\inf_{u \in U} V(t,x-u) = V(t,x-u^*(t,x))$ and for each fixed t the map $(t,x) \to u^*(t,x)$ from $\mathbb{R}^p \to U$ is continuous.

A3 The set $U \subseteq \mathbb{R}_+^p$ of admissible jump values is a cone.

A4 The cost function c is continuously differentiable in both variables.

A5 There exists an optimal control.

A1, A2 and A3 are similar to the assumption of Bensoussan and Lions [4] that V be a good function. In Assumption A1, what we mean by the regularity of the boundary becomes clear in the analysis between Lemma 3.2 and Theorem 3.1. We note that the continuity of V implies that C is an open set. We prepare the proof of differentiability with the following Lemmas which say that V is continuously differentiable in C and in Int S (where Int denotes "Interior.")

Lemma 3.1. V is continuously differentiable in x in the region C.

Proof: Fix $(t,x) \in C$. Since C is open there is a Neighbourhood (Nbd) of (t,x), $N(t,x)$, s.t. $N(t,x) \subset C$. Let S be the first exit time for the Brownian Motion which starts from x at time t from the Nbd $N(t,x)$. Now since in C, V satisfies $V(t,x) < 1 + V(t,x-u^*(t,x))$, it follows that the optimal control law does not jump in C, in particular in $N(t,x)$. Thus the optimality Principle yields.

$$V(T,B_T) = E\{\int_T^S c(s,\ B_s)ds/B_T\} + E\ V(S,B_S)$$

for all stopping times T such that $t \leq T \leq S$. Now since c is continuously differentiable in both variables, it follows from the solution of Problem 6, pg. 59 of McKean [12] that V is continuously differentiable in x in the

Nbd $N(t,x)$. This completes the proof of the Lemma ⌑

Next we show that V is continuously idfferentiable in x (for fixed t) in int S = int $\tilde{C}$. It will then remain to show that V is differentiable on ∂S.

Lemma 3.2. For each t, V is continuously differentiable in x on int S.

Proof: The proof is in two steps. We first show that for each fixed t, $(t,x) \in S \Rightarrow x - u^*(t,x) \in C$. Then we use this to establish the statement of the lemma

Step 1: Fix $(t,x) \in S$. Let $z = x - u^*(t,x)$. We need to show that $z \in C$. Now since $(t,x) \in S$ we have

(a) $V(t,x) = 1 + V(t,z)$. If z is not in C then $z \in S$. In that case we have

(b) $V(t,z) = 1 + \inf_{u \in U} V(t,z-u)$. Now since U is a cone in $\mathbb{R}^p_+$, every point which is accessible from z is accessible from x. Thus $\inf_{u \in U} V(t,z-u) \geq \inf_{u \in U} V(t,x-u) = V(t,z)$. So certainly we have $1 + \inf_{u \in U} V(t,z-u) > V(t,z)$ which contradicts (b). Thus we must have $z \in C$. This concludes the proof of Step 1.

Step 2: We now show that for fixed t, V is continuously differentiable in x for $x \in$ int S. Towards this end fix $(t,x) \in$ int S. Let $u = u^*(t,x)$. Then by step 1, $x-u \in C$. Now since C and int S are disjoint open sets, there are Nbds. N of x-u and N_1 of x in $\mathbb{R}^p$ such that $N \subset C$ and $N_1 \subset$ int S. By the continuity of the addition Map on $\mathbb{R}^p$, there exists a Nbd. N_2 of x in $\mathbb{R}^p$ such that $x'-u \in N$ for every $x' \in N_2$. By the continuity of the map u^* as assumed in A2, there exists a Nbd N_3 of x such that $x'-u^*(t,x') \in N$ for every $x' \in N_3$. Let now $N' = N_1 \cap N_2 \cap N_3$. We shall work locally in N'. Let $x' \in N'$. The we have:

$$V(t,x') - V(t,x) = [1 + V(t,x'-u')] - [1 + V(t,x-u)] \quad \text{where } u' = u^*(t,x') \text{ and } u = u^*(t,x)$$

$$= V(t,x'-u') - V(t,x-u)$$

$$= [V(t,x'-u') - V(t,x'-u)] + [V(t,x'-u) - V(t,x-u)] \qquad \text{(a)}$$

Now $x'-u'$, $x'-u$, $x-u \in C$ and by the previous lemma V is continuously differentiable in C. Therefore taking recourse to a Taylor expansion (a) can be rewritten as

$$V(t,x') - V(t,x) = \left.\frac{\partial V}{\partial x}\right|_z \cdot (u-u') + \left.\frac{\partial V}{\partial x}\right|_{x-u} \cdot (x'-x) + o(x'-x) \qquad \text{(b)}$$

where z is an interior point of the line segment joining $x'-u'$ to $x'-u$. We show that the first term on the right in (b) is zero. Towards this end consider

$$V(z) \leq V(z+ (u-u')) \leq \varepsilon \left.\frac{\partial V}{\partial x}\right|_z \cdot (u-u') + V(z) \qquad \text{(c)}$$

for all $|\varepsilon|$ sufficiently small. We note that the first inequality in (c) holds for ε both positive and negative by the continuity of the map u^* and because U is a cone.

Thus from (c) we obtain

$$0 \leq \varepsilon \left.\frac{\partial V}{\partial x}\right|_z \cdot (u-u') \quad \text{for all } |\varepsilon|. \text{ sufficiently small}$$

Taking $\varepsilon > 0$ and $\varepsilon < 0$ we see that

$$\left.\frac{\partial V}{\partial x}\right|_z \cdot (u-u') = 0$$

Thus (b) becomes

$$V(t,x') - V(t,x) = \frac{\partial V}{\partial x}\bigg|_{x-u}\cdot(x'-x) + o(x'-x).$$

This shows that $\frac{\partial V}{\partial x}\big|_x$ exists and equals $\frac{\partial V}{\partial x}\big|_{x-u}$. Next we show that $\frac{\partial V}{\partial x}$ is continuous at all points $x \in \text{int } S$. Towards this end fix t and let $x_n \to x$ in $\mathbb{R}^p$, where $(t,x_n) \in \text{int } S$ for all n. Let $u^*(x_n) = u^*(t,x_n)$

Then $\frac{\partial V}{\partial x}\big|_{x_n} = \frac{\partial V}{\partial x}\big|_{x_n-u^*(x_n)}$. Now since $x_n \to x$ and u^* is continuous we have $u^*(x_n) \to u^*(x)$ or $x_n - u^*(x_n) \to x - u^*(x)$. But $x_n - u^*(x_n) \in C$ for all n and V is continuously differentiable in C. Thus $\frac{\partial V}{\partial x}\big|_{x_n-u^*(x_n)} \to \frac{\partial V}{\partial x}\big|_{x-u^*(x)}$ $= \frac{\partial V}{\partial x}\big|_x$. This shows that V is continuously differentiable in x in int S. This concludes the proof of the lemma. ¤

We are now ready to prove that V is continuously differentiable in x on ∂S. In order to prove this we fix $(t,x) \in \partial S$ and consider a control law which does not jump in a small Nbd of (t,x) and follows the optimal policy outside this Nbd. This will enable us to compare the right and left partial derivatives at x, in every direction, which exist by lemmas 3.1 and 3.2. We shall see that they are equal. This will enable us to conclude that $\frac{\partial V}{\partial x}\big|_x$ exists and is continuous for $x \in \partial\bar{C} = \partial S$. Before we prove the theorem we introduce some notation and make an observation. We remark that this cumbersome notation is necessary because of the variety of cases that occur.

Fix $(t,x) \in \partial\bar{C} = \partial S$. Let e_i denote the unit vector in the direction of the $\underline{i}$th co-ordinate axis. For each ε let $\delta_i(\varepsilon) = e_i + \varepsilon e_i$. Now fix i. Since $(t,x) \in \partial\bar{C} = \partial S$ and $\partial C = \partial S$ is regular the following mutually exclusive cases arise:

1) $x + \delta_i(\varepsilon) \in S$ for all $|\varepsilon|$ sufficiently small.

2) $x + \delta_i(\varepsilon) \in C$ for all $|\varepsilon|$ sufficiently small.

3) $x + \delta_i(\varepsilon) \in S$ and $x + \delta_i(-\varepsilon) \in C$ for $\varepsilon > 0$, ε sufficiently small.

4) $x + \delta_i(-\varepsilon) \in S$ and $x + \delta_i(\varepsilon) \in C$ for $\varepsilon > 0$, ε sufficiently small.

We remark that these are the exclusive cases which arise because of the assumption that $\partial\bar{C} = \partial S$ is regular.

Now let
$$D_{e+_i}(x) = \lim_{\substack{\varepsilon \to 0 \\ \varepsilon > 0}} \frac{\partial V}{\partial x_i}(x+\delta_i(\varepsilon))$$
and let
$$D_{e-_i}(x) = \lim_{\substack{\varepsilon \to 0 \\ \varepsilon < 0}} \frac{\partial V}{\partial x_i}(x+\delta_i(\varepsilon))$$
where we have suppressed the variable t since we have fixed it.

Then by lemma 3.1 and 3.2 both $D_{e+_i}(x)$ and $D_{e-_i}(x)$ exist. We have to show that they are equal. If the index i satisfies case 1) then by lemma 3.2 we have $D_{e+_i}(x) = D_{e-_i}(x)$. If the index i satisfies case 2) then by lemma 3.1 we reach the same conclusion. The only cases that remain are cases 3) and 4). Let now I_S be the set of all indices i which satisfy case 3) and let I_C be the set of all indices i which satisfy case 4). Fix an index $i \in I_S$. Then we claim that $D_{e+_i}(x) \le D_{e-_i}(x)$. Here we recall that $(t,x) \in \partial.\bar{C}$. To see the above claim, fix $\varepsilon > 0$. Then since $i \in I_S$ we have $x + \delta_i(\varepsilon) \in S$, $x + \delta_i(-\varepsilon) \in C$ for ε sufficiently small. Since the set U of admissible control values is a cone, we have $V(x+\delta_i(\varepsilon)) \le V(x)$. Thus $D_{e+_i}(x) \le 0$. On the other hand since $(t,x) \in \partial S$, we have $V(x) \ge V(x+\delta_i(-\varepsilon))$ for ε sufficient small. Since if not we would obtain $V(x) < 1 + \inf_{u \in U} V(x-u)$ which contradicts the fact that $x \in S$. Thus $D_{e-_i}(x) \ge 0$. We have thus shown that $D_{e+_i}(x) \le 0 \le D_{e-_i}(x)$ or $D_{e+_i}(x) \le D_{e-_i}(x)$. If the index $i \in I_C$ a similar argument holds. Thus we have shown that $D_{e+_i}(x) \le D_{e-_i}(x)$ for all indices i when $(t,x) \in \partial S = \partial\bar{C}$. In the theorem we show that the reverse inequality holds.

Theorem 3.1.

For each fixed $t \in [0,1]$, the Markovian value function V is continuously differentiable in x on $\mathbb{R}^p$.

Proof: Fix t. Let $C|_t = \{x \in \mathbb{R}^p | (t,x) \in C\}$. Similarly define int $S|_t$. By lemmas 3.1 and 3.2, the statement of the theorem holds in $C|_t$ and int $S|_t$. We now show that it holds on $\partial\bar{C}|_t$. Towards this end, fix $x \in \partial\bar{C}|_t$ and fix $\varepsilon > 0$. Let $N(\varepsilon;(t,x)) = \{(s,y) \in [t,1] \times \mathbb{R}^p / |s-t| < t, |y-x| < \varepsilon\}$. Then $N(\varepsilon;(t,x))$ is a Nbd of radius ε around (t,x). Let S_ε be the first exit time of the Brownian Motion starting at x, from the Nbd. $N(\varepsilon;(t,x))$. Consider the control law which maintains a no jumps policy in $N(\varepsilon;(t,x))$ and follows the optimal policy outside $N(\varepsilon;(t,x))$. Compute the cost J of this policy. We have

$$J = E[\int_t^{S_\varepsilon} c(s,x+B_s)ds] + E[V(S_\varepsilon,x+B_{S_\varepsilon})] \le K_1\varepsilon + K_2\varepsilon + EV(t,x+B_{S_\varepsilon}) \qquad \text{(a)}$$

where K_1 is the uniform bound for c and K_2 is the time Lipschitz Constant for V. Now since S is the first exit time for the Brownian Motion starting at x from $N(\varepsilon;(t,x))$ we have $|B^i_{S_\varepsilon}| < \varepsilon$ where $B^i_{S_\varepsilon}$ is the ith component of B_{S_ε}. Thus using lemmas 3.1 and 3.2, the inequality (a) can be rewritten as

$$J \le K\cdot\varepsilon + V(t,x) + \sum_i D_{e+_i}(x)\ E[B^i_{S_\varepsilon} / 0 \le B^i_{S_\varepsilon} < \varepsilon]$$

$$+ \sum_i D_{e-_i}(x)\ E[B^i_{S_\varepsilon} / -\varepsilon < B^i_{S_\varepsilon} < 0] + o(\varepsilon^{1/2}) \qquad \text{(b)}$$

where we have used the fact that $Eo(B^i_{S_\varepsilon}) = o(\varepsilon^{1/2})$. But

$$-E[B^i_{S_\varepsilon} / -\varepsilon < B^i_{S_\varepsilon} < 0] = E[B^i_{S_\varepsilon} / 0 \le B^i_{S_\varepsilon} < \varepsilon] \le \frac{\sigma_i}{(2\pi)^{1/2}}\varepsilon^{1/2} + o(\varepsilon^{1/2})$$

Thus (b) can be rewritten as

$$J \leq K\cdot\varepsilon + V(t,x) + \sum_i [D_{e+_i}(x) - D_{e-_i}(x)][\frac{\sigma_i}{(2\pi)^{1/2}} \varepsilon^{1/2} + o(\varepsilon^{1/2})] + o(\varepsilon^{1/2}). \quad \text{(c)}$$

Now looking at (c) we see that if any term under the summation sign is strictly negative we can take ε sufficiently small and obtain $J < V(t,x)$ which contradicts the fact that V is the value function. Thus we have $D_{e+_i}(x) \geq D_{e-_i}(x)$ for all i. Combining this with the result $D_{e+_i}(x) \leq D_{e-_i}(x)$ proved just before the statement of the theorem, we obtain $D_{e+_i}(x) = D_{e-_i}(x)$. We have thus shown that V is continuously differentiable at x in every co-ordinate direction. This concludes the proof of the theorem. ¤

We remark that when the state space $\mathbb{R}^p$ is one dimensional, the assumption that int S is non-empty implies that V is constant in x for each fixed t for $x \in$ int S. This statement is fairly easily proved. We shall not go into details here. Thus in the one dimensional case we obtain that V is twice continuously differentiable in int $S|_t$ and in fact both the derivatives are 0. Furthermore the technique used in lemma 3.1 could be used to show twice continuous differentiability on $C|_t$. Furthermore the technique used in Theorem 3.1 could be used to show that V is twice continuously differentiable on $\partial\bar{C}|_t$. In addition the technique used in Theorem 3.1 could be used as a basis for discovering a heuristic algorithm which tells us whether we are following an optimal policy or not. This would be particularly useful if we restricted ourselves to (s,S) policies. Then we could discover whether the reorder point we operate on should be increased or decreased in order to maintain the equality of the derivatives.

CHAPTER 4

CONCLUDING REMARKS

In chapter one we have considered a fairly general form of the stochastic control problem and a class of control laws which are important whenever physical or institutional constraints require that controls cannot be changed continuously. This is particularly meaningful in an Economic Context when control laws are modelled as prices or have components which represent capacity levels. Another important class of problems which can be tackled under this framework is that of optimal stopping problems and quickest detection problems. The results presented here could be combined with those of Boel [7] to obtain a theory of Impulse Control of jump processes. The detailed analysis of the continuous time Inventory control problem carried out in chapter two shows that the abstract conditions lead to equations which are difficult to solve computationally but which nevertheless indicate important properties of the optimal control as remarked at the end of Chapter 3.

REFERENCES

1. Bather, J. Appl. Prob., 3, 1966, p. 538-549.
2. Benes, V., "Existence of Optimal Strategies Based on Specified Information, for a Class of Stochastic Decision Problems," SIAM J. CONTROL, 8, 179-188, 1970.
3. Benes, V. "Existence of Optimal Stochastic Control Laws," SIAM J. CONTROL, 9, 446-475, 1971.
4. Bensoussan A., Lions J-L, "Nouvelle Formulation de problémes de contrôle impulsionnel et applications" C. R. Acad. Sc. Paris, t. 276, Serie A, 1189-1192, 2 mai 1973.
5. ________________________, "Contrôle impulsionnel et inéquations quasi-variationelles stationnaires" C. R. Acad. Sc. Paris, t. 276, Serie A, 1279-1284, 9 mai 1973.
6. ________________________, "Inèquation variationelles non lineaires du premier et du second ordre," C. R. Acad. Sc. Paris, t. 276, 1411-1415, 21 mai 1973.
7. Boel, R. K., "Optimal Control of Jump Processes," Electronics Research Laboratory, University of California, Berkeley, Calif., Memo #M448, July 1974.
8. Boel, R. K., and Varaiya, P. P., "Optimal Control fo Jump Processes." submitted to SIAM J. CONTROL, 1975.
9. Boel, R. K., Varaiya, P. P., and Wong, E., "Martingales on Jump Processes I: Representation Results," Electronics Research Laboratory, University of California, Berkeley, Calif., Memo #M407, September 1973, submitted to SIAM J. CONTROL.
10. Davis, M. H. A., and Varaiya, P. P., "Dynamic Programming Conditions for Partially Observable Stochastic Systems," SIAM J. CONTROL, 11,

226-261, 1973.

11. Kushner, H, "Stochastic Stability and Control," New York: Academic Press, New York, London, 1969.

12. McKean, H. P., Jr, "Stochastic Integrals," Academic Press, New York, London, 1969.

13. Meyer, P. A., "Probability and Potentials," Blaisdell Publishing Co., 1966.

14. ____________, "Temps d'arret algebriquement prévisibles," Séminaire de Probilités VI, Lecture Notes in Mathematics, #258, Berlin: Springer Verlag, 158-163, 1972.

15. Rishel, R. W, "Necessary and Sufficient Dynamic Programming Conditions for Continuous Time Stochastic Optimal Control," SIAM J. CONTROL, 6, 131-147, 1968.

16. ____________, "Weak Solutions of a Partial Differential Equation of Dynamic Programming," SIAM J. CONTROL, 9, #4, 519-528, 1971.

17. Rudin, W, "Functional Analysis," McGraw Hill, 1973.

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

19. Scarf, H, Math. Math. in the Social Sciences, Stanford University Press, Stanford, Calif., 1960.

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

21. Vienott, J. SIAM, App. Math., 14, #5, September 1966.

22. Wong, E, "Representation of Martingales, Quadratic Variation and Applications," SIAM J. CONTROL, 9, 621-632, 1971.

23. Yosida, K, "Functional Analysis," Springer Verlag, vol. 123, 1965.