Printed at the Mathematical Centre, Kruislaan 413, Amsterdam, The Netherlands.

The Mathematical Centre, founded 11 February 1946, is a non-profit institution for the promotion of pure and applied mathematics and computer science. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

MATHEMATICAL CENTRE TRACTS 164

MARKOV DECISION PROCESSES WITH CONTINUOUS TIME PARAMETER

F.A. VAN DER DUYN SCHOUTEN

MATHEMATISCH CENTRUM AMSTERDAM 1983

1980 Mathematics subject classification: 60B10, 60G17, 60J25, 90B05, 90B25, 90C47, 93E20

ISBN 90 6196 261 7

ACKNOWLEDGEMENT

With thanks I acknowledge the support provided by Arie Hordijk, Manfred Schäl and Henk Tijms during the preparation of this monograph.
In particular the suggestions of Arie Hordijk resulted in many substantial improvements.
I am indebted to the "Stichting Mathematisch Centrum" for offering me the opportunity to publish this book in the series "Mathematical Centre Tracts".

CONTENTS

CHAPTER 1

PROBABILITY THEORY ON METRIC SPACES

1.1. INTRODUCTION.

Since a stochastic process can be seen as a probability measure on a function space, weak convergence of stochastic processes is equivalent with weak convergence of probability measures on metric spaces. Which function space has to be considered depends on the behaviour of the sample paths of the stochastic processes under consideration. PROHOROV (1956) and BILLINGSLEY (1968) studied among others the function space C[0,1], being the class of all continuous functions on the unit interval with values in some metric space S. Weak convergence on the space $C[0,\infty)$ of S-valued continuous functions defined on $[0,\infty)$ has been treated by STONE (1963) and WHITT (1970). Another important function space is $D[0,\infty)$, the set of all S-valued functions on $[0,\infty)$, which are right continuous and have left hand limits at every $t>0$. For an extensive study of this space we refer to LINDVALL (1973) and WHITT (1980).

In the theory of controlled stochastic processes, however, we encounter processes, which cannot be seen as random elements of $C[0,\infty)$ or $D[0,\infty)$, since the sample paths are neither right- nor left continuous. Although most of these processes can be put into the framework of the function spaces $C[0,\infty)$ or $D[0,\infty)$ by methods ad-hoc, a unifying approach to this kind of controlled stochastic processes requires the introduction of a new sample space. This will be illustrated by some examples in section 1.3.

1.2. GENERAL CONCEPTS.

In this section we give an outline of some general concepts of probability theory on metric spaces. For a more extensive introduction the reader is referred to BILLINGSLEY (1968) or PARTHASARATHY (1967). The concepts and results of this section are rather disconnected, but have in common, that they all will be used later on. In this section S (or S_i, i=1,2) is a metric space and $\mathcal{S}$ ($\mathcal{S}_i$, i=1,2) its Borelfield i.e. the σ-field generated by the open subsets of S.

<u>DEFINITION 1.2.1</u>. A function h from S_1 into S_2 is called *measurable* if $h^{-1}\mathcal{S}_2 \subset \mathcal{S}_1$ and a function f from S_1 into S_2 is called *h-measurable* if $f^{-1}\mathcal{S}_2 \subset h^{-1}\mathcal{S}_2$.

<u>DEFINITION 1.2.2</u>. A sequence of probability measures $(P_n)_{n=1}^{\infty}$ on S *converges weakly* to a probability measure P on S if

$$\lim_{n\to\infty} \int_S f \, dP_n = \int_S f \, dP,$$

for all bounded continuous real-valued functions f on S.
Notation: $P_n \overset{w}{\to} P$.

<u>THEOREM 1.2.3</u>. (Portmanteau theorem). Let P, P_n, n≥1 be probability measures on S. The following five assertions are equivalent:

(i) $P_n \overset{w}{\to} P$.

(ii) $\lim_{n\to\infty} \int_S f \, dP_n = \int_S f \, dP$ for all bounded, uniformly continuous real-valued functions f.

(iii) $\limsup_{n\to\infty} P_n(F) \le P(F)$ for all closed $F \in \mathcal{S}$.

(iv) $\liminf_{n\to\infty} P_n(G) \ge P(G)$ for all open $G \in \mathcal{S}$.

(v) $\lim_{n\to\infty} P_n(B) = P(B)$ for all $B \in \mathcal{S}$ for which $P(\delta B) = 0$, where δB denotes the boundary of the set B.

PROOF. See page 12 and 13 of BILLINGSLEY (1968). □

DEFINITION 1.2.4. A *random element* of S is a measurable function X from some probability space $(\Omega, \mathcal{F}, P)$ into S (measurable means $X^{-1}\mathcal{S} \subset \mathcal{F}$). The probability measure PX^{-1} on S defined by

$$PX^{-1}(B) := P(X^{-1}B) := P(X \in B) \text{ for } B \in \mathcal{S}$$

is called the *probability measure induced by X* or the *probability distribution of X.*

If S is a function space then a random element is often called a *random function* or *stochastic process*. If $S = \mathbb{R}$ we call a random element a *random variable* and if $S = \mathbb{R}^k$, $k > 1$ a *random vector*.
If random elements of S are introduced we often omit to specify explicitly the underlying probability space $(\Omega, \mathcal{F}, P)$ on which the random elements are defined. We always assume that all random elements under consideration are defined on one common probability space.

DEFINITION 1.2.5. A sequence of random elements $(X_n)_{n=1}^{\infty}$ of S *converges in distribution* to a random element X of S if

$$PX_n^{-1} \overset{w}{\to} PX^{-1}$$

Notation: $X_n \overset{d}{\to} X$.

We write $X \overset{d}{=} Y$ if the random elements X and Y have the same probability distribution.

NOTATION 1.2.6. If h is a function from S_1 into S_2 then Disc (h) is the set of points in S_1 where h is discontinuous.

PROPOSITION 1.2.7. Disc (h) $\in \mathcal{S}_1$ even if h is not measurable.

PROOF. See page 225 of BILLINGSLEY (1968). □

DEFINITION 1.2.8. A *transition probability from* S_1 *to* S_2 is a mapping P from $S_1 \times \mathcal{S}_2$ into $[0,1]$, such that

(i) $P(s_1,.)$ is a probability measure on S_2 for every $s_1 \in S_1$

(ii) $P(.,A_2)$ is measurable on S_1 for every $A_2 \in \mathcal{S}_2$.

The set of all probability measures on S will be denoted by $\mathcal{P}(S)$. Hence, a transition probability from S_1 to S_2 can be considered as a mapping from S_1 into $\mathcal{P}(S_2)$.

THEOREM 1.2.9. (Continuous mapping theorem). Let h be a measurable function from S_1 into S_2 and X, X_n, $n \geq 1$, random elements of S_1. If $X_n \overset{d}{\to} X$ and $PX^{-1}(\text{Disc }(h)) = 0$ then

$$h(X_n) \overset{d}{\to} hX$$

as random elements of S_2.

PROOF. See page 30 of BILLINGSLEY (1968). □

DEFINITION 1.2.10. Let X, X_n, $n \geq 1$, be random elements of S. Then $(X_n)_{n=1}^{\infty}$ *converges almost surely* to X, if X, X_n, $n \geq 1$, are all defined on a common probability space $(\Omega,\mathcal{F},P)$ and if there exists a set $\Omega_0 \in \mathcal{F}$ with $P(\Omega_0) = 1$ such that $\lim_{n\to\infty} X_n(\omega) = X(\omega)$ for all $\omega \in \Omega_0$. Notation: $X_n \overset{a.s.}{\to} X$.

LEMMA 1.2.11. Let X, X_n, $n \geq 1$, be random elements of S. If $X_n \overset{a.s.}{\to} X$ then $X_n \overset{d}{\to} X$.

PROOF. For every continuous real-valued function f on S we have $f(X_n) \overset{a.s.}{\to} f(X)$. If, moreover, f is bounded then $\int_\Omega f(X_n)\, dP \to \int_\Omega f(X)\, dP$, by Lebesgue's theorem on bounded convergence. But $\int_\Omega f(X_n)\, dP = \int_S f\, dPX_n^{-1}$ and $\int_\Omega f(X)\, dP = \int_S f\, dPX^{-1}$. Hence $PX_n^{-1} \overset{w}{\to} PX^{-1}$ which implies by definition $X_n \overset{d}{\to} X$. □

REMARK 1.2.12. The converse of lemma 1.2.11 is not true even if X, X_n, $n \geq 1$, are defined on a common probability space. For instance let $P(X=1) = P(X=-1) = \frac{1}{2}$ and define $X_n = (-1)^n X$. Then $X_n \overset{d}{=} X$ for all n, whereas $(X_n)_{n=1}^{\infty}$ does not converge almost surely.

THEOREM 1.2.13. Let X, X_n, $n \geq 1$, be random elements of a complete and separable space S. If $X_n \overset{d}{\to} X$ then there exist random elements X', X'_n, $n \geq 1$, of S defined on a common probability space such that

$$X'_n \overset{d}{=} X_n, \qquad n \geq 1$$

$$X' \overset{d}{=} X$$

and

$$X'_n \overset{a.s.}{\to} X'.$$

PROOF. See SKOROHOD (1956). □

DUDLEY (1968) and WICHURA (1970) proved this theorem under weaker assumptions.
A useful application of theorem 1.2.13. is put into words in theorem 1.2.16. below. First we need the following definitions.

DEFINITION 1.2.14. Let (Ω, F, P) be a probability space. An assertion H holds *almost everywhere w.r.t. P* if there exists a $F \in F$ with $P(F) = 1$ such that H holds for all $x \in F$. Notation: H holds P-a.e.

DEFINITION 1.2.15. Let f, f_n, $n \geq 1$, be measurable functions from S_1 into S_2. The sequence $(f_n)_{n=1}^{\infty}$ is said to be *continuously convergent at x* to the function f if $\lim_{n\to\infty} x_n = x$ implies $\lim_{n\to\infty} f_n(x_n) = f(x)$. The sequence $(f_n)_{n=1}^{\infty}$ is *continuously convergent* to f if it is continuously convergent at all $x \in S_1$ to the function f. Notation: $f_n \overset{c}{\to} f$ (see for example page 197 of KURATOWSKI (1966).

THEOREM 1.2.16. Let X, X_n, $n \geq 1$, be random elements of S_1 and h, h_n, $n \geq 1$, measurable functions from S_1 into S_2. Assume that S_1 is separable and complete. If $X_n \overset{d}{\to} X$ and $h_n \overset{c}{\to} h$, PX^{-1}-a.e., then

$$h_n(X_n) \overset{d}{\to} h(X)$$

as random elements of S_2.

PROOF. According to theorem 1.2.13. there exist random elements X', X_n', $n\geq 1$, of S_1 such that $X_n' \overset{d}{=} X_n$, $X' \overset{d}{=} X$ and $X_n' \overset{a.s.}{\to} X'$. But then $h_n(X_n') \overset{a.s.}{\to} h(X')$. From lemma 1.2.11. it follows that $h_n(X_n') \overset{d}{\to} h(X')$ and consequently $h_n(X_n) \overset{d}{\to} h(X)$ since $h_n(X_n') \overset{d}{=} h_n(X_n)$ and $h(X') \overset{d}{=} h(X)$. □

We will apply this theorem in the setting of the next corollary.

COROLLARY 1.2.17. Let P, P_n, $n\geq 1$, be probability measures on S_1 and h, h_n, $n\geq 1$, measurable functions from S_1 into S_2. Assume that S_1 is separable and complete. If $P_n \overset{w}{\to} P$ and $h_n \overset{c}{\to} h$, P-a.e. then

$$P_n h_n^{-1} \overset{w}{\to} Ph^{-1}$$

as probability measures on S_2.

A useful notion in the theory of convergence in distribution of random variables is the concept of uniform integrability.

DEFINITION 1.2.18. A sequence of random variables $(X_n)_{n=1}^{\infty}$ is *uniformly integrable* if all X_n are defined on a common probability space (Ω,F,P) and

$$\lim_{\alpha\to\infty} \sup_n \int_{\{|X_n|\geq\alpha\}} |X_n| \, dP = 0.$$

The following theorem can be found on page 32 of BILLINGSLEY (1968).

THEOREM 1.2.19. Suppose that X, X_n, $n\geq 1$, are random variables such that $X_n \overset{d}{\to} X$.

(i) If $(X_n)_{n=1}^{\infty}$ is uniformly integtable then $EX_n \to EX$.

(ii) If X and X_n, $n\geq 1$, are non-negative and integrable then $EX_n \to EX$ implies that $(X_n)_{n=1}^{\infty}$ is uniformly integrable.

Combining this theorem with theorem 1.2.16. we get a useful convergence theorem.

THEOREM 1.2.20. Let X, X_n, $n\geq 1$, be random elements of a complete separable space S and f, h, f_n, h_n, $n\geq 1$, real-valued measurable functions on S. Suppose that $X_n \overset{d}{\to} X$ and that $f_n \overset{c}{\to} f$, PX^{-1}-a.e. and $h_n \overset{c}{\to} h$, PX^{-1}-a.e. If $E\, h_n(X_n) \to E\, h(X)$ and $|f_n(x)| \le h_n(x)$ for all $x \in S$ then $E\, f_n(X_n) \to E\, f(X)$.

PROOF. From theorem 1.2.16. it follows that $f_n(X_n) \overset{d}{\to} f(X)$ and $h_n(X_n) \overset{d}{\to} h(X)$. Since $h_n(X_n)$ and $h(X)$ are non-negative we find from theorem 1.2.19.(ii) that $(h_n(X_n))_{n=1}^{\infty}$ is uniformly integrable. Hence $(f_n(X_n))_{n=1}^{\infty}$ is uniformly integrable. The theorem follows from theorem 1.2.19.(i). □

COROLLARY 1.2.21. Let P, P_n, $n\ge1$, be probability measures on a complete separable space S and f, h, f_n, h_n, $n\ge1$, real-valued measurable functions on S. Suppose that $P_n \overset{w}{\to} P$ and that $f_n \overset{c}{\to} f$, P-a.e. and $h_n \overset{c}{\to} h$, P-a.e. If

$$\lim_{n\to\infty} \int h_n \, dP_n = \int h \, dP$$

and

$$|f_n(x)| \le h_n(x) \qquad \text{for all } x \in S$$

then

$$\lim_{n\to\infty} \int f_n \, dP_n = \int f \, dP.$$

Next theorem can be seen as a generalization of Fatou's lemma.

THEOREM 1.2.22. Let X, X_n, $n\ge1$, be random elements of a complete separable space S and f, f_n, $n\ge1$, non-negative, measurable functions on S. Assume that $X_n \overset{d}{\to} X$. If for any sequence $(x_n)_{n=1}^{\infty}$ with $\lim_{n\to\infty} x_n = x$ we have

$$\liminf f_n(x_n) \ge f(x), \qquad PX^{-1}\text{-a.e.}$$

then

$$\liminf E\, f_n(X_n) \ge E\, f(X).$$

PROOF. From theorem 1.2.13. follows the existence of random elements X', X'_n, $n\ge1$ of S such that $X'_n \overset{d}{=} X_n$, $X' \overset{d}{=} X$ and $X'_n \overset{a.s.}{\to} X$!

Hence

$$E\, f_n(X_n) = \int_S f_n \, dPX_n'^{-1} = \int_\Omega f_n(X_n') \, dP.$$

With Fatou's lemma follows

$$\lim_{n\to\infty} E\, f_n(X_n) \geq \int_\Omega \liminf_n f_n(X_n') \, dP$$

and by assumption

$$\int_\Omega \liminf_n f_n(X_n') \, dP \geq \int_\Omega f(X') \, dP = E\, f(X). \qquad \square$$

COROLLARY 1.2.23. Let P, P_n, $n\geq 1$, be probability measures on a complete separable space S and f, f_n, $n\geq 1$, non-negative measurable functions on S. Assume that $P_n \overset{w}{\to} P$. If for any sequence $(x_n)_{n=1}^{\infty}$ with $\lim_{n\to\infty} x_n = x$ we have

$$\liminf_{n\to\infty} f_n(x_n) \geq f(x), \qquad P\text{-a.e.}$$

then

$$\liminf \int_S f_n \, dP_n \geq \int_S f \, dP.$$

The theorems 1.2.20. and 1.2.22., which seem to be new, have been obtained independently by LANGEN (1981) with different proofs.

REMARK 1.2.24. From the proofs of the theorems 1.2.16., 1.2.20. and 1.2.22. follows easily that the conditions $h_n \overset{c}{\to} h$, PX^{-1}-a.e. (theorem 1.2.16.), $f_n \overset{c}{\to} f$ and $h_n \overset{c}{\to} h$, PX^{-1}-a.e. (theorem 1.2.20.) respectively $\liminf f_n(x_n) \geq f(x)$, PX^{-1}-a.e. (theorem 1.2.22.) can be relaxed in the sense that these conditions only have to hold for any $x\in C$ and any sequence $(x_n)_{n=1}^{\infty}$ with $\lim x_n = x$ and $x_n \in C_n$, where C, C_n, $n\geq 1$ are subsets of S such that $PX^{-1}(C)=1$ and $PX_n^{-1}(C_n)=1$, $n\geq 1$.

To conclude this section we derive an inequality concerning weak convergent probability measures.

NOTATION 1.2.25. For any set $B \in \mathcal{S}$ we denote by

B°: = the set of interior points of B.

δB: = the boundary of B.

$\bar{B}$: = the closure of B, i.e. $\bar{B} = B \cup \delta B = B^\circ \cup \delta B$.

THEOREM 1.2.26. Let P, P_n, $n \geq 1$, be probability measures on S and $A_n \in \mathcal{S}$, $n \geq 1$. If $P_n \xrightarrow{w} P$ then

(i) $$\limsup_{n\to\infty} P_n(A_n) \leq P\Big(\bigcap_{k=1}^{\infty} \overline{\bigcup_{m=k}^{\infty} A_m}\Big)$$

(ii) $$\liminf_{n\to\infty} P_n(A_n) \geq P\Big(\bigcup_{k=1}^{\infty} \Big(\bigcap_{m=k}^{\infty} A_m\Big)^\circ\Big).$$

PROOF. $P_n(A_n) \leq P_n(\bigcup_{m=k}^{\infty} A_m) \leq P_n(\overline{\bigcup_{m=k}^{\infty} A_m})$ for $n \geq k$. Hence

$$\limsup_{n\to\infty} P_n(A_n) \leq \limsup_{n\to\infty} P_n(\overline{\bigcup_{m=k}^{\infty} A_m}) \leq P(\overline{\bigcup_{m=k}^{\infty} A_m}) \text{ for all } k.$$

The last inequality follows from theorem 1.2.3.(iii).

Since $\lim_{k\to\infty} P(\overline{\bigcup_{m=k}^{\infty} A_m}) = P(\bigcap_{k=1}^{\infty} \overline{\bigcup_{m=k}^{\infty} A_m})$ we find (i) and the proof of (ii) is by complementation. □

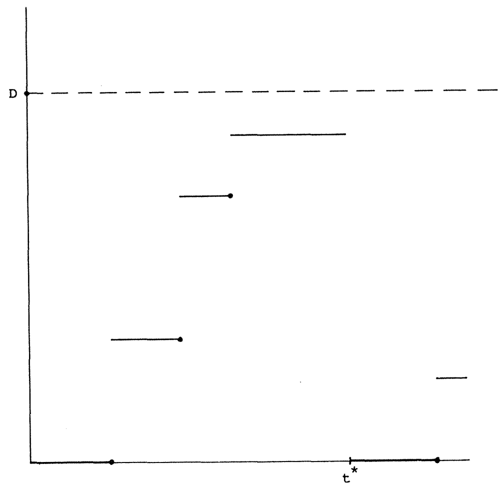
1.3. THE DRIFT FUNCTION.

In this section S is again a metric space with metric ρ and Borelfield $\mathcal{S}$. From now on we will often refer to S as the *state space*. In the wide field of the theory of stochastic processes several spaces of S-valued functions have been studied extensively. For example the space C[0,∞) consisting of all continuous functions defined on the non-negative half-line with values in S. A lot of stochastic processes (e.g. the Brownian motion) can be considered as random elements of C[0,∞). Another important function space is D[0,∞), the set of all S-valued functions defined on the non-negative half-line, which are right-continuous and have left hand limits at every t > 0. A great deal of the stochastic processes with non-continuous sample paths can be considered as random elements of D[0,∞).

However, in the theory of controlled stochastic processes there arise interesting stochastic processes which cannot be seen as random elements of either $C[0,\infty)$ or $D[0,\infty)$, especially because the sample paths are neither right- nor leftcontinuous. A space with sufficient generality for these processes is the collection of all S-valued functions on the positive half-line, which have left hand limits at every $t>0$ and right hand limits at every $t\geq 0$. However, in this monograph we have chosen an easier to handle but more restrictive space, not consisting of functions but of sequences of elements from the cartesian product of the non-negative half-line and S. Before we introduce this space in a formal way in section 1.4 we give some examples to illustrate the insufficiency of the spaces $C[0,\infty)$ and $D[0,\infty)$. These examples also show that these processes can be seen as random elements of a sequence space.

EXAMPLE 1.3.1. A certain device is subject to shocks which occur randomly in time according to a Poisson process. Every shock causes independently of the other ones a certain amount of damage. The damage accumulates additively. The amount of damage caused by a single shock is a random variable with known distribution function. There are operating costs for the device which in general will be an increasing function of the cumulative damage. The system can be controlled by replacing the device by a new one against specified costs. An easy to handle control-rule is the one which prescribes to replace the device as soon as the cumulative damage exceeds a threshold D. Consider the stochastic process describing at every epoch the cumulative damage of the device under operation, given that the above specified control rule is used. Figure 1.3.2. shows a typical sample path of this process, which is neither right- nor leftcontinuous at t^*.

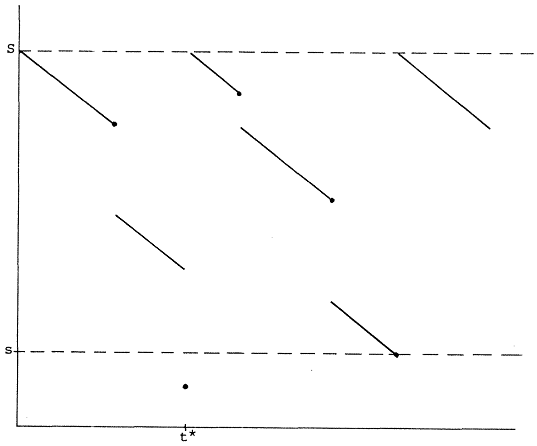
FIGURE 1.3.2.



Note that in this example the sample path is completely determined by the epochs at which a jump (shock or replacement) occurs and the cumulative damage immediately after the jump. This property is essential for all processes to be studied in this monograph and it is this property which leads us to the introduction of sequence spaces. In this rather simple example the state of the system is constant between jumps. This property, however, is not substantial as the following example shows.

EXAMPLE 1.3.3. The demand process at a warehouse where a certain commodity is stocked consists of two independent processes. There is a deterministic continuous demand at rate $\sigma>0$ per unit time and a stochastic demand, which is described by a compound Poisson process. So customers arrive according to a Poisson process and the demands of the customers are independent and identically distributed random variables with known distribution function. Unfilled demands will be backlogged and the warehouse management can place an order at any time which will be delivered without lead time. Suppose that the system is controlled by an (s,S)-rule i.e. as soon as the inventory level reaches or drops below s it is raised by ordering upto S. Consider the stochastic process describing at every epoch the inventory level (where negative values denote the amount to be backlogged) given that the (s,S) control rule is used. A typical sample path of this process is drawn in figure 1.3.4., where we have chosen $s>0$. Again this sample path is neither right- nor leftcontinuous at t^*.

FIGURE 1.3.4.



In this example, as in the previous one, the sample paths are determined by the epochs at which a jump (demand or delivery) occurs and the inventory level immediately after the jump.
In general we require the existence of one function f defined on the cartesian product of state space S and the non-negative half-line with values in S, to describe the behaviour of all possible sample paths between jumps under all possible control rules*, i.e. f(i,t) denotes the state of the system at time t+u given that the system is in state $i \in S$ at epoch u and no jump occurs between u and t+u, for all u.
In example 1.3.1. the appropriate function is given by

$$f(i,t) = i \qquad \text{for } i \in S,\ t \geq 0$$

and in example 1.3.3. by

$$f(i,t) = i-\sigma t, \text{ for } i\in S,\ t\geq 0.$$

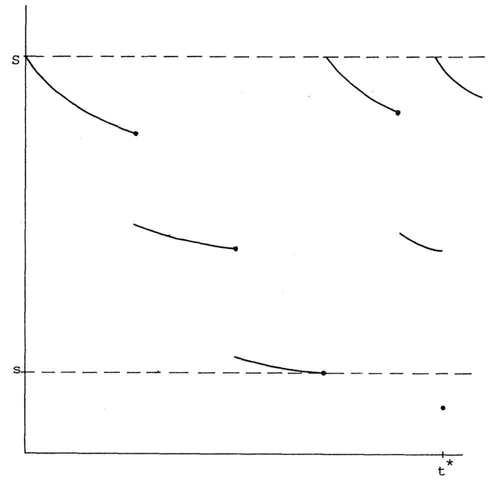
EXAMPLE 1.3.5. Consider again example 1.3.3.with the modification that the rate of the deterministic demand depends linearly on the physical inventory level, i.e. when the inventory level is $x>0$ then the deterministic demand equals σx per unit time. The appropriate function f is then defined by

* the formal definition of a control rule or policy will be given later on.

$$f(i,t) = \begin{cases} i & \text{for } i \le 0,\ t \ge 0 \\ ie^{-\sigma t} & \text{for } i > 0,\ t \ge 0. \end{cases}$$

Figure 1.3.6. shows a typical sample path.

FIGURE 1.3.6.



DEFINITION 1.3.7. A function $f\colon S\times[0,\infty)\to S$ is a *drift function for* S if:

(i) f is continuous on $S \times [0,\infty)$

(ii) $f(i,0) = i$ for all $i \in S$

(iii) $f(i,t) = f(f(i,u),\ t-u)$ for all $i \in S$, $0 \le u \le t$ and $t \ge 0$.

Note that property (iii) is in fact the deterministic version of the well known Markov property for stochastic processes.

In this monograph we restrict ourselves to processes with sample paths whose behaviour between jumps can be described by *one* drift function under all possible control rules. This excludes at first glance a number of interesting problems in the area of production-inventory control and controlled waiting line models.

EXAMPLE 1.3.8. Consider the M/G/1 queueing model as an inventory system with the inventory at time t being the virtual waiting time. The epochs at which customers arrive are generated by a Poisson process with rate λ. Any arriving customer enlarges the inventory of the system with a stochastic amount with known probability distribution. At any moment the controller can choose a service rate σ from a finite set $\{\sigma_1,\sigma_2,\ldots,\sigma_n\}$. As long as the inventory is positive and service rate σ is used, the inventory decreases between arrival epochs linearly at rate σ. Consider the stochastic process describing at every epoch the inventory level. Then it is obvious that there exists no drift function describing the behaviour of the sample paths of this process between jumps under *all* possible control rules. However, if we consider the stochastic process describing at every epoch the inventory level *and* the service rate used then the following drift function will do

$$f((i,\sigma),t) = \begin{cases} (i-\sigma t,\sigma) & \text{for } 0 \le t \le \frac{i}{\sigma},\ i \in S = [0,\infty) \\ (0,\sigma) & \text{for } t > \frac{i}{\sigma},\ i \in S. \end{cases}$$

When $S=[0,\infty)$ an important class of drift functions $f(i,.)$ can be obtained by solving the differential equation

$$z'(t) = r(z(t))$$
$$z(0) = i,$$

where $r(.)$ is a Lipschitz continuous function on $[0,\infty)$ with $r(0)=0$ (see also ÇINLAR and PINSKY (1971)). For example, $r(x)=-\sigma x$ yields $f(i,t)=ie^{-\sigma t}$.

1.4. THE SEQUENCE SPACES $J[0,t]$ AND $J[0,\infty)$.

The stochastic processes introduced in the preceding section can all be considered as random elements of sequence spaces. In this section we formally introduce and analyse some of these sequence spaces.
For any finite $t>0$ denote

$$(1.4.1) \qquad J[0,t]:= \{(t_j,i_j)_{j=1}^n : n \in \mathbb{N};\ 0 = t_1 \le t_2 \le \ldots \le t_n \le t;$$
$$t_{j+2} - t_j > 0,\ 1 \le j \le n-2;\ i_j \in S,\ 1 \le j \le n\}.$$

Let ζ be a "fictitious state" not belonging to S and denote $[0,\infty]:= [0,\infty) \cup \{\infty\}$ and $S^+:= S \cup \{\zeta\}$.

Next we define

$$(1.4.2) \qquad J[0,\infty):=\{(t_j,i_j)_{j=1}^{\infty}: 0=t_1\le t_2\le\ldots;\ t_j\in[0,\infty],\ i_j\in S^+,\ j\ge 1;$$

$$\lim_{j\to\infty} t_j=\infty;\ t_{j+2}-t_j>0 \text{ if } t_j<\infty;\ i_j=\zeta \text{ iff } t_j=\infty\}.$$

The space $J[0,t]$ consists of all finite sequences of elements from the cartesian product of the finite interval $[0,t]$ (time axis) and state space S, with the restriction that the sequence of time points is ordered and at most two subsequent time points can be equal. The reason for the introduction of the space $J[0,t]$ is twofold. First of all $J[0,t]$ is important in its own right, since a lot of stochastic processes with finite time horizon t can be seen as random elements of $J[0,t]$. On the other hand $J[0,t]$ will be used as auxiliary space in analyzing $J[0,\infty)$. The space $J[0,\infty)$ contains all infinite sequences of elements from the cartesian product of the non-negative half line included ∞ and the state space included ζ. The sample paths drawn in the figures 1.3.2., 1.3.4. and 1.3.6. are representable as elements of $J[0,\infty)$ by their sequence of jump epochs and states immediately after the jumps. The elements ∞ and ζ are introduced to enable us to represent sample paths with a finite number of jumps as infinite sequences. Sample paths with a finite number of jumps occur for example in processes for which there is a positive probability to stay for an infinite long period in a single state.
On the other hand the requirement that no more than two subsequent time points can be equal opens the possibility to represent for a given drift function these sequences as S-valued functions on the time axis. The spaces $J[0,t]$ and $J[0,\infty)$ will now be endowed with metrics, such that they become complete and separable. Completeness and separability facilitate the characterization of weak convergence. First we analyse $J[0,t]$ for fixed $t>0$. For the duration of this analysis we agree upon the following notation:

$$x:=(t_j,i_j)_{j=1}^{n};$$

$$y:=(s_j,h_j)_{j=1}^{m};$$

and

$$x(k):=(t_j(k),i_j(k))_{j=1}^{n(k)},\ k\ge 1$$

which are assumed to be elements of $J[0,t]$.

Furthermore we use the notation

$$w \vee z := \max(w,z)$$

and

$$w \wedge z := \min(w,z)$$

for any two real numbers w and z.
We define

$$(1.4.3) \qquad d_t(x,y) := |n-m| \vee \{1 \wedge \max_{j\le n,m} \{|t_j-s_j| \vee \rho(i_j,h_j)\}\}$$

PROPOSITION 1.4.1. d_t defines a metric on $J[0,t]$.

PROOF. The proof is straightforward and will be omitted. □

However, the metric space $(J[0,t],d_t)$ is not complete. For instance, put

$$(1.4.4) \qquad x_n := (0,\ i,\ s-\tfrac{1}{n},\ j,\ s,\ i,\ s,\ j)$$

for some $s \in (0,t)$ and $i,j \in S$ with $i \neq j$.
Then $d_t(x_n,x_m) = |\frac{1}{n} - \frac{1}{m}|$, which implies that $(x_n)_{n=1}^{\infty}$ is d_t-fundamental. However, $(x_n)_{n=1}^{\infty}$ does not converge in $J[0,t]$. In order to make $J[0,t]$ a complete metric space we need another, more complicated metric.
Define

$$(1.4.5) \qquad \tilde{d}_t(x,y) := |n-m| \vee \{1 \wedge \max_{j\le n,m} \{|t_j-s_j| \vee \rho(i_j,h_j) \vee \vee \left|\log \frac{t_j-t_{j-2}}{s_j-s_{j-2}}\right|\}\}$$

where $t_{-1} = s_{-1} = t_0 = s_0 = -1$.

Note that for the sequence $(x_n)_{n=1}^{\infty}$ defined by (1.4.4) above $\tilde{d}_t(x_n,x_{n^2}) = \log n$ for n sufficiently large. Hence $(x_n)_{n=1}^{\infty}$ is not $\tilde{d}_t$-fundamental.

PROPOSITION 1.4.2. (i) $\tilde{d}_t$ defines a metric on $J[0,t]$
(ii) $\tilde{d}_t(x,y) \geq d_t(x,y)$ for all $x,y \in J[0,t]$.

PROOF. The proof follows immediately from the definitions of d_t and $\tilde{d}_t$. □

LEMMA 1.4.3. If S is complete then $J[0,t]$ is complete w.r.t. $\tilde{d}_t$.

PROOF. Suppose S is complete and $(x(k))_{k=1}^{\infty}$ is $\tilde{d}_t$-fundamental in $J[0,t]$. Then $(n(k))_{k=1}^{\infty}$ converges as $k \to \infty$. Hence there exist numbers K and n such that $n(k) = n$ for all $k > K$. For $j = 1,\ldots,n$ the sequences $(t_j(k))_{k=1}^{\infty}$ and $(i_j(k))_{k=1}^{\infty}$ are fundamental in $[0,t]$ and S respectively and therefore convergent. Put $t_j^*: = \lim_{k\to\infty} t_j(k)$ and $i_j^*: = \lim_{k\to\infty} i_j(k)$, for $1 \le j \le n$. Since the sequence $(\log(t_{j+2}(k) - t_j(k)))_{k=1}^{\infty}$ is fundamental in $\mathbb{R}$ it follows that $t_{j+2}^* - t_j^* > 0$ for $j=1,\ldots,n-2$. This implies that $x^*: = (t_j^*,i_j^*)_{j=1}^{n}$ is an element of $J[0,t]$. Finally it is straightforward to prove that $\lim_{k\to\infty} \tilde{d}_t\ (x(k),x^*) = 0$. □

Although $J[0,t]$ is not complete under the simpler metric d_t, we can use d_t as well, as far as topological properties are concerned. This is justified by lemma 1.4.5. to follow.

DEFINITION 1.4.4. Let (S,ρ) be a metric space. For $i \in S$ the *ε-sphere about i* is defined by

$$S_\rho(i,\varepsilon): = \{j \in S: \rho(i,j) < \varepsilon\}.$$

LEMMA 1.4.5. The metrics d_t and $\tilde{d}_t$ are equivalent.

PROOF. From proposition 1.4.2. (ii) we know that $S_{\tilde{d}_t}(x,\varepsilon) \subset S_{d_t}(x,\varepsilon)$ for all $x \in J[0,t]$ and therefore it is sufficient to prove that for every $x \in J[0,t]$ and $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$S_{d_t}(x,\delta) \subset S_{\tilde{d}_t}(x,\varepsilon).$$

Choose $x=(t_j,i_j)_{j=1}^{n} \in J[0,t]$ and $\varepsilon>0$. Without loss of generality we assume that $\varepsilon<1$ and $\varepsilon < t_{j+2}-t_j$, $1\le j\le n-2$. Put $\delta=\frac{1}{4}\varepsilon^2$ and take $y=(s_j,h_j)_{j=1}^{m} \in S_{d_t}(x,\delta)$. Then $d_t(x,y)<\delta<1$, which implies $n=m$. Moreover, for $j=1,\ldots,n-2$

$$|t_{j+2} - t_j - s_{j+2} + s_j| \le 2\delta = \tfrac{1}{2}\varepsilon^2 < \tfrac{1}{2}\varepsilon\ |t_{j+2} - t_j|.$$

Hence we have for $j = 1,\ldots,n-2$

$$-\varepsilon < -\tfrac{1}{2}\varepsilon - \tfrac{1}{4}\varepsilon^2 < \log(1-\tfrac{1}{2}\varepsilon) < \log \frac{s_{j+2}-s_j}{t_{j+2}-t_j} < \log(1+\tfrac{1}{2}\varepsilon) < \tfrac{1}{2}\varepsilon < \varepsilon.$$

This implies $y \in S_{\tilde{d}_t}(x,\varepsilon)$ and hence $S_{d_t}(x,\delta) \subset S_{\tilde{d}_t}(x,\varepsilon)$. □

This concludes the analysis of the space $J[0,t]$.

Next we consider the space $J[0,\infty)$ and agree for the rest of this section upon the notation:

$$x: = (t_j,i_j)_{j=1}^{\infty};$$

$$y: = (s_j,h_j)_{j=1}^{\infty}$$

and

$$x(k): = (t_j(k),i_j(k))_{j=1}^{\infty}, \qquad k \geq 1$$

which are assumed to be elements of $J[0,\infty)$.

<u>DEFINITION 1.4.6.</u> (i) For $t \geq 0$ the *t-restriction* is a function r_t: $J[0,\infty) \to J[0,t]$ defined by

$$r_t x = \begin{cases} (t_j,i_j)_{j=1}^{n-1}, & \text{if } t=t_n \\ (t_j,i_j)_{j=1}^{n}, & \text{otherwise} \end{cases}, \text{ where } n:=\sup\{j:t_j \leq t\}.$$

(ii) For $t \geq 0$ the *t-extension* is a function e_t: $J[0,t] \to J[0,\infty)$ defined by

$$e_t((t_j,i_j)_{j=1}^{n}) = (t_1,i_1,\ldots,t_n,i_n,\infty,\zeta,\ldots)$$

On $J[0,\infty)$ we define the metrics

$$(1.4.6) \qquad d(x,y): = \int_0^{\infty} e^{-t} (1 \wedge d_t(r_t x, r_t y))dt$$

and

$$(1.4.7) \qquad \tilde{d}(x,y): = \int_0^{\infty} e^{-t} (1 \wedge \tilde{d}_t(r_t x, r_t y))dt.$$

<u>PROPOSITION 1.4.7.</u> d and $\tilde{d}$ are well-defined metrics on $J[0,\infty)$.

PROOF. For $x=(t_j,i_j)_{j=1}^{\infty}$ and $y=(s_j,h_j)_{j=1}^{\infty}$ fixed $d_t(r_tx,r_ty)$ and $\tilde{d}_t(r_tx,r_ty)$ are piecewise constant functions of t and therefore measurable. This implies that the integrals are well-defined. The proof of the symmetry of d and $\tilde{d}$ and the triangle inequality are easily checked. From proposition 1.4.2.(ii) and the definition of d and $\tilde{d}$ it follows that $\tilde{d}(x,y)\geq d(x,y)$, which implies that $\tilde{d}$ separates if d separates. So it remains to prove that d separates. Suppose $d(x,y)=0$ and $(t_j,i_j)\neq(s_j,h_j)$ for some j. We consider three cases:

(i) $i_j \neq h_j$ and $t_j < \infty$ or $s_j < \infty$. Then $d_t(r_tx, r_ty) > 0$ for all $t > (t_j \wedge s_j) < \infty$. This contradicts $d(x,y) = 0$.

(ii) $i_j \neq h_j$ and $t_j = s_j = \infty$. This contradicts $x,y \in J[0,\infty)$.

(iii) $i_j = h_j$. Then $d_t(r_tx, r_ty) > 0$ for all $(t_j \wedge s_j) < t < (t_j \vee s_j)$. This again contradicts $d(x,y) = 0$. □

The next theorem connects the convergence of sequences in $J[0,\infty)$ with that in $J[0,t]$.

THEOREM 1.4.8. $\lim_{k\to\infty} d(x(k), x) = 0$ iff $\lim_{k\to\infty} d_t(r_tx(k), r_tx) = 0$ for all $t \neq t_j$, $j \geq 1$.

PROOF. *Necessity*. Choose $t \in (t_j,t_{j+1})$ for some j. Since $\lim_{k\to\infty}d(x(k),x)=0$ it follows that $\lim_{k\to\infty}t_j(k) = t_j$ and $\lim_{k\to\infty}t_{j+1}(k) = t_{j+1}$. Choose $\varepsilon > 0$ such that $t \in (t_j+\varepsilon, t_{j+1}-\varepsilon)$. Then there exists a number K such that for all $k > K$ the function $d_s(r_sx(k), r_sx)$ is constant on $(t_j+\varepsilon, t_{j+1}-\varepsilon)$ as a function of s. Hence it follows from

$$d(x(k), x) \geq \int_{t_j+\varepsilon}^{t_{j+1}-\varepsilon} e^{-s} (1 \wedge d_s(r_sx(k), r_sx))ds \qquad \text{for all } k > K,$$

that $\lim_{k\to\infty} d_t(r_tx(k), r_tx) = 0$.

Sufficiency. An immediate consequence of Lebesgue's theorem on bounded convergence. □

Note that $J[0,\infty)$ is not complete w.r.t. d. For instance put $x_n:=(0, i, s-\frac{1}{n}, j, s, i, s, j, \infty, \zeta,\ldots)$ for $s > 0$ and $i,j \in S$ with $i \neq j$. Then

$$d(x_n,x_m) = \int_{s-\frac{1}{n}}^{s-\frac{1}{m}} e^{-s}\, ds + \int_{s-\frac{1}{m}}^{\infty} e^{-s}\, \left|\frac{1}{n} - \frac{1}{m}\right| ds \qquad \text{for } n < m.$$

It follows that $d(x_n, x_m) \le 2|\frac{1}{n} - \frac{1}{m}|$, which implies that $(x_n)_{n=1}^{\infty}$ is d-fundamental. However, $(x_n)_{n=1}^{\infty}$ does not converge in $J[0,\infty)$.

<u>THEOREM 1.4.9</u>. If S is complete then $J[0,\infty)$ is complete w.r.t. $\tilde{d}$.

<u>PROOF</u>. Suppose S is complete and $(x(k))_{k=1}^{\infty}$ is $\tilde{d}$-fundamental in $J[0,\infty)$. Consider for fixed $j \ge 1$ the sequence $(t_j(k))_{k=1}^{\infty}$. We consider two cases:

(i) $(t_j(k))_{k=1}^{\infty}$ is bounded. Choose $0 < \varepsilon < 1$ and $t > 0$ such that $t_j(k) \le t$ for all k. Put $\varepsilon' := \varepsilon e^{-t}$. Since $(x(k))_{k=1}^{\infty}$ is $\tilde{d}$-fundamental there exists a number K such that for all $k > K$ and all ℓ we have

$$\tilde{d}(x_{k+\ell}, x_k) < \varepsilon'.$$

From this inequality follows the existence of a real number $s > t$ (s may depend on k and ℓ) such that

$$\tilde{d}_s(r_s x_{k+\ell}, r_s x_k) < \varepsilon.$$

Since $s > t$ and $t_j(k) \le t$, $k \ge 1$ it follows that $|t_j(k+\ell) - t_j(k)| < \varepsilon$ and $\rho(i_j(k+\ell), i_j(k)) < \varepsilon$. So $(t_j(k))_{k=1}^{\infty}$ and $(i_j(k))_{k=1}^{\infty}$ are fundamental in $[0,\infty)$ and S respectively and hence convergent. Put $t_j^* := \lim_{k\to\infty} t_j(k)$ and $i_j^* := \lim_{k\to\infty} i_j(k)$.

(ii) In case that $(t_j(k))_{k=1}^{\infty}$ is unbounded we put $t_j^* := \infty$ and $i_j^* := \zeta$.

Now we define $x^* := (t_j^*, i_j^*)_{j=1}^{\infty}$. Then it is straightforward to prove that $x^* \in J[0,\infty)$. To show for example that $\lim_{j\to\infty} t_j^* = \infty$ suppose $\lim_{j\to\infty} t_j^* = t < \infty$ and put $\varepsilon = \int_{t+1}^{t+2} e^{-s}\,ds$. Choose $K > 0$ and $k > K$. Since $x(k) \in J[0,\infty)$ it follows that $\lim_{j\to\infty} t_j(k) = \infty$ which implies the existence of a number j such that $t_j(k) > t+2$. However, since $\lim_{k\to\infty} t_j(k) = t_j^*$ and $\lim_{j\to\infty} t_j^* = t$ we can also find a number $\ell > K$ such that $t_j(\ell) < t+1$. This implies that $\tilde{d}(x_k, x_\ell) > \int_{t+1}^{t+2} e^{-s}\,ds = \varepsilon$ which is in contradiction with the fact that $(x_n)_{n=1}^{\infty}$ is $\tilde{d}$-fundamental. Finally the proof that $\lim_{k\to\infty} \tilde{d}(x(k), x^*) = 0$ is straightforward. □

Since the metric d is easier to handle than the metric $\tilde{d}$ it is useful to know that these metrics are equivalent. To prove this we need the following lemma.

LEMMA 1.4.10. Consider the metric spaces (S,ρ_1) and (S,ρ_2). The following two statements are equivalent:

(i) ρ_1 and ρ_2 are equivalent on S.

(ii) $\lim_{n\to\infty} \rho_1(s_n, s) = 0$ iff $\lim_{n\to\infty} \rho_2(s_n, s) = 0$ for all $s, s_n \in S$.

PROOF. The implication (i) $\Rightarrow$ (ii) follows immediately from the definition of equivalent metrics.
Assume that (ii) holds and that ρ_1 and ρ_2 are not equivalent, i.e. there exists an open set O in (S, ρ_1), which is not open in (S, ρ_2). Hence there is a point $s \in O$ which is not an interior point of O in (S, ρ_2). This implies that the n^{-1}-sphere $S_{\rho_2}(s, n^{-1})$ contains a point $s_n \notin O$ for all $n \geq 1$. Hence there exists a sequence $(s_n)_{n=1}^{\infty}$ such that $\lim_{n\to\infty}\rho_2(s_n, s) = 0$ while $(s_n)_{n=1}^{\infty}$ does not converge to s in (S, ρ_1), because $s_n \notin O$, $s \in O$ and O is open in (S, ρ_1). Contradiction. □

THEOREM 1.4.11. The metrics d and $\tilde{d}$ are equivalent on $J[0,\infty)$.

PROOF. From lemma 1.4.10. follows that it is sufficient to show that $\lim_{k\to\infty} d(x(k),x) = 0$ iff $\lim_{k\to\infty} \tilde{d}(x(k),x) = 0$.
From theorem 1.4.8 . follows that $\lim_{k\to\infty} d(x(k),x) = 0$ iff $\lim_{k\to\infty} d_t(r_t x(k), r_t x) = 0$ for all $t \neq t_j$, $j \geq 1$. Combining this with lemmas 1.4.5. and 1.4.10. yields $\lim_{k\to\infty} d(x(k), x) = 0$ iff $\lim_{k\to\infty} \tilde{d}_t(r_t x(k), r_t x) = 0$ for all $t \neq t_j$, $j \geq 1$. Applying theorem 1.4.8. with respect to $\tilde{d}$ we get $\lim_{k\to\infty} d(x(k), x) = 0$ iff $\lim_{k\to\infty} \tilde{d}(x(k), x) = 0$. □

THEOREM 1.4.12. If S is separable then $J[0,\infty)$ is separable.

PROOF. Let W be a countable dense subset of S. Define

$$Z := \{(t_j,i_j)_{j=1}^{\infty} \in J[0,\infty) : t_k = \infty \text{ for some } k;\ i_j \in W,\ t_j \text{ is rational, } j=1,\dots,k-1\}.$$

Then Z is a countable subset of $J[0,\infty)$. To prove that Z is dense in $J[0,\infty)$ choose $x=(t_j,i_j)_{j=1}^{\infty} \in J[0,\infty)$ and define

$$t_j^*(k) = \begin{cases} \dfrac{[t_j k^2]}{k^2} & \text{if } t_j<k \text{ and } j\leq k \\ \infty & \text{else} \end{cases}$$

where $[w]$ denotes the *entier of* w for any real number w. Choose

$\overset{*}{i}_j(k) \in W \cup \{\zeta\}$ such that $\overset{*}{i}_j(k) = \zeta$ iff $\overset{*}{t}_j(k) = \infty$ and $\rho(\overset{*}{i}_j(k), i_j) < k^{-1}$ if $\overset{*}{t}_j(k) < \infty$.
Put

$$\overset{*}{x}(k) := (\overset{*}{t}_j(k), \overset{*}{i}_j(k))_{j=1}^{\infty}, \; k \geq 1.$$

Then $\overset{*}{x}(k) \in Z$, $k \geq 1$ and

$$d(\overset{*}{x}(k), x) \leq \exp(-(k \wedge t_k)) + \int_0^{(k \wedge t_k)} e^{-s} d_s(r_s x(k), r_s x) ds \leq$$

$$\leq \exp(-(k \wedge t_k)) + \int_0^k e^{-s} k^{-1} ds + \sum_{j=1}^{k} \int_{\overset{*}{t}_j(k)}^{t_j} e^{-s} ds \leq$$

$$\leq \exp(-(k \wedge t_k)) + k^{-1} + k.k^{-2}.$$

Hence $\lim_{k\to\infty} d(\overset{*}{x}(k), x) = 0$. □

From now on we assume that S is separable and complete, which implies, according to theorems 1.4.9 and 1.4.12. that $J[0,\infty)$ is separable and complete.
As mentioned before the definitions of $J[0,t]$ and $J[0,\infty)$ enable us to represent the elements of these spaces as S-valued functions on the time axis provided a drift function has been given. There are situations in which it is more convenient to use this representation in stead of the sequence itself. We define this function representation as follows.

<u>DEFINITION 1.4.13</u>. Let f be a drift function for S. For $t \geq 0$ the *t-projection* is a S-valued function π_t on $J[0,\infty)$ defined for $x=(t_j,i_j)_{j=1}^{\infty}$ by

$$\pi_t x = \begin{cases} i_1 & \text{for } t = 0 \\ f(i_j, t-t_j) & \text{for } t \in (t_j, t_{j+1}) \\ f(i_j, t_{j+1}-t_j) & \text{for } t = t_{j+1} < t_{j+2} \\ i_{j+1} & \text{for } t = t_{j+1} = t_{j+2} < \infty. \end{cases}$$

REMARK 1.4.14.

(i) Note that $\pi_t x$ as a function of t for fixed x is left continuous except in the time points t for which there exists a number j such that $t = t_j = t_{j+1}$.

(ii) From the definition of $\pi_t x$ follows easily that for fixed x the function $\pi_t x$ as function of t is measurable.

This section is closed with a lemma concerning the functions r_t, e_t and π_t introduced before (see definitions 1.4.6. and 1.4.13).
These functions will be repeatedly used in the sequel and therefore it is worthwhile to know that they have certain regularity properties.
Note that for π_t to be well-defined we have to specify a drift function.
This is not necessary for r_t and e_t.

LEMMA 1.4.15. Let f be a drift function for S and choose $t \geq 0$. Then

(i) e_t is continuous on $J[0,t]$

(ii) r_t is measurable on $J[0,\infty)$

(iii) π_t is measurable on $J[0,\infty)$.

PROOF.

(i) Let $x_0 := (t_j,i_j)_{j=1}^m \in J[0,t]$ and $x_n \in J[0,t]$, $n \geq 1$ such that $\lim_{n\to\infty} d_t(x_n,x_0) = 0$. In the same way as the necessary part of theorem 1.4.8. one proves that $\lim_{n\to\infty} d_s(r_s e_t x_n, r_s e_t x_0) = 0$ for all $t \geq s \neq t_j$, $j = 1,\dots,m$. Moreover, for $s > t$ we have $d_s(r_s e_t x_n, r_s e_t x_0) = d_t(x_n,x_0)$. Hence $\lim_{n\to\infty} d_s(r_s e_t x_n, r_s e_t x_0) = 0$ for all $s \neq t_j$, $j = 1,\dots,m$, which implies that $\lim_{n\to\infty} d(e_t x_n, e_t x_0) = 0$ by the sufficient part of theorem 1.4.8.

(ii) Put

$B := \{x \in J[0,\infty)$: there is a number j such that $t = t_j = t_{j+1}\}$

$C := \{x \in J[0,\infty)$: there is a number j such that $t_{j-1} < t = t_j < t_{j+1}\}$

$D := \{x \in J[0,\infty)$: there is a number j such that $t_j < t < t_{j+1}\}$.

The set B is closed in $J[0,\infty)$ and D is open in $J[0,\infty)$. Since $C = (B \cup D)^c$ the sets B, C and D are all measurable. Let A be a closed subset of $J[0,t]$. Then $r_t^{-1}(A) = (r_t^{-1}(A) \cap B) \cup (r_t^{-1}(A) \cap C) \cup (r_t^{-1}(A) \cap D)$. Since $r_t^{-1}(A) \cap B$ is closed in $J[0,\infty)$, $r_t^{-1}(A) \cap C$ is closed in C and $r_t^{-1}(A) \cap D$ is closed in D all these sets are measurable in $J[0,\infty)$. Hence $r_t^{-1}(A)$ is measurable.

(iii) The proof of (iii) proceeds quite similar to (ii). □

1.5. COMPARING $J[0,\infty)$ WITH RELATED SPACES.

In this section we compare the space $J[0,\infty)$ with some related spaces. In the first place the resemblance of $J[0,\infty)$ to the space $D[0,\infty]$, introduced in section 1.1, forces itself upon us. In some sense $J[0,\infty)$ is more general than $D[0,\infty)$, because for an element x of $J[0,\infty)$ the function $\pi_t(x)$ may be neither right- nor leftcontinuous as a function of t, while this function is right continuous for the elements of $D[0,\infty)$. On the other other hand $D[0,\infty)$ is more general than $J[0,\infty)$, because the elements of $J[0,\infty)$ can be represented as functions on $[0,\infty)$ only with use of a drift function, which implies that the behaviour between jumps for all elements of $J[0,\infty)$ is the same.
In chapter 2 it will turn out to be useful to consider $J[0,\infty)$ as a subset of the infinite product space $([0,\infty)\times S\cup\{(\infty,\zeta)\})^{\infty}$.
The question arises whether the set $[0,\infty)\times S\cup\{(\infty,\zeta)\}$ can be endowed with a metric $\tilde{\rho}$ such that on $J[0,\infty)$ the infinite product metric on $([0,\infty)\times S\cup\{(\infty,\zeta)\})^{\infty}$ is equivalent to the metric d defined by (1.4.6). The answer is positive as is shown by theorem 1.5.1. below. Put for abbreviation

$$\overline{[0,\infty)\times S} := [0,\infty)\times S\cup\{(\infty,\zeta)\}$$

and define for $(t_j,i_j)\in\overline{[0,\infty)\times S}$, j=1,2

$$\tilde{\rho}((t_1,i_1),(t_2,i_2)) := |e^{-t_1}-e^{-t_2}|+\min(e^{-t_1},e^{-t_2})\{\rho^+(i_1,i_2)\wedge 1\} \tag{1.5.1}$$

where ρ^+ is a metric on S^+ defined by

$$(1.5.2)\qquad \rho^+(i,j) := \begin{cases} \rho(i,j)\wedge 1 & \text{for } i,j\in S \\ 1 & \text{for } i\in S,\ j=\zeta \text{ or } j\in S,\ i=\zeta \\ 0 & \text{for } i=j=\zeta. \end{cases}$$

Let $x=(t_j,i_j)_{j=1}^{\infty}$ and $y:=(s_j,h_j)_{j=1}^{\infty}$ be elements of $(\overline{[0,\infty)\times S})^{\infty}$ and define

$$(1.5.3)\qquad d^{\infty}(x,y) := \sum_{j=1}^{\infty} 2^{-j}\tilde{\rho}((t_j,i_j),(s_j,h_j)).$$

<u>THEOREM 1.5.1</u>. The metrics d and d^{∞} defined by (1.4.6) and (1.5.3) are equivalent on $J[0,\infty)$.

<u>PROOF</u>. It is sufficient to show that $\lim_{k\to\infty} d^{\infty}(x(k),x)=0$ iff $\lim_{k\to\infty} d(x(k),x)=0$ for all x, $x(k)\in J[0,\infty)$. Let $x=(t_j,i_j)_{j=1}^{\infty}$ and $x(k)=(t_j(k),i_j(k))_{j=1}^{\infty}$, $k\geq 1$ be elements of $J[0,\infty)$ and suppose that $\lim_{k\to\infty} d^{\infty}(x(k),x)=0$. Choose $\varepsilon>0$ and $t>0$ such that $e^{-t}<\frac{1}{2}\varepsilon$. Choose $n\geq 1$ such that $t_n\leq t$ and $t_{n+1}>t$. Then there exists a number $K>0$ such that for all $k\geq K$

$$\tilde{\rho}((t_j(k),i_j(k)),(t_j,i_j))<(1-e^{-\frac{\varepsilon}{4n}})e^{-t-\varepsilon},\qquad j=1,\dots,n$$

$$t_{n+1}(k)>t$$

and

$$t_n(k)\leq t+\varepsilon.$$

From the definition of $\tilde{\rho}$ follows that $|t_j(k)-t_j|<\frac{1}{4}\varepsilon n^{-1}$ and $\rho(i_j(k),i_j)<\frac{1}{4}\varepsilon n^{-1}$ for $k\geq K$ and $j=1,\dots,n$.
Hence we have for $k\geq K$,

$$d(x(k),x)\leq e^{-t}+n.\tfrac{1}{4}\varepsilon n^{-1} + \int_0^t e^{-s}.\tfrac{1}{4}\varepsilon ds<\varepsilon.$$

To prove the other implication suppose that $\lim_{k\to\infty} d(x(k),x)=0$. Choose $\varepsilon>0$ and $t\in(t_n,t_{n+1})$ for some n, such that $e^{-t}<\frac{1}{2}\varepsilon$. From theorem 1.4.8. follows that $\lim_{k\to\infty} d_t(r_t x(k),r_t x)=0$. Hence there exists a number $K>0$ such that $d_t(r_t x(k),r_t(x)<\frac{1}{2}\varepsilon$ and $t_{n+1}(k)>t$ for all $k\geq K$. Hence $\tilde{\rho}((t_j(k),i_j(k)),(t_j,i_j))<\varepsilon$ for all $k\geq K$ and all j which implies $d^{\infty}(x(k),x)<\varepsilon$ for $k\geq K$. □

1.6. WEAK CONVERGENCE ON $J[0,\infty)$.

According to theorem 1.2.3. a sufficient condition for the weak convergence of a sequence of probability measures $(P_n)_{n=1}^{\infty}$ to P on the metric space S is $\lim_{n\to\infty} P_n(B)=P(B)$ for all $B\in\mathcal{S}$ with $P(\delta B)=0$. For $J[0,\infty)$, however, this condition is not easy to verify. Hence we are interested in a relaxation of this condition in the sense that the limiting relation only holds for a subclass of the Borelsets of $J[0,\infty)$. To describe an appropriate subclass we need the following functions defined on $J[0,\infty)$ with values in $[0,\infty]$ and S^+ respectively.

$$T_n((t_j,i_j)_{j=1}^{\infty}) := t_n$$

$$S_n((t_j,i_j))_{j=1}^{\infty} := i_n.$$

Note that (T_n,S_n) as a function on $J[0,\infty)$ with values in $(\overline{[0,\infty)\times S},\tilde{\rho})$ is continuous for all $n\geq 1$.

<u>THEOREM 1.6.1</u>. Let P, P_k, $k\geq 1$ be probability measures on $J[0,\infty)$ and define

$$\mathcal{F} := \{ \bigcap_{n=1}^{m} (T_n^{-1}(B_n)\cap S_n^{-1}(A_n)) : m\geq 1;\ A_n\in\mathcal{S},\ B_n \text{ finite interval}, \tag{1.6.1}$$

$$1\leq n\leq m-1;\ A_m\in\mathcal{S},\ B_m \text{ finite interval or}$$

$$A_m=S^+,\ B_m \text{ infinite interval with } \infty\in B_m\}.$$

If $\lim_{k\to\infty} P_k(F)=P(F)$ for all $F\in\mathcal{F}$ with $P(\delta F)=0$ then $P_k \overset{w}{\to} P$.

<u>PROOF</u>. Put $\mathcal{F}_0:=\{F\in\mathcal{F}: P(\delta F)=0\}$. Since $\delta(A\cap B)\subset\delta A\cup\delta B$ it follows that $\mathcal{F}_0$ is closed under finite intersections. Choose $x\in J[0,\infty)$ and $\varepsilon>0$. Since $J[0,\infty)$ is separable, it is according to corollary 1 on page 14 of BILLINGSLEY (1968) sufficient to prove the existence of a set $F\in\mathcal{F}_0$ such that $x\in F^{\circ}\subset F\subset S_d(x,\varepsilon)$. Choose $t>0$ such that $e^{-t}<\frac{1}{3}\varepsilon$. Consider two cases separately.

(i) $T_2(x)=\infty$. Then we put

$$F:=\{y\in J[0,\infty):\rho(S_1(x),S_1(y))<\tfrac{1}{3}\varepsilon,\ T_2(y)\in(t,\infty],S_2(y)\in S^+\}.$$

(ii) $T_2(x)<\infty$. Choose m such that $T_m(x)\le t$ and $T_{m+1}(x)>t$ and put

$$F:=\bigcap_{n=1}^{m+1} T_n^{-1}(B_n)\cap S_n^{-1}(A_n),$$

where $A_n=\{j\in S:\ \rho(S_n(x),j)<\frac{1}{3}\varepsilon m^{-1}\}$, $B_n=(T_n(x)-\frac{1}{3}\varepsilon m^{-1},T_n(x)+\frac{1}{3}\varepsilon m^{-1})$, $1\le n\le m$;

$$A_{m+1}=S^+ \text{ and } B_{m+1}=[t-\tfrac{1}{3}\varepsilon m^{-1},\infty].$$

In both cases $x\in F^\circ$ and $d(x,y)<\frac{1}{3}\varepsilon+m\,\frac{\varepsilon}{3m}+\int_0^t e^{-s}.\frac{\varepsilon}{3}\,ds<\varepsilon$ for all $y\in F$. From the observation that

$$\delta F\subset\bigcup_{n=1}^{m+1}(T_n^{-1}(\delta B_n)\cup S_n^{-1}(\delta A_n))$$

we easily conclude that the set F, which obviously belongs to $\mathcal{F}$, also belongs to $\mathcal{F}_0$. □

CHAPTER 2

MARKOV DECISION DRIFT PROCESSES ON $J[0,\infty)$

2.1. INTRODUCTION.

Markov decision processes with continuous time parameter have been introduced by BELLMAN (1957) in chapter 11 of his book. HOWARD (1960) considers in chapter 8 of his book also these kind of processes, but he emphasizes the infinite horizon case, where BELLMAN is concerned exclusively with the finite horizon model. During the last two decades a large number of papers have appeared on this subject. MILLER (1968, 1968a), KAKUMANU (1971, 1975), DOSHI (1974, 1976) and PLISKA (1975) treated continuous time Markov decision processes under increasing generality of state and action spaces. A partial overview of the relevant literature will be given in the last chapter of this monograph.

The continuous time Markov decision drift processes (CTMDP) to be studied in this monograph are generalisations of Markov decision processes with continuous time parameter in two different aspects.
In the first place a CTMDP permits both control of the infinitesimal generator of the process as well as impulsive controls. Impulsive controls are not allowed in the models studied by the authors mentioned above. The difference between these two kinds of control can at best be explained by an example. Consider the inventory model of example 1.3.3. of chapter 1. In this inventory system the decision maker may have control on the arrival rate of customers by advertising or by levying toll. In this situation at every instant a rate for the Poisson arrival process of customers has to be chosen by fixing an advertising- or toll level. This kind of control affects the state of the system (the economic inventory) only in an indirect way. What is controlled is the infinitesimal generator of the inventory process. On the other hand the decision maker may have the opportunity to

place an order at any desired epoch. By such a control the economic inventory changes immediately.
In all papers mentioned above the absence of this second type of control is assumed, by which a number of interesting applications (like inventory-, production- and replacement models) is excluded.

We will call the decisions which cause an immediate change of state *impulsive controls*. The decisions which only affect the infinitesimal generator are called *controls*. A control can affect the system only if it is chosen during a time interval of positive length. Note that the difference between controls and impulsive controls is not meaningful for discrete time processes. Since these processes are considered only on equidistant decision epochs immediate changes of state do not occur.

The main reason for a model which contains both generator controls and impulsive controls is not the abundance of models in which both types of controls simultaneously occur. However, it is useful to have a unifying approach to two distinct classes of important continuous time decision models. In the area of control theory several results have been obtained on decision processes with both generator- and impulsive control. For example the work of KUSHNER (1977) and ROBIN (1978) contains valuable contributions to this field. Our approach, however, is quite different.

A second aspect of generalisation of a CTMDP compared with a continuous time Markov decision process is the behaviour of the process between two successive jump epochs. In a Markov decision process with continuous time parameter it is assumed that the state of the system is constant between jumps. In a CTMDP, however, we assume the existence of a deterministic drift function, according to which the system evolves between jump epochs, independent of the chosen policy.

In most of the literature concerning Markov decision processes with continuous time parameter the policies are restricted to Markov policies, i.e. those policies which prescribe a decision, which may depend only on the current epoch and the state of the system at this epoch. As in YUSHKEVIC (1977) and YUSHKEVIC and FAINBERG (1979) we consider randomised policies which are history remembering with respect to past *states*. An appropriate modeling of decision processes in continuous time with randomized policies which are history remembering with respect to *actions* seems not yet to exist. The problem is that for a process which is controlled by a randomized policy the trajectories of the realized actions are in general extremely irregular functions of time. Consider for example the simple case of a system with only one state and two actions and suppose that this system is controlled by the policy which prescribes at every instant to choose each action with probability ½. We do not know whether the trajectories of the realized decisions are measurable functions of time or not.

The existing results in the literature on continuous time processes both in Markov decision theory and control theory mainly concern conditions for the existence of stationary optimal policies, which are obtained by analysis of the continuous time optimality equation. Structural results of optimal policies are scarce. For the analysis of the CTMDP we use in this paper the method of *discrete time approximation.* In many places in control and decision theory this method plays some role, e.g. in the definition of stochastic integrals or in establishing the existence of a solution of the continuous time optimality equation. Given a CTMDP we construct a sequence of discrete time Markov decision processes (DTMDP) with decreasing distance between two successive decision epochs. On one hand these discrete time Markov decision processes are in accordance with the usual definitions, on the other hand the formulation is chosen in such a way that the connection with the CTMDP becomes clear. We will study a CTMDP by analyzing the conditions under which it can be approximated by a sequence of DTMDP's, say with decision epochs $\{nk^{-1}, n \in \mathbb{N}_0\}$, $k \geq 1$ and then by obtaining properties in discrete time which are invariant under taking limits as $k \to \infty$.

In our approach we cannot avoid to go through a rather technical derivation of sufficient conditions for the approximation procedure. However, because of the generality of our model and theorems, many interesting results for more specific models e.g. waiting line, maintenance-replacement and inventory models can be obtained without much additional effort. Especially for establishing the structure of optimal policies the approximation approach seems to be very powerful.

Structural results of optimal policies in discrete time models can often be obtained by the method of induction to the number of time periods. The discrete time approximation then in its turn is used to carry over these structural results to the continuous time model. In chapters 5, 6 and 7 we will give some applications of this type. Of course the method of discrete time approximation is also important from a numerical point of view. We have, however, ourselves not addressed to questions of this type.

2.2. MARKOV DECISION DRIFT PROCESSES WITH CONTINUOUS TIME PARAMETER.

DEFINITION 2.2.1. *A Markov decision drift process with continous time parameter (CTMDP)* is a nine-tuple $(S, A_1, A_2, q, \Pi, p, c_1, c_2, f)$, where

(i) S is a complete separable metric space, with metric ρ and Borelfield $\mathcal{S}$.

(ii) A_i is a complete separable metric space with metric ρ_i and Borelfield $\mathcal{A}_i$, i=1,2.

(iii) $S \times A_i$ is endowed with a metric generating the product topology, i=1,2.

(iv) q is a real-valued, non-negative measurable mapping on $S \times A_1$.

(v) Π is a transition probability from $S \times A_1$ to S.

(vi) p is a transition probability from $S \times A_2$ to S.

(vii) c_i is a real-valued measurable function on $S \times A_i$, i=1,2.

(viii) f is a drift function for S.

The following interpretation will be given to the component parts of a CTMDP.

(i) S denotes the *state space* of the process.

(ii) A_1 denotes the *set of (generator) controls.*

(iii) A_2 denotes the *set of impulsive controls.*

(iv) q(.,.) is the *jump rate* i.e. if at epoch t the actual state of the system is s and control a is chosen during $(t, t+\Delta t)$ the probability that no jump will occur during $(t, t+\Delta t)$ is equal to $1 - q(s,a)\Delta t + o(\Delta t)$, for small positive Δt.

(v) $\Pi(.,.,.)$ is the *jump distribution* i.e. if at epoch t the actual state of the system is s and control a is chosen then $\Pi(s,a,\Lambda)$ is the conditional probability that a jump at t brings the state of the system in $\Lambda \in \mathcal{S}$, given that a jump occurs at t.

(vi) $p(.,.,.)$ is the *impulsive jump distribution* i.e. if the actual state of the system is s and impulsive control a is chosen then $p(s,a,\Lambda)$ is the probability that the process jumps instantaneously to $\Lambda \in S$.

(vii) $c_1(.,.)$ is the *cost rate* i.e. whenever the system is in state s and control a is chosen a cost of $c_1(s,a)$ per unit time is incurred.

(viii) $c_2(.,.)$ is the *lump cost* i.e. whenever the system is in state s and impulsive control a is chosen an immediate lump cost of $c_2(s,a)$ is incurred.

(ix) f is the drift function describing the behaviour of the process between two successive jump epochs independently of the policy.

In addition to the assumptions that are made tacitly in definition 2.2.1., the following explicit assumptions will be made throughout this monograph. Remember the notation $\mathcal{P}(S)$ for the class of all probability measures on a metric space S. We assume that the space $\mathcal{P}(S)$ is endowed with the topology of weak convergence. With this topology $\mathcal{P}(S)$ can be represented as a complete separable metric space (see theorems 6.2. and 6.4. in chapter II of PARTHASARATHY (1967) or pages 237, 238 of BILLINGSLEY (1968)). By identifying the probability measures degenerated in one point with that point, S can be considered as a subset of $\mathcal{P}(S)$.

<u>ASSUMPTION 2.2.2</u>. Let $(S, A_1, A_2, q, \Pi, p, c_1, c_2, f)$ be a CTMDP. We assume:

(i) $q(.,.)$ is bounded and continuous on $S \times A_1$.

(ii) $\Pi(.,.)$ is continuous as a function from $S \times A_1$ into $\mathcal{P}(S)$.

(iii) $p(.,.)$ is continuous as a function from $S \times A_2$ into $\mathcal{P}(S)$.

(iv) $c_i(.,.)$ is continuous on $S \times A_i$, $i=1,2$.

Next we introduce for a given CTMDP the class of possible policies.

<u>DEFINITION 2.2.3</u>. Let $(S, A_1, A_2, q, \Pi, p, c_1, c_2, f)$ be a CTMDP. A *policy* consists of a closed subset V of S (the *impulsive control set*) and a pair $R=(R_1,R_2)$ of transition probabilities from $J[0,\infty)\times[0,\infty)$ to A_1 and A_2 respectively (R_1 is called the *control rule* and R_2 the *impulsive control rule*), such that

(i) $R_i(.,t)$ is r_t-measurable as a function from $J[0,\infty)$ into $P(A_i)$ for all $t\geq 0$ and $i=1,2$.

(ii) there exists a $\delta>0$ such that for all $t\geq 0$ and all $x\in J[0,\infty)$ for which $\pi_t x\in V$

$$\int_{A_2} p(\pi_t x,a,W(\delta))dR_2(x,t)(a) = 1$$

where

$$W(\delta) = \{j\in S: \inf\{u: f(j,u)\in V\}>\delta\}.$$

The following interpretation and comments can be given to the component parts of a policy.

(i) $R_1(x,t)$ denotes the randomized control that is chosen at time t when the history of the process upto t is given by $r_t x$. (the requirement that $R_i(.,t)$ is r_t-measurable ensures that the policy is *non-anticipating*).

(ii) $R_2(x,t)$ denotes the randomized impulsive control that is chosen at time t under the history $r_t x$ provided the impulsive control rule is active.

(iii) V denotes the subset of the state space where the impulsive control rule is active.

(iv) the second requirement of definition 2.2.3 ensures that whenever the impulsive control becomes active the system jumps instantaneously to the set $W(\delta)$ from which the set V cannot be reached along the drift function within a time interval of length δ.

(v) in general a policy will be denoted by the pair (V,R), where V denotes a closed subset of S and R a pair (R_1,R_2) of transition probabilities.

The definition of a policy (V,R) implies that the epochs at which impulsive controls are chosen are generated by the successive entrance times in the closed subset V of the state space. Hence all policies are stationary w.r.t. the impulsive controls in the following weak sense. Once the decision maker has decided to choose an impulsive control in some state he has to choose an impulsive control at every entrance in that state, although the specific impulsive control that is chosen may depend on the entire history. A natural generalization is obtained by replacing the entrance times of closed sets by arbitrary stopping times. In this monograph we restrict ourselves to entrance times by which we get round several technical difficulties.

<u>DEFINITION 2.2.4</u>. Let $(S, A_1, A_2, q, \Pi, p, c_1, c_2, f)$ be a CTMDP and (V,R) a policy, with $R=(R_1,R_2)$

(i) (V,R) is *deterministic* (or *pure*) if $R_i(x,t) \in A_i$ for all $(x,t) \in J[0,\infty) \times [0,\infty)$, i=1,2. (Recall that $A_i \subset P(A_i)$).

(ii) (V,R) is *memoryless* if $R_i(.,t)$ is π_t-measurable for all $t \geq 0$, i.e. for all $t \geq 0$ and any Borelset B of $P(A_i)$ there exists a $C \in \mathcal{S}$ such that $\{x: R_i(x,t) \in B\} = \pi_t^{-1}C$.

(iii) (V,R) is *stationary* if $R_i(.,.)$ as a function from $J[0,\infty) \times [0,\infty)$ into $P(A_i)$ is π-measurable, where π is a mapping from $J[0,\infty) \times [0,\infty)$ on S defined by $\pi(x,t) = \pi_t x$.

So far we have given definitions of a CTMDP and of a policy for these processes. A first step in comparing two policies for a given CTMDP is made by defining for any given CTMDP and any given policy a probability measure on $J[0,\infty)$. In this way we can consider a CTMDP under a fixed policy as a random element of $J[0,\infty)$. First we give some lemma's which will be used hereafter.

LEMMA 2.2.5. Let S, S_i, i=1,2,3, be metric spaces. If f_i is a measurable function from S into S_i, i=1,2, and g a measurable function from $S_1 \times S_2$ into S_3, then the function h from S into S_3, defined by

$$h(s) := g(f_1(s), f_2(s))$$

is measurable.

PROOF. Define the function f from S into $S_1 \times S_2$ by $f(s) := (f_1(s), f_2(s))$. Then $h = g \circ f$, which implies $h^{-1}(U) = f^{-1}(g^{-1}(U))$ for all $U \in \mathcal{S}_3$. Since g is measurable it is sufficient to show that f is measurable. From the measurability of f_1 and f_2 follows that

$$f^{-1}(U_1 \times U_2) = f_1^{-1}(U_1) \cap f_2^{-1}(U_2) \in \mathcal{S}, \text{ for all } U_i \in \mathcal{S}_i, \ i=1,2.$$

On the other hand, the collection $\{A \subset S_1 \times S_2: f^{-1}(A) \in \mathcal{S}\}$ is a σ-algebra. Hence $\{A \subset S_1 \times S_2: f^{-1}(A) \in \mathcal{S}\}$ is a σ-algebra including $\mathcal{S}_1 \times \mathcal{S}_2$. □

LEMMA 2.2.6. Let S_1 and S_2 be metric spaces, P a transition probability from S_1 to S_2 and f a real-valued measurable function on $S_1 \times S_2$. The real-valued function h defined by

$$h(s_1):=\int_{S_2} f(s_1,s_2)\ \ P(s_1,ds_2)$$

is measurable on S_1.

PROOF. See page 74 of NEVEU (1965). □

LEMMA 2.2.7. Let V be a closed subset of a metric space (S,ρ) and f a drift function for S. The function τ_V from S into $[0,\infty]$ defined by

$$(2.2.1)\qquad \tau_V(j):=\inf\{t\colon f(j,t)\in V\}$$

is lower semicontinuous.

PROOF. It is sufficient to show that the set $\{j\in S\colon \tau_V(j)\le t\}$ is closed in S for all $t\ge 0$. Choose $0\le t<\infty$ and let j, j_n, $n\ge1$ be elements of S such that $\lim_{n\to\infty}\rho(j_n,j)=0$ and $\tau_V(j_n)\le t$ for $n\ge 1$. Since V is closed and f is continuous there exists a sequence $(s_n)_{n=1}^{\infty}$ such that $f(j_n,s_n)\in V$ and $s_n\le t$ for $n\ge 1$. Hence $(s_n)_{n=1}^{\infty}$ has a convergent subsequence $(s_{n_k})_{k=1}^{\infty}$. Put $s:=\lim_{k\to\infty}s_{n_k}$. Then $s\le t$ and $f(j,s)\in V$ which implies $\tau_V(j)\le t$. It follows that the set $\{j\colon \tau_V(j)\le t\}$ is closed. □

LEMMA 2.2.8. Let V be a closed subset of a metric space (S,ρ) and f a drift function for S. The function σ_V from S into $V\cup\{\zeta\}$ defined by

$$(2.2.2)\qquad \sigma_V(j):=\begin{cases} f(j,\ \tau_V(j)) & \text{if } \tau_V(j)<\infty\\ \zeta & \text{otherwise}\end{cases}$$

is measurable on S.

PROOF. A direct consequence of lemma 2.2.5. and lemma 2.2.7. □

At this point we have to consider for $n\geq 1$ the finite product space $(\overline{[0,\infty)\times S},\tilde{\rho})^n$ endowed with product metric. (See (1.5.1) for the definition of $\tilde{\rho}$). Put

$$J_n: =\{(t_j,i_j)_{j=1}^n \in (\overline{[0,\infty)\times S})^n: \quad 0 = t_1 \leq t_2 \leq \ldots \leq t_n;$$
$$t_{j+2} - t_j > 0 \text{ if } t_j < \infty\}.$$

<u>LEMMA 2.2.9</u>. Let $n \geq 1$, V a closed subset of a metric space (S,ρ) and f a drift function for S. The mapping g_V from J_n into $J[0,\infty)$ defined by

$$(2.2.3) \qquad g_V(t_1,i_1,\ldots,t_n,i_n) := \begin{cases} (t_1,i_1,\ldots,t_n,i_n,\infty,\zeta,\ldots) \text{ if } i_n=\zeta \\ \text{or } t_n=t_{n-1}=t_n+\tau_V(i_n) \text{ or } \tau_V(i_n)=\infty \\ (t_1,i_1,\ldots,t_n,i_n,t_n+\tau_V(i_n),i_n,\infty,\zeta,\ldots) \\ \text{otherwise} \end{cases}$$

is measurable on J_n.

<u>PROOF</u>. Define on $[0,\infty] \times J_n$ the function h by

$$h(s,(t_1,i_1,\ldots,t_n,i_n)) := \begin{cases} (t_1,i_1,\ldots,t_n,i_n,\infty,\zeta,\ldots) \text{ if } i_n=\zeta \\ \text{or } t_n=t_{n-1}=t_n+s \text{ or } s=\infty \\ (t_1,i_1,\ldots,t_n,i_n,t_n+s,i_n,\infty,\zeta,\ldots) \\ \text{otherwise.} \end{cases}$$

Then $h(.,.)$ is measurable on $[0,\infty] \times J_n$ and $g_V(t_1,i_1,\ldots,t_n,i_n) = h(\tau_V(i_n), (t_1,i_1,\ldots,t_n,i_n))$. The lemma follows from lemma 2.2.5. and lemma 2.2.7. $\square$

The function g_V assigns to every history upto the n^{th} jump a complete path from $J[0,\infty)$. One might wonder why not the simpler definition $\tilde{g}(t_1,i_1,\ldots,t_n,i_n) = (t_1,i_1,\ldots,t_n,i_n,\infty,\zeta,\ldots)$ is chosen. Note that $\tilde{g}$ is essentially the t_n-extension although defined on the set J_n instead of $J[0,t_n]$ (see also definition 1.4.6). The reason why we need the function g_V is the following. From the definition of a policy (V,R) follows that for every $x\in J[0,\infty)$ and $t\geq 0$ the t-projection $\pi_t x$ indicates whether the impulsive control rule is active or not. Hence we should extend a history

$z=(t_1,i_1,\ldots,t_n,i_n)\in J_n$ with $i_n \in V$ in such a way that the t_n-projection of this extension is an element of V. Now consider the case where $i_n \in V$, $t_{n-1}<t_n<\infty$ and $f(i_{n-1},t_n-t_{n-1})\notin V$. Since $\tau_V(i_n)=0$ we have $g_V(z) = (t_1,i_1,\ldots,t_n,i_n,t_n,i_n,\infty,\zeta,\ldots)$. From the definition of t-projection (definition 1.4.13) follows that $\pi_{t_n} g_V(z)=i_n$ while $\pi_{t_n}\tilde{g}(z)=f(i_{n-1},t_n-t_{n-1})\notin V$. So the impulsive control rule is active at epoch t_n on the sample path $g_V(z)$ but not on the sample path $\tilde{g}(z)$.

Let $(S, A_1, A_2, q, \Pi, p, c_1, c_2, f)$ be a CTMDP, (V,R) a policy and P_0 an (initial) distribution on S. The construction of a probability measure on $J[0,\infty)$ induced by this CTMDP, this policy and this initial distribution proceeds as follows. First we construct for any fixed $z\in J_n$ a probability measure $Q^{(n)}(z)$ on $\overline{[0,\infty)\times S}$. Next we show in theorem 2.2.14. below that for all $n\geq 1$ the mapping $z\to Q^{(n)}(z)$ is a transition probability from J_n to $\overline{[0,\infty)\times S}$. With these transition probabilities we construct by induction probability measures $P^{(n)}_{(V,R)}$ on $(\overline{[0,\infty)\times S})^n$.
The theorem of Ionescu Tulcea then yields a probability measure $P_{(V,R)}$ on $(\overline{[0,\infty)\times S})^\infty$. Finally we show that $P_{(V,R)}$ is concentrated on $J[0,\infty)$.
For fixed CTMDP and initial distribution $P_{(V,R)}$ will be referred to as *the probability measure induced by (V,R)*.
For the construction of the probability measures $Q^{(n)}(z)$ on $\overline{[0,\infty)\times S}$ we need the following lemma.

<u>LEMMA 2.2.10</u>. Let

(2.2.4) $\quad \mathcal{G} := \{G \subset \overline{[0,\infty)\times S} : G = B\times F$ with B an interval in $[0,\infty)$ and $F\in\mathcal{S}$ or $G=B\times F\cup\{(\infty,\zeta)\}$ with B an interval in $[0,\infty)$ and $F\in\mathcal{S}\}$.

Any σ-additive set function $\tilde{P}$ mapping $\mathcal{G}$ into $[0,1]$ such that $\tilde{P}(\overline{[0,\infty)\times S})=1$ can be uniquely extended to a probability measure on the Borelfield of $\overline{[0,\infty)\times S}$.

<u>PROOF</u>. Since $\mathcal{G}$ is a semi algebra which generates the Borelfield of $\overline{[0,\infty)\times S}$ the lemma is an immediate consequence of proposition 1.6.1. on page 25 of NEVEU (1965). □

<u>DEFINITION 2.2.11</u>. For any $z:=(t_j,i_j)_{j=1}^{n} \in J_n$ we define a σ-additive set function $Q^{(n)}(z)$ on G as follows (with suppression of the index V in the functions τ_V, σ_V and g_V).

(2.2.5) $$Q^{(n)}(z)\ ([0,t]\times S):=0 \qquad \text{for } t<t_n;$$

(2.2.6) $$Q^{(n)}(z)\ ([0,t]\times F):=$$

$$\int_{s=0}^{(t-t_n)\wedge\tau(i_n)} \exp\Big(-\int_{u=0}^{s}\int_{A_1} q(f(i_n,u),a)\,dR_1(g(z),t_n+u)(a)\,du\Big)\ .$$

$$\cdot \int_{A_1} q(f(i_n,s),a)\ .\ \Pi(f(i_n,s),a,F)\ dR_1(g(z),t_n+s)(a)\,ds\ +$$

$$+1_{\{\tau(i)\le t-t_n\}}(i_n)\ \exp\Big(-\int_{u=0}^{\tau(i_n)}\int_{A_1} q(f(i_n,u),a)\,dR_1(g(z),t_n+u)(a)du\Big).$$

$$\cdot \int_{A_2} p(\sigma(i_n),a,F)\ dR_2(g(z),t_n+\tau(i_n))(a)$$

for $t_n \le t < \infty$ and $F \in S$;

(2.2.7) $$Q^{(n)}(z)([t,\infty)\times F) = Q^{(n)}(z)([t,\infty]\times F):=1_{\{\tau(i)\ge t-t_n\}}(i_n)\ .$$

$$\cdot \int_{s=t-t_n}^{\tau(i_n)} \exp\Big(-\int_{u=0}^{s}\int_{A_1} q(f(i_n,u),a)\ dR_1(g(z),t_n+u)(a)\,du\Big)\ .$$

$$\cdot \int_{A_1} q(f(i_n,s),a)\,\Pi(f(i_n,s),a,F)\ dR_1(g(z),t_n+s)(a)\ ds\ +$$

$$+ 1_{\{t-t_n \leq \tau(i) < \infty\}}(i_n) \cdot$$

$$\cdot \exp\left(- \int_{u=0}^{\tau(i_n)} \int_{A_1} q(f(i_n,u),a) \, dR_1(g(z),t_n+u)(a) \, du\right) \cdot$$

$$\cdot \int_{A_2} p(\sigma(i_n),a,F) \, dR_2(g(z),t_n+\tau(i_n))(a)$$

for $t_n \leq t < \infty$ and $F \in S$;

(2.2.8) $\quad Q^{(n)}(z)\{(\infty,\zeta)\} :=$

$$1_{\{\tau(i)=\infty\}}(i_n) \exp\left(- \int_{u=0}^{\infty} q(f(i_n,u),a) \, dR_1(g(z),t_n+u)(a) \, du\right)$$

for $t_n < \infty$;

(2.2.9) $\quad Q^{(n)}(z)([t,\infty) \times F) := Q^{(n)}(z)([t_n,\infty) \times F)$

for $t < t_n < \infty$ and $F \in S$;

(2.2.10) $\quad Q^{(n)}(z)(\{\infty,\zeta\}) := 1 \qquad$ for $t_n = \infty$.

Finally the definition of $Q^{(n)}(z)$ on G is completed in a straightforward way on unions and differences.

REMARK 2.2.12. (i) The expressions on the right hand side of (2.2.6), (2.2.7) and (2.2.8) are well defined. This follows as a special consequence from theorem 2.2.14. below.

(ii) $Q^{(n)}(z)$ denotes for $z=(t_j,i_j)_{j=1}^{n}$ the simultaneous conditional probability distribution of the first jump epoch after t_n and the state of the system after that jump, given the history z. In defining $Q^{(n)}(z)$ we distinguish between the events that a jump is caused by an impulsive control and that a jump occurs in the natural process. Consider for example formula (2.2.6). The first term denotes the probability that the next jump will take place before the next impulsive control time $\tau(i_n)+t_n$ and that this jump brings the state of the system into the set F. Since $\tau(i_n)+t_n$ is the first epoch after t_n at which an impulsive control is chosen, it is clear that a jump before $\tau(i_n)+t_n$ corresponds to a jump in the natural process. The second term in (2.2.6) denotes the probability that before $\tau(i_n)+t_n$ no jump in the natural process occurs and that the jump at $\tau(i_n)+t_n$, caused by an impulsive control, brings the state of the system into the set F. Note that in the integrals in both terms the state of the system is continuously updated with the drift function f.

LEMMA 2.2.13.

(i) $Q^{(n)}(z)$ defined above is a σ-additive set function mapping $\mathcal{G}$ into $[0,1]$ such that $Q^{(n)}(z)(\overline{[0,\infty)\times S})=1$.

(ii) $Q^{(n)}(z)$ can be uniquely extended to a probability measure on $\overline{[0,\infty)\times S}$. This extension will also be denoted by $Q^{(n)}(z)$.

PROOF.

(i) By (2.2.5) through (2.2.10) $Q^{(n)}(z)\ (G)$ is defined for all $G\in\mathcal{G}$. The proof of the σ-additivity of $Q^{(n)}(z)$ is straightforward from the definition. Finally (2.2.7) and (2.2.8) yield for t = 0

$$Q^{(n)}(z)(\overline{[0,\infty)\times S}) = 1.$$

(ii) A direct consequence of lemma 2.2.10. □

THEOREM 2.2.14. $Q^{(n)}(.)$ defines a transition probability from J_n to $\overline{[0,\infty)\times S}$.

PROOF. According to lemma 2.2.13. above $Q^{(n)}(z)$ is a probability measure on $\overline{[0,\infty)\times S}$ for all $z\in J_n$.

What remains to show is that $Q^{(n)}(.)(C)$ is measurable on J_n for all Borelsets C of $\overline{[0,\infty) \times S}$. Since $\{C: Q^{(n)}(.)(C)$ is measurable$\}$ is a monotone class and the σ-algebra generated by a (semi)-algebra $\mathcal{A}$ is identical with the monotone class generated by $\mathcal{A}$, it follows that it is sufficient to prove that $Q^{(n)}(.)(G)$ is measurable on J_n for all $G \in \mathcal{G}$.
We will show this for $G = [0,t] \times F$, with $0 < t < \infty$ and $F \in \mathcal{S}$ (the proof for other sets from $\mathcal{G}$ is quite similar). Since the set $\{(t_1,i_1,\ldots,t_n,i_n) \in J_n: t_n > t\}$ is a Borelset in J_n it is sufficient to show that the right hand side of (2.2.6) is measurable in z. This follows by systematic application of the lemma's 2.2.5. upto 2.2.9. For example to show that the mapping on $J_n \times [0,\infty)$ into $\mathbb{R}$ defined by

(2.2.11) $$((t_1,i_1,\ldots,t_n,i_n),u) \to \int_{A_1} q(f(i_n,u),a)\, dR_1(g(t_1,i_1,\ldots,t_n,i_n),t_n+u)(a)$$

is measurable, we note that the mapping on $S \times J[0,\infty) \times [0,\infty) \times A_1$

$$(s, x, u, a) \to q(s,a)$$

is measurable, while

$$(s, x, u) \to R_1(x, T_n(x)+u)(A)$$

is measurable on $S \times J[0,\infty) \times [0,\infty)$, for all $A \in \mathcal{A}_1$. Hence, by lemma 2.2.6.

$$(s, x, u) \to \int_{A_1} q(s,a)\, dR_1(x, T_n(x)+u)(a)$$

is measurable on $S \times J[0,\infty) \times [0,\infty)$.
Since

$$((t_1,i_1,\ldots,t_n,i_n), u) \to f(i_n,u)$$

and

$$((t_1,i_1,\ldots,t_n,i_n), u) \to g(t_1,i_1,\ldots,t_n,i_n)$$

are measurable on $J_n \times [0,\infty)$ (see lemma 2.2.9.), it follows from lemma 2.2.5. that the mapping defined by (2.2.11) is measurable on $J_n \times [0,\infty)$. □

REMARK 2.2.15. From definition 2.2.11. (in particular formula (2.2.6)) follows that

(i) for $z = (t_1, i_1, \ldots, t_n, i_n)$ we have

$$Q^{(n)}(z)(\{t_n\} \times S) > 0 \text{ iff } i_n \in V$$

and

(ii) for $z = (t_1, i_1 \ldots, t_n, i_n)$ with $i_n \in V$ we have, by part (iii) of definition 2.2.3.

$$Q^{(n)}(z)(\{t_n\} \times W(\delta)) = 1.$$

DEFINITION 2.2.16. Let P_0 be a given (initial) distribution on S. We define probability measures $P^{(n)}_{(V,R)}$ on $(\overline{[0,\infty)\times S})^n$, $n \geq 1$ and $P_{(V,R)}$ on $(\overline{[0,\infty)\times S})^\infty$ as follows:

(2.2.12) $$P^{(1)}_{(V,R)}(B \times F) := 1_B(0).P_0(F), \text{ for } B \text{ a Borelset in } [0,\infty) \text{ and } F \in \mathcal{S}.$$

(2.2.13) $$P^{(n)}_{(V,R)}(G_1 \times \ldots \times G_n) := \int_{G_1 \times \ldots \times G_{n-1}} Q^{(n-1)}(z)(G_n)\, dP^{(n-1)}_{(V,R)}(z),$$

for all Borelsets G_i in $\overline{[0,\infty)\times S}$, $i=1,\ldots,n$; $n \geq 2$.

(2.2.14) $$P_{(V,R)}(G_1 \times \ldots \times G_n \times \overline{[0,\infty)\times S} \times \ldots) = P^{(n)}_{(V,R)}(G_1 \times \ldots \times G_n)$$

for $n \geq 1$ and Borelsets G_i of $\overline{[0,\infty)\times S}$, $i=1,\ldots,n$.

THEOREM 2.2.17. By (2.2.14) a unique probability measure on $(\overline{[0,\infty)\times S})^\infty$ is defined.

PROOF. A direct consequence of the theorem of Ionescu Tulcea (see NEVEU (1965)). □

THEOREM 2.2.18. The probability measure $P_{(V,R)}$ is concentrated on $J[0,\infty)$.

PROOF. With remark 2.2.15. one easily shows that the probability measures $P^{(n)}_{(V,R)}$ are concentrated on J_n for $n\geq 1$. What remains to prove is that for all $t\geq 0$

$$(2.2.15) \qquad \lim_{n\to\infty} P_{(V,R)}\{(t_j,i_j)_{j=1}^{\infty} : t_n\leq t\} = 0,$$

i.e. $P_{(V,R)}$ does not assign a positive probability to those elements $(t_j,i_j)_{j=1}^{\infty}\in\overline{([0,\infty)\times S)}^{\infty}$ for which the sequence $(t_j)_{j=1}^{\infty}$ is bounded.
For $z=(t_j,i_j)_{j=1}^{n}\in J_n$ with $i_n\in W(\delta)$ we have, according to (2.2.6)

$$(2.2.16) \qquad Q^{(n)}(z)([0,t_n+\delta]\times S) =$$

$$= \int_{s=0}^{\delta} \exp(-\int_{u=0}^{s} \int_{A_1} q(f(i_n,u),a)\,dR_1(g(z),t_n+u)(a)du) \;.$$

$$\cdot \int_{A_1} q(f(i_n,s),a).dR_1(g(z),t_n+s)(a)ds.$$

With a well-known theorem on changing independent variables (see page 377 of TITCHMARSH (1939)) it follows from (2.2.16) that

$$(2.2.17) \qquad Q^{(n)}(z)([0,t_n+\delta]\times S) =$$

$$= 1 - \exp(-\int_{u=0}^{\delta} \int_{A_1} q(f(i_n,u),a)\,dR_1(g(z),t_n+u)(a)du).$$

Combining this with assumption 2.2.2.(i) we find that there exists a number b, such that

$$(2.2.18) \qquad Q^{(n)}(z)([0,t_n+\delta]\times S) \leq 1-\exp(-b\delta)$$

for all $z = (t_1,i_1,\ldots,t_n,i_n) \in J_n$, with $i_n \in W(\delta)$.
Reasoning as above and using part (iii) of definition 2.2.3. we have

$$(2.2.19) \qquad Q^{(n)}(z)([0,t_n+\delta]\times (W(\delta))^c) \leq 1 - \exp(-b\delta)$$

for all $z = (t_1,i_1,\ldots,t_n,i_n) \in J_n$.

From remark 2.2.15, (2.2.18) and (2.2.19) follows that for all $u \le \delta$ and all $t > 0$ the conditional probability

$$(2.2.20) \qquad P_{(V,R)}(\{(t_j,i_j)_{j=1}^{\infty}: t_{n+2} \le t+u\} \mid \{(t_j,i_j)_{j=1}^{\infty}: t_n \ge t\}) \le 2(1-\exp(-bu)).$$

Choose $0 < \delta' < \delta$ such that $2(1-\exp(-b\delta')) < 1$ and define a sequence $(X_m)_{m=1}^{\infty}$ of independent random variables on a common probability space $(\Omega, A, \mathbb{P})$ such that $\mathbb{P}(X_m = \delta') = 1-2(1-\exp(-b\delta'))$ and $\mathbb{P}(X_m = 0) = 2(1-\exp(-b\delta'))$.
Then it follows from (2.2.20) by induction on n, that for $n \ge 1$ and $t \ge 0$

$$(2.2.21) \qquad P_{(V,R)}\{(t_j,i_j)_{j=1}^{\infty}: t_{2n+1} \le t\} \le \mathbb{P}(\sum_{m=1}^{n} X_m \le t).$$

Since the right hand side of (2.2.21) converges to zero as $n \to \infty$ it follows that $\lim_{n\to\infty} P_{(V,R)}\{(t_j,i_j)_{j=1}^{\infty}: t_{2n+1} \le t\} = 0$. □

We conclude this section with a remark.

<u>REMARK 2.2.19.</u>

(i) It is worthwhile to observe that we can deduce $P_{(V,R)}$-a.e. from a path $(t_j,i_j)_{j=1}^{\infty} \in J[0,\infty)$ the epochs at which impulsive controls are chosen. Since an impulsive control causes a jump the impulsive control epochs belong to the set $\{t_j: j \ge 1\}$ and are identified by the set $\{t_j: t_j - t_{j-1} = \tau_V(i_{j-1})\}$, $P_{(V,R)}$-a.e.

(ii) From the definition of $Q^{(n)}$ and $P_{(V,R)}$ follows easily that for $n \ge 1$

$$P_{(V,R)}(\{x: T_n(x) < t < T_{n+1}(x);\ \pi_t x \in V \text{ for some } t > 0\}) = 0$$

and

$$P_{(V,R)}(\{x: \pi_t x \in V;\ \pi_{t^+} x \notin W(\delta) \text{ for some } t > 0\}) = 0$$

where for $x \in J[0,\infty)$

$$\pi_{t^+} x := \lim_{s \downarrow t} \pi_s x \text{ and } \pi_{t^-} x := \lim_{s \uparrow t} \pi_s x.$$

(see for the definition of $\pi_t x$ and $T_n x$ definition 1.4.13 and section 1.6 respectively).

2.3. MARKOV DECISION DRIFT PROCESSES WITH DISCRETE TIME PARAMETER.

DEFINITION 2.3.1. *A Markov decision drift process with discrete time parameter (DTMDP)* is a nine tuple $(S,A_1,A_2,p_1,p_2,c_1,c_2,f,k)$, where

(i) S is a complete separable metric space with metric ρ and Borelfield $\mathcal{S}$.

(ii) A_i is a complete separable metric space with metric ρ_i and Borelfield $\mathcal{A}_i$, i=1,2.

(iii) $S\times A_i$ is endowed with a metric generating the product topology, i=1,2.

(iv) p_i is a transition probability from $S\times A_i$ to S, i=1,2.

(v) c_i is a real-valued measurable function on $S\times A_i$, i=1,2.

(vi) f is a drift function for S.

(vii) k is a natural number.

The following interpretation can be given to the component parts of a DTMDP:

(i) S denotes the *state space* of the process.

(ii) A_1 denotes the *set of controls.*

(iii) A_2 denotes the *set of impulsive controls.*

(iv) $p_i(.,.,.)$, i=1,2 are the *one step transition distributions* i.e. if at the n-th decision epoch the actual state of the system is s and control a (i=1) or impulsive control a (i=2) is chosen, then $p_i(s,a,F)$ is the probability that the state of the system at the (n+1)-th decision epoch belongs to the set $\{f(s,k^{-1}):s\in F\}$.

(v) $c_i(.,.)$ are the *one step cost functions,* i.e. whenever the system is in state s and control a (i=1) or impulsive control a (i=2) is chosen a direct cost $c_i(s,a)$ is incurred.

(vi) the drift function f plays a role in the interpretation of the one step transition distributions.

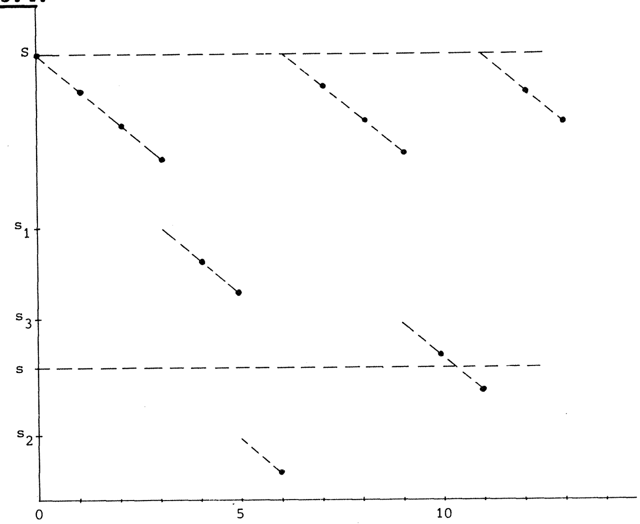
(vii) k denotes the *time parameter,* which indicates that two successive decision epochs are k^{-1} apart.

REMARK 2.3.2. When one is merely interested in Markov decision (drift) processes with discrete time parameter the definition of a DTMDP can be simplified considerably, since in general it is not necessary to distinguish between A_1 and A_2, p_1 and p_2 and between c_1 and c_2. However, since we will compare CTMDP's and DTMDP's we have chosen a definition of a DTMDP, which is as close as possible to the definition of CTMDP. Note that the classical definition of a Markov decision process with discrete time parameter can be obtained from the definition of a DTMDP by choosing $A_2=\emptyset$ (which implies that p_2 and c_2 are undefined), $f(s,t)=s$ and $k=1$.

First we give an example to explain how a sample path of a DTMDP with time parameter k under a fixed policy can be seen as element of $J[0,\infty)$. Moreover, this example illustrates the role of the drift function.

EXAMPLE 2.3.3. The total demand of the customers arriving at a warehouse during a time interval of length k^{-1} is a random variable with known distribution function. At the end of each time period of length k^{-1} the demands of the customers, arrived during that period, are fulfilled as long as the stock level is positive. Unfilled demands are backlogged. The warehouse management can place orders at the beginning of each interval which will be delivered at the end of this interval before the demands of the customers are fulfilled. Moreover, there is a deterministic demand of σ per unit time.
Suppose that the system is controlled by a (s,S)-rule i.e. as soon as the inventory level reaches or drops below s it is raised upto S. Consider the discrete time stochastic process describing at any decision epoch the inventory level (where negative values denote the amount to be backlogged), given that the (s,S)-rule is used. Note that this process is a discrete time version of the continuous time process in example 1.3.3.
A typical sample path of this process is drawn in figure 2.3.4. below.
This sample path can be represented by
$(0,S,3k^{-1},s_1,5k^{-1},s_2,6k^{-1},S,9k^{-1},s_3,11k^{-1},S,\ldots) \in J[0,\infty)$.

FIGURE 2.3.4.



Stochastic processes generated by a DTMDP can only be observed and controlled at equidistant decision epochs. Therefore the sample paths of these processes have no jumps between two successive decision epochs. It is intuitively clear that a DTMDP under a fixed policy can be considered as a random element of $J[0,\infty)$. However, it is useful to define a discretized subset of $J[0,\infty)$ on which these processes are concentrated. Define for $k \geq 1$

$$L_k: = \{nk^{-1}: n \geq 0\}$$

$$L_k^+: = L_k \cup \{\infty\}$$

and

$$J_k[0,\infty): = \{(t_j,i_j)_{j=1}^{\infty} \in J[0,\infty): t_j \in L_k^+, t_j < t_{j+1} \text{ iff } t_j < \infty, j \geq 1\}.$$

Next we introduce for a given DTMDP the class of possible policies.

DEFINITION 2.3.5. Let $(S,A_1,A_2,p_1,p_2,c_1,c_2,f,k)$ be a DTMDP. *A policy* consists of a closed subset V of S and a pair $R=(R_1,R_2)$ of transition probabilities from $J_k[0,\infty)\times[0,\infty)$ to A_1 and A_2 respectively such that $R_i(.,t)$ is r_t-measurable as a function from $J_k[0,\infty)$ into $\mathcal{P}(A_i)$ for all $t\in L_k$, $i=1,2$.

The interpretation of the component parts of a policy is the same as for the continuous time case. For the classical Markov decision process with discrete time parameter the definition of a policy simplifies to merely a transition probability R_1 from $J_k[0,\infty)\times[0,\infty)$ to A_1.

DEFINITION 2.3.6. Let $(S,A_1,A_2,p_1,p_2,c_1,c_2,f,k)$ be a DTMDP and (V,R) a policy with $R=(R_1,R_2)$.

(i) (V,R) is *deterministic* (or *pure)* if $R_i(x,t)\in A_i$ for all $(x,t)\in J_k[0,\infty)\times L_k$, $i=1,2$.

(ii) (V,R) is *memoryless* if $R_i(.,t)$ is π_t-measurable for all $t\in L_k$, $i=1,2$.

(iii) (V,R) is *stationary* if $R_i(.,.)$ as a function from $J_k[0,\infty)\times L_k$ into $\mathcal{P}(A_i)$ is π-measurable, $i=1,2$.

In the classical theory of Markov decision processes with discrete time parameter we assign to every policy R_1 a probability measure $P^{(k)}_{R_1}$ on S^∞ by defining the conditional one-dimensional distributions:

$$(2.3.1)\qquad P^{(k)}_{R_1}(\{\pi_{(m+1)k^{-1}}(x)\in F\}\mid\{\pi_{mk^{-1}}(x)=s;\ R_1(x,mk^{-1})=\nu\}) := = \int_{A_1} p_1(s,a,F)d\nu(a).$$

However, in order to compare stochastic processes generated by a CTMDP with processes generated by a DTMDP we have to construct for a DTMDP and any policy (V,R) a probability law $P^{(k)}_{(V,R)}$ on $(\overline{L_k \times S})^{\infty}$ (where $\overline{L_k \times S} := L_k \times S \cup \{(\infty,\zeta)\}$), in a way similar to definitions 2.2.11 and 2.2.16.
Put for $k \geq 1$

$$(2.3.2) \qquad J_n(k) := \{(t_j,i_j)_{j=1}^n \in J_n : t_j \in L_k^+,\ t_j < t_{j+1} \text{ iff } t_j < \infty,\ 1 \leq j \leq n\}$$

and define for any closed subset V of S and drift function f for S the functions $\tau_V^{(k)}$, $\sigma_V^{(k)}$ on S and $g_V^{(k)}$ on $J_n(k)$ as follows:

$$(2.3.3) \qquad \tau_V^{(k)}(j) := \inf\{\ell k^{-1} : f(j,\ell k^{-1}) \in V,\ \ell = 1,2,\ldots\}.$$

$$(2.3.4) \qquad \sigma_V^{(k)}(j) := \begin{cases} f(j,\tau_V^{(k)}(j)) & \text{if } \tau_V^{(k)}(j) < \infty \\ \zeta & \text{if } \tau_V^{(k)}(j) = \infty \end{cases}$$

$$(2.3.5) \qquad g_V^{(k)}((t_j,i_j)_{j=1}^n) := \begin{cases} (t_1,i_1,\ldots,t_n,i_n,\infty,\zeta,\ldots) \\ \text{if } i_n = \zeta \text{ or } \tau_V^{(k)}(i_n) = \infty \\ (t_1,i_1,\ldots,t_n,i_n,t_n + \tau_V^{(k)}(i_n),i_n,\infty,\zeta,\ldots) \\ \text{otherwise.} \end{cases}$$

The functions $\tau_V^{(k)}$, $\sigma_V^{(k)}$ and $g_V^{(k)}$, which are discrete analogues of τ_V, σ_V and g_V (see (2.2.1), (2.2.2) and (2.2.3)) are measurable functions.

Let $(S,A_1,A_2,p_1,p_2,c_1,c_2,f,k)$ be a DTMDP and (V,R) a policy. Next we define for $n \geq 1$ transition probabilities $Q_k^{(n)}(.)$ from $J_n(k)$ to $\overline{L_k \times S}$, where $Q_k^{(n)}(z)$ is the conditional simultaneous distribution of the next jump epoch and the state of the system after the jump.

DEFINITION 2.3.7. For $z=(\ell_j k^{-1}, i_j)_{j=1}^{n} \in J_n(k)$ we define (with suppression of the index V in the function $\tau_V^{(k)}$, $\sigma_V^{(k)}$ and $g_V^{(k)}$):

$$Q_k^{(n)}(z)(\{\ell k^{-1}\} \times S) := 0 \quad \text{for } \ell \le \ell_n < \infty; \tag{2.3.6}$$

$$Q_k^{(n)}(z)(\{\ell k^{-1}\} \times F) := \tag{2.3.7}$$

$$\prod_{m=1}^{\ell-\ell_n-1} \int_{A_1} p_1(f(i_n, mk^{-1}), a, \{f(i_n, mk^{-1})\})\, dR_1(g^{(k)}(z), (m+\ell_n)k^{-1})(a) \,.$$

$$\{1_{\{\tau^{(k)}(i) > (\ell-\ell_n)k^{-1}\}}(i_n) \,.$$

$$\int_{A_1} p_1(f(i_n, (\ell-\ell_n)k^{-1}), a, F | \{f(i_n, (\ell-\ell_n)k^{-1})\})\, dR_1(g^{(k)}(z), \ell k^{-1})(a) +$$

$$+\, 1_{\{\tau^{(k)}(i) = (\ell-\ell_n)k^{-1}\}}(i_n) \,. \int_{A_2} p_2(\sigma^{(k)}(i_n), a, F)\, dR_2(g^{(k)}(z), \ell k^{-1})(a)\}$$

for $\ell_n < \ell < \infty$ and $F \in S$;

$$Q_k^{(n)}(z)\{(\infty,\zeta)\} := 1_{\{\tau^{(k)}(i)=\infty\}}(i_n) \,. \tag{2.3.8}$$

$$\prod_{m=1}^{\infty} \int_{A_1} p_1(f(i_n, mk^{-1}), a, \{f(i_n, mk^{-1})\})\, dR_1(g^{(k)}(z), (m+\ell_n)k^{-1})(a)$$

for $\ell_n < \infty$;

$$Q_k^{(n)}(z)\{(\infty,\zeta)\} := 1 \quad \text{for } \ell_n = \infty; \tag{2.3.9}$$

$$Q_k^{(n)}(z)(B \times F) := \sum_{\{\ell : \ell k^{-1} \in B\}} Q_k^{(n)}(z)(\{\ell k^{-1}\} \times F) \tag{2.3.10}$$

for $B \subset L_k$ and $F \in S$.

THEOREM 2.3.8. $Q_k^{(n)}(.)$ defines a transition probability from $J_n(k)$ to $\overline{L_k \times S}$.

PROOF. Let $z := (\ell_j k^{-1}, i_j)_{j=1}^{n} \in J_n(k)$. By definition $Q_k^{(n)}(z)(.)$ is σ-additive on the measurable rectangles of $\overline{L_k \times S}$. Hence $Q_k^{(n)}(z)(.)$ defines a probability measure on the Borelsets of $\overline{L_k \times S}$ if

$$Q_k^{(n)}(z)(\overline{L_k \times S}) = 1. \tag{2.3.11}$$

For $\ell_n=\infty$ (2.3.11) follows from (2.3.9) and for $\ell_n<\infty$ (2.3.11) follows from (2.3.7) (with F=S), (2.3.8), (2.3.10) and the fact that

$$\sum_{\ell=1}^{p} \left(\prod_{m=1}^{\ell-1} r_m\right).(1-r_\ell) = 1 - \prod_{m=1}^{p} r_m$$

for any sequence of real numbers $(r_m)_{m=1}^{\infty}$ and any $p\in\mathbb{N}\cup\{\infty\}$.
On the other hand we have to show that $Q_k^{(n)}(.)(C)$ is measurable on $J_n(k)$ for all Borelsets $\overline{L_k\times S}$. As in the proof of theorem 2.2.14. a successive application of the lemma's 2.2.5. upto 2.2.9. yields this proof.
There is, however, one additional complication. We have to show that for fixed m the mapping on $J_n(k)\times A$ defined by

$$((\ell_j k^{-1}, i_j)_{j=1}^{n}, a) \to p_1(f(i_n, mk^{-1}), a, \{f(i_n, mk^{-1})\})$$

is measurable.
This follows from the measurability on $S\times A$ of the mapping

$$(s,a) \to p_1(s,a,\{s\}).$$

□

Now we are ready to construct a probability measure $P_{(V,R)}^{(k)}$ on $J_k[0,\infty)$ for any DTMDP, any policy (V,R)and any initial probability distribution P_0 on S.

<u>DEFINITION 2.3.9</u>. By induction on n we define probability laws $P_{k,(V,R)}^{(n)}$ on $J_n(k)\subset(\overline{L_k\times S})^n$, n≥1 and $P_{(V,R)}^{(k)}$ on $J_k[0,\infty)\subset(\overline{L_k\times S})^{\infty}$ as follows:

(2.3.12) $\quad P_{k,(V,R)}^{(1)}(B\times F) := 1_B(0).P_0(F)$, for B a Borelset in L_k and $F\in\mathcal{S}$

(2.3.13) $$P_{k,(V,R)}^{(n)}(G_1\times\ldots\times G_n) := \int_{G_1\times\ldots\times G_{n-1}} Q_k^{(n-1)}(z)(G_n)\,dP_{k,(V,R)}^{(n-1)}(z)$$

for G_i a Borelset in $\overline{L_k\times S}$, i=1,...,n, n≥2.

(2.3.14) $$P_{(V,R)}^{(k)}(G_1\times\ldots\times G_n\times\overline{L_k\times S}\times\ldots) = P_{k,(V,R)}^{(n)}(G_1\times\ldots\times G_n)$$

for n≥1 and Borelsets G_i of $\overline{L_k\times S}$, i=1,...,n.

THEOREM 2.3.10. By (2.3.14) a unique probability measure $P^{(k)}_{(V,R)}$ on $(\overline{L_k \times S})^{\infty}$ is defined for which $P^{(k)}_{(V,R)}(J_k[0,\infty))=1$.

PROOF. The first part follows from the theorem of Ionescu Tulcea, while the second assertion is a consequence of the construction of $P^{(n)}_{k,(V,R)}$, $n \geq 1$. □

We conclude this section with a theorem which is useful in verifying the conditions of the main theorem 2.4.3. in the next section.

THEOREM 2.3.11. Let (V,R) be a policy for a CTMDP and $(V_k)^{\infty}_{k=1}$ a sequence of closed subsets of S. If

$$(2.3.15) \qquad P^{(n)}_{(V,R)}\{z=(t_j,i_j)^n_{j=1}: \tau^{(k)}_{V_k}(.) \xrightarrow{c} \tau_V(.) \text{ at } i_n\} = 1$$

then

$$(2.3.16) \qquad g^{(k)}_{V_k}(.) \xrightarrow{c} g_V(.), \quad P^{(n)}_{(V,R)}\text{-a.e.}$$

PROOF. Since $P_{(V,R)}$ is concentrated on $J[0,\infty)$ we know that $P^{(n)}_{(V,R)}\{z=(t_j,i_j)^n_{j=1}: t_n=t_{n-1}=t_n+\tau_V(i_n)\}=0$. Hence for $z=(t_j,i_j)^n_{j=1}\in J_n$

$$g_V(z) = \begin{cases} (t_1,i_1,\dots,t_n,i_n,\infty,\zeta,\dots) & \text{if } i_n=\zeta \text{ or } \tau_V(i_n)=\infty \\ (t_1,i_1,\dots,t_n,i_n,t_n+\tau_V(i_n),i_n,\infty,\zeta,\dots) & \text{otherwise,} \end{cases}$$

while

$$g^{(k)}_{V_k}(z) = \begin{cases} (t_1,i_1,\dots,t_n,i_n,\infty,\zeta,\dots) & \text{if } i_n=\zeta \text{ or } \tau^{(k)}_{V_k}(i_n)=\infty \\ (t_1,i_1,\dots,t_n,i_n,t_n+\tau^{(k)}_{V_k}(i_n),i_n,\infty,\zeta,\dots) & \text{otherwise.} \end{cases}$$

Hence (2.3.16) follows by assumption (2.3.15). □

2.4. DISCRETIZATION AND WEAK CONVERGENCE.

In this section first we construct for a given CTMDP a sequence of approximating DTMDP's with time parameter k approaching infinity. Next we provide sufficient conditions on the relation between a policy (V,R) for the CTMDP and policies $(V^{(k)},R^{(k)})$ for the approximating DTMDP's in order that the induced probability laws $P^{(k)}_{(V^{(k)},R^{(k)})}$ converge weakly to the law $P_{(V,R)}$ on $J[0,\infty)$.
This kind of limiting result is extremely useful in analyzing a CTMDP. For example a lot is known in literature about the structure of optimal policies for certain classes of DTMDP's. The standard way in which this kind of results is obtained for DTMDP's is the method of induction to the number of decision epochs in finite horizon problems and passing to limits when the time horizon is infinite. Obviously this procedure does not work for CTMDP's. However, the approximation result of this section provides the key to carry over structural results from a DTMDP to a CTMDP.

<u>DEFINITION 2.4.1.</u> Let $(S,A_1,A_2,q,\Pi,p,c_1,c_2,f)$ be a CTMDP. For $k\geq 1$ the *k th approximating DTMDP* is defined by $(S,A_1,A_2,p_1^{(k)},p_2^{(k)},c_1^{(k)},c_2^{(k)},f,k)$, where

(i) $$p_1^{(k)}(s,a,B) := k^{-1}q(s,a)\Pi(s,a,B)+1_B(s)(1-k^{-1}q(s,a)),\quad (s,a)\in S\times A_1$$

(ii) $$p_2^{(k)}(s,a,B) := p(s,a,B) \quad \text{for } (s,a)\in S\times A_2.$$

(iii) $$c_1^{(k)}(s,a) := k^{-1}c_1(s,a) \quad \text{for } (s,a)\in S\times A_1.$$

(iv) $$c_2^{(k)}(s,a) := c_2(s,a) \quad \text{for } (s,a)\in S\times A_2.$$

<u>REMARK 2.4.2.</u> (i) From the definition of a DTMDP (definition 2.3.1.) it is clear that the discrete time processes defined above are well defined DTMDP's for all $k>\sup\{q(s,a):(s,a)\in S\times A_1\}$.
(ii) Suppose that

(2.4.1) $$\Pi(s,a,\{s\}) = 0 \quad \text{for all } (s,a)\in S\times A_1.$$

Then the one step transition distribution $p_1^{(k)}$ of the k-th approximating DTMDP trivially satisfies:

$$(2.4.2)\qquad \begin{cases} p_1^{(k)}(s,a,B)=k^{-1}q(s,a)\Pi(s,a,B)+k^{-1}o(1) & \text{if } s\notin B \\ p_1^{(k)}(s,a,\{s\})=1-k^{-1}q(s,a)+k^{-1}o(1), & \end{cases}$$

where o(1) represents a (bounded) measurable function on $S\times A_1$, which may depend on s, a, k and B and which converges uniformly in (s,a,B) to zero on compact subsets of $S\times A_1$ as $k\to\infty$. Since representation (2.4.2) will be used in the proof of our theorems, we can replace the strict definition of $p_1^{(k)}$ in definition 2.4.1 by (2.4.2). Also the definitions of $p_2^{(k)}$, $c_1^{(k)}$ and $c_2^{(k)}$ in definition 2.4.1 can be relaxed to

$$\begin{aligned} p_2^{(k)}(s,a,B) &:= p(s,a,B) + o(1) \\ c_1^{(k)}(s,a) &:= k^{-1}c_1(s,a) + k^{-1}o(1) \\ c_2^{(k)}(s,a) &:= c_2(s,a) + o(1) \end{aligned}$$

Next theorem gives sufficient conditions for the weak convergence of $P^{(k)}_{(V^{(k)},R^{(k)})}$ to $P_{(V,R)}$, where (V,R) and $(V^{(k)},R^{(k)})$ are policies for a CTMDP and the k-th approximating DTMDP respectively. In section 2.5. we will illustrate the rather technical conditions of this theorem with some examples.

<u>THEOREM 2.4.3</u>. Let $(S,A_1,A_2,q,\Pi,p,c_1,c_2,f)$ be a CTMDP for which assumption 2.2.2. and (2.4.1) hold. Let P_0 be an initial distribution on S, (V,R) a policy for the CTMDP with $R=(R_1,R_2)$ and $(V^{(k)},R^{(k)})$ a policy for the k-th approximating DTMDP with $R^{(k)}=(R_1^{(k)},R_2^{(k)})$, $k\geq 1$. Then

$$P^{(k)}_{(V^{(k)},R^{(k)})} \xrightarrow{w} P_{(V,R)} \qquad \text{as } k\to\infty$$

if the following conditions are satisfied:

$$(2.4.3)\qquad P^{(n)}_{(V,R)}\{z=(t_j,i_j)_{j=1}^n : \tau^{(k)}_{V^{(k)}}(.) \xrightarrow{c} \tau_V(.) \text{ at } i_n\} = 1, \qquad n\geq 1.$$

(2.4.4) $\quad P^{(n)}_{(V,R)}\{z=(t_j,i_j)^n_{j=1}:\ R_2^{(k)}(g^{(k)}_{V^{(k)}}(.),.) \overset{c}{\rightarrow} R_2(g_V(.),.)$

at $(z,t_n+\tau_V(i_n))\} = 1, \quad n \geq 1.$

(2.4.5) $\quad (P^{(n)}_{(V,R)}\times\lambda)\{(z,t):\ z=(t_j,i_j)^n_{j=1},\ t \geq t_n,$

$R_1^{(k)}(g^{(k)}_{V^{(k)}}(.),.) \overset{c}{\not\rightarrow} R_1(g_V(.),.)$ at $(z,t)\} = 0, \quad n \geq 1.$

Here λ denotes the Lebesgue measure on $[0,\infty)$, while $h_k(.)\overset{c}{\not\rightarrow}h(.)$ at x means that $h_k(.)$ is not continuously convergent to $h(.)$ at x.

PROOF. In the proof of this theorem we put for abbreviation:

$$P_k:=P^{(k)}_{(V^{(k)},R^{(k)})},\ P:=P_{(V,R)},\ P_k^{(n)}:=P^{(n)}_{k,(V^{(k)},R^{(k)})} \text{ and } P^{(n)}:=P^{(n)}_{(V,R)}.$$

According to theorem 1.6.1. it is sufficient to show that

(2.4.6) $\quad P_k(F) \rightarrow P(F)$, as $k \rightarrow \infty$

for all $F = \bigcap_{n=1}^{m} T_n^{-1}(B_n) \cap S_n^{-1}(F_n) \in \mathcal{F}$ with $P(\delta F) = 0$.

We prove (2.4.6) by induction on m. For m = 1 we have by (2.3.12), (2.3.14), (2.2.12) and (2.2.14)

$$P_k(T_1 \in B,\ S_1 \in F) = P_k^{(1)}(B \times F) = 1_B(0)\ .\ P_0(F) =$$

$$= P^{(1)}(B \times F) = P(T_1 \in B,\ S_1 \in F), \text{ for } k \geq 1.$$

Suppose that (2.4.6) is true for all $F = \bigcap_{n=1}^{m} (T_n^{-1}(B_n) \cap S_n^{-1}(F_n)) \in \mathcal{F}$ with $P(\delta F) = 0$ for some fixed m.
From the induction hypothesis follows that

(2.4.7) $\quad P_k^{(m)} \overset{w}{\rightarrow} P^{(m)} \quad$ on $J_m \subset (\overline{[0,\infty)\times S},\tilde{\rho})^m$

Now choose $F_0 := \bigcap_{n=1}^{m} T_n^{-1}(B_n) \cap S_n^{-1}(F_n) \cap T_{m+1}^{-1}(B) \cap S_{m+1}^{-1}(F) \in \mathcal{F}$ with

(2.4.8) $P(\delta F_0) = 0.$

Note that by the definition of F the sets B_i are finite intervals for $1 \le i \le m$. From (2.3.13) and (2.2.13) follows that

$$P_k(F_0) = P_k^{(m+1)}(F_0) = \int_{B_1\times F_1\times\ldots\times B_m\times F_m} Q_k^{(m)}(z)(B\times F)\,dP_k^{(m)}(z)$$

and

$$P(F_0) = \int_{B_1\times F_1\times\ldots\times B_m\times F_m} Q^{(m)}(z)(B\times F)\,dP^{(m)}(z).$$

Combining (2.4.7) with corollary 1.2.21. we conclude that F_0 satisfies (2.4.6) if

(2.4.9) $P^{(m)}(\delta(B_1\times F_1\times\ldots\times B_m\times F_m)\cap\{z:Q^{(m)}(z)(B\times F)>0\}) = 0$

and

(2.4.10) $Q_k^{(m)}(.)(B\times F) \overset{c}{\to} Q^{(m)}(.)(B\times F)$, $P^{(m)}$-a.e. on $\overline{B_1\times F_1\times\ldots\times B_m\times F_m}$.

Condition (2.4.9) follows from (2.4.8). To prove (2.4.10) we construct an exception subset E of $\overline{B_1\times F_1\times\ldots\times B_m\times F_m}$ with $P^{(m)}(E)=0$. Define subsets E_i of $\overline{B_1\times F_1\times\ldots\times B_m\times F_m}$, $i=1,\ldots,6$ as follows:

$$E_1 := \{z=(t_j,i_j)_{j=1}^m : \tau^{(k)}_{V^{(k)}}(.) \overset{c}{\not\to} \tau_V(.) \text{ at } i_m\}$$

$$E_2 := \{z=(t_j,i_j)_{j=1}^m : R_2^{(k)}(g^{(k)}_{V^{(k)}}(.),.) \overset{c}{\not\to} R_2(g_V(.),.)$$
$$\text{at } (z,t_m+\tau_V(i_m))\}$$

$$E_3 := \{z=(t_j,i_j)_{j=1}^m : \lambda\{(t\ge t_m : R_1^{(k)}(g^{(k)}_{V^{(k)}}(.),.) \overset{c}{\not\to} R_1(g_V(.),.)$$
$$\text{at } (z,t)\}>0\}.$$

$$E_4:=\{z=(t_j,i_j)_{j=1}^m:\ t_m+\tau_V(i_m)\in\delta B;$$
$$\int_{A_2} p(\sigma_V(i_m),\ a,F)\ dR_2(g_V(z),\ t_m+\tau_V(i_m))(a)>0\}.$$

$$E_5:=\{z=(t_j,i_j)_{j=1}^m:\ t_m+\tau_V(i_m)\in B;$$
$$\int_{A_2} p(\sigma_V(i_m),\ a,\delta F)\ dR_2(g_V(z),t_m+\tau_V(i_m))(a)>0\}.$$

$$E_6:=\{z=(t_j,i_j)_{j=1}^m:\ \lambda\{t:\ t\le\tau_V(i_m);\ t+t_m\in B;$$
$$\int_{A_1} q(f(i_m,t),a)\Pi(f(i_m,t),a,\delta F)\,dR_1(g_V(z),t_m+t)(a)>0\}>0\}.$$

We conclude from (2.4.3), (2.4.4) and (2.4.5) that $P^{(m)}(E_1\cup E_2\cup E_3)=0$. Since (2.4.8) implies that $P^{(m+1)}(\overline{B_1\times F_1\times\ldots\times B_m\times F_m}\times\delta B\times\bar{F})=0$ and $P^{(m+1)}(\overline{B_1\times F_1\times\ldots\times B_m\times F_m}\times\bar{B}\times\delta F)=0$ it follows with (2.2.6) and (2.2.7) that $P^{(m)}(E_4\cup E_5\cup E_6)=0$. Put

(2.4.11) $\qquad E:=\bigcup_{i=1}^{6} E_i,$

then $P^{(m)}(E)=0$.

Hence (2.4.10) holds if

(2.4.12) $\qquad Q_k^{(m)}(.)(B\times F)\ \subsetneq\ Q^{(m)}(.)(B\times F)$ on $\overline{B_1\times F_1\times\ldots\times B_m\times F_m}\cap E^c$.

From the definition of $\mathcal{F}$ follows that either B is a finite interval and $F\in\mathcal{S}$ òr B is an infinite interval including ∞ and $F=S^+$. We restrict ourselves to the first case (the proof in the second case is simpler). We put without loss of generality $B=[0,t]$.

Choose $z:=(t_j,i_j)_{j=1}^m\in\overline{B_1\times F_1\times\ldots\times B_m\times F_m}\cap E^c$ and let $z(k)=(\ell_j(k)k^{-1},i_j(k))_{j=1}^m\in J_m(k)$, $k\ge 1$ such that $z(k)\to z$. Assume that $t\ge t_m$. By replacing n by m in formula (2.2.6) we have an explicit expression for $Q^{(m)}(z)(B\times F)$.

By the definition of the k-th approximating DTMDP combined with remark 2.4.4. and formula (2.3.7) we have (with suppression of the index $V^{(k)}$)

$$(2.4.13)\qquad Q_k^{(m)}(z(k))(B\times F) = \sum_{\ell=1}^{([tk]-\ell_m(k))\wedge(\tau^{(k)}(i_m(k))k-1)} \prod_{n=1}^{\ell-1}$$

$$\int_{A_1}\{1-k^{-1}q(f(i_m(k),nk^{-1}),a)+k^{-1}o(1)\}dR_1^{(k)}(g^{(k)}(z(k)),(n+\ell_m(k))k^{-1})(a)\,.$$

$$\cdot\int_{A_1}\{k^{-1}q(f(i_m(k),\ell k^{-1}),a)\Pi(f(i_m(k),\ell k^{-1}),a,F)+k^{-1}o(1)\}$$

$$dR_1^{(k)}(g^{(k)}(z(k)),(\ell+\ell_m(k))k^{-1})(a)\ +$$

$$+\ 1_{\{\tau^{(k)}(i)k\le[tk]-\ell_m(k)\}}(i_m(k))\ .$$

$$\cdot\prod_{n=1}^{\tau^{(k)}(i_m(k))k-1}\int_{A_1}\{1-k^{-1}q(f(i_m(k),nk^{-1}),a)+k^{-1}o(1)\}$$

$$dR_1^{(k)}(g^{(k)}(z(k)),(n+\ell_m(k))k^{-1})(a)\ .$$

$$\cdot\int_{A_2}\{p(\sigma^{(k)}(i_m(k)),a,F)+o(1)\}dR_2^{(k)}(g^{(k)}(z(k)),\tau^{(k)}(i_m(k))+\ell_m(k)k^{-1})(a)\,.$$

Since $z\notin E_1$ we know that $\tau_{V_k}^{(k)}(i_m(k))\to\tau_V(i_m)$.

Hence the first term of the right hand side of (2.4.13) is for $k\to\infty$ asymptotically equivalent to

$$\int_{s=0}^{(t-t_m)\wedge\tau_V(i_m)}\prod_{n=1}^{[sk]-1}\int_{A_1}\{1-k^{-1}q(f(i_m(k),nk^{-1}),a)+k^{-1}o(1)\}$$

$$dR_1^{(k)}(g^{(k)}(z(k)),(n+\ell_m(k))k^{-1})(a)\,.$$

$$\cdot\int_{A_1}\{q(f(i_m(k),[sk]k^{-1}),a)\Pi(f(i_m(k),[sk]k^{-1}),a,F)+o(1)\}$$

$$dR_1^{(k)}(g^{(k)}(z(k)),([sk]+\ell_m(k))k^{-1})(a)ds.$$

From lemma's 2.4.4. and 2.4.5. below follows that the first term of (2.4.13) converges as $k\to\infty$ to the first term of (2.2.6).
The convergence of the second factor of the second term of (2.4.13) to the corresponding factor of (2.2.6) follows also from lemma 2.4.4.
Finally the convergence of the product of the first and third factor of the second term of (2.4.13) to the product of the corresponding factors of the second term of (2.2.6) is established in lemma 2.4.6.
This completes the proof of theorem 2.4.3. □

LEMMA 2.4.4 . Let E be defined by (2.4.11) in the proof of theorem 2.4.3. If $z = (t_j,i_j)_{j=1}^m \in E^c$ and $z_k = (\ell_j(k)k^{-1},i_j(k))_{j=1}^m \in J_m(k)$ such that $z_k \to z$ and if $s_k \to s \geq 0$ then

$$(2.4.14)\qquad \prod_{n=1}^{[s_k k]-1} \int_{A_1} \{1-k^{-1}q(f(i_m(k),nk^{-1}),a)+k^{-1}o(1)\}\, dR_1^{(k)}(g^{(k)}(z_k),(n+\ell_m(k))k^{-1})(a)$$

converges as $k\to\infty$ to

$$\exp(-\int_{u=0}^{s}\int_{A_1} q(f(i_m,u),a)\ dR_1(g(z),t_m+u)(a)du).$$

(Here o(1) represents a bounded measurable function on $S\times A_1$, depending on k, which converges uniformly to zero on compact subsets of $S\times A_1$).

PROOF. The natural logarithm of (2.4.14) is equal to

$$(2.4.15)\qquad \sum_{n=1}^{[s_k k]-1} \log(1-k^{-1}\int_{A_1}\{q(f(i_m(k),nk^{-1}),a)+o(1)\})\, dR_1^{(k)}(g^{(k)}(z_k),(n+\ell_m(k))k^{-1})(a).$$

Since $-x-x^2 \leq \log(1-x) \leq -x$ for $0\leq x<\frac{1}{2}$ we conclude that (2.4.15) is asymptotically equivalent to

$$-\int_{u=0}^{s}\int_{A_1}\{q(f(i_m(k),[uk]k^{-1}),a)+o(1)\}dR_1^{(k)}(g^{(k)}(z_k),([uk]+\ell_m(k))k^{-1})(a)du.$$

Note that by the continuity of $f(.,.)$ and $q(.,.)$

$$q(f(i_m(k),[uk]k^{-1}),.)+o(1) \overset{c}{\to} q(f(i_m,u),.) \qquad \text{on } A_1.$$

Since $z \notin E_3$ it follows that

$$R_1^{(k)}(g^{(k)}(z_k),([uk]+\ell_m(k))k^{-1}) \overset{w}{\to} R_1(g_V(z),t_m+u)$$

for λ-almost all u.
The lemma follows from corollary 1.2.21. □

<u>LEMMA 2.4.5.</u> Let z, s, z_k, s_k, $k\geq 1$ be as in the previous lemma. Then

$$\int_{A_1} \{q(f(i_m(k),[s_k k]k^{-1}),a)\Pi(f(i_m(k),[s_k k]k^{-1}),a,F) + o(1)\} \, dR_1^{(k)}(g^{(k)}(z_k),([s_k k]+\ell_m(k))k^{-1})(a)$$

converges as $k \to \infty$ to

$$\int_{A_1} q(f(i_m,s),a)\Pi(f(i_m,s),a,F) \; dR_1(g_V(z),t_m+s)(a)$$

for λ-almost all $s \in [0,t-t_m] \cap [0,\tau_V(i_m)]$.

<u>PROOF</u>. Since $z \notin E_3$ we have

$$R_1^{(k)}(g^{(k)}(z_k),([s_k k]+\ell_m(k))k^{-1}) \overset{w}{\to} R(g_V(z),t_m+s)$$

for λ-almost all s.
Since $z \notin E_6$ we know that for λ-almost all $s \in [0,t-t_m] \cap [0,\tau_V(i_m)]$

$$\int_{A_1} q(f(i_m,s),a)\Pi(f(i_m,s),a,\delta F) \; dR_1(g_V(z),t_m+s)(a) = 0.$$

Moreover, we have by the continuity of $f(.,.)$, $q(.,.)$ and $\Pi(.,.)$ for all $s\geq 0$ and $a \in A_1$ for which $\Pi(f(i_m,s),a,\delta F) = 0$,

$$q(f(i_m(k),[s_k k]k^{-1}),.)\Pi(f(i_m(k),[s_k k]k^{-1}),.,F) + o(1) \overset{c}{\to}$$

$$q(f(i_m,s),.)\Pi(f(i_m,s),.,F) \qquad \text{at } a \in A_1.$$

Application of corollary 1.2.21. yields the lemma. □

LEMMA 2.4.6. Let z, z_k, $k\geq 1$ be as in lemma 2.4.4. Then

$$(2.4.16)\qquad 1_{\{\tau^{(k)}(i)k\leq[tk]-\ell_m(k)\}}(i_m(k))\ .$$

$$\cdot\int_{A_2}\{p(\sigma^{(k)}(i_m(k)),a,F)+o(1)\}dR_2^{(k)}(g^{(k)}(z_k),\tau^{(k)}(i_m(k))+\ell_m(k)k^{-1})(a)$$

converges as $k\to\infty$ to

$$(2.4.17)\qquad 1_{\{\tau_V(i)\leq t-t_m\}}(i_m)\cdot\int_{A_2} p(\sigma_V(i_m),a,F)\,dR_2(g_V(z),t_m+\tau_V(i_m))(a).$$

PROOF. Since $z\notin E_1$ the proof is immediate if $\tau_V(i_m)>t-t_m$. Therefore we assume $\tau_V(i_m)\leq t-t_m$. Since $z\notin E_1\cup E_2$ it follows that

$$R_2^{(k)}(g^{(k)}(z_k),\tau^{(k)}(i_m(k))+\ell_m(k)k^{-1})\overset{w}{\to}R_2(g_Vz),t_m+\tau_V(i_m)).$$

Since $z\notin E_5$ it follows that

$$\int_{A_2} p(\sigma_V(i_m),a,\delta F)\,dR_2(g_V(z),t_m+\tau_V(i_m))(a) = 0.$$

Moreover, since $z\notin E_1$ it follows that $\sigma^{(k)}(i_m(k))\to\sigma_V(i_m)$. Hence we have by the continuity of $p(.,.)$ for all $a\in A_2$ for which $p(\sigma_V(i_m),a,\delta F) = 0$:

$$p(\sigma^{(k)}(i_m(k)),.,F) + o(1)\overset{c}{\to}p(\sigma_V(i_m),.,F)\quad \text{at } a\in A_2.$$

Hence, by corollary 1.2.21., the second factor of (2.4.16) converges to the second factor of (2.4.17).
If $\tau_V(i)=t-t_m$ the proof is completed by noticing that $z\notin E_4$. If $\tau_V(i)<t-t_m$ the first factor of (2.4.16) converges to the first factor of (2.4.17), which also completes the proof. □

This section will be closed with a definition of a *regular* policy.

DEFINITION 2.4.7. Let $(S,A_1,A_2,q,\Pi,p,c_1,c_2,f)$ be a CTMDP and P_0 an initial distribution on S. A policy (V,R) is called *regular* if for all $k\geq 1$ there exists a policy (V_k,R_k) for the k-th approximating DTMDP, such that $P^{(k)}_{(V_k,R_k)}\overset{w}{\to}P_{(V,R)}$.

2.5. EXAMPLES.

In this section we treat two specific CTMDP's in order to give an impression about the difficulty to verify whether or not the conditions of theorem 2.4.3. hold.

EXAMPLE 2.5.1. Consider the *replacement model* of example 1.3.1. with some refinements. Assume that besides replacement the decision maker has control on the arrival rate ν of the shocks. Suppose that ν can be varied within the interval $[\nu_1,\nu_2]$ with $0<\nu_1\leq\nu_2<\infty$. The amount of damage caused by a single shock is a random variable with probability distribution F, with $F(0)=0$. Besides the damage caused by shocks the device decays continuously at rate $\sigma\geq 0$, i.e. between shocks the cumulative damage increases linearly at rate σ. This replacement model can be formulated as a CTMDP $(S,A_1,A_2,q,\Pi,p,c_1,c_2,f)$ where the state of the system denotes the total damage incurred by the device under operation, the controls are the possible arrival rates for the shock arrival process and the only impulsive control available is replacement, which brings the state of the system back to zero. We denote the replacement action with o.
Put

$$\begin{aligned}
&S: = [0,\infty)\\
&A_1: = [\nu_1,\nu_2]\\
&A_2: = \{o\}\\
(2.5.1)\qquad &q(s,\nu): = \nu \quad \text{for } (s,\nu)\in S\times A_1\\
&\Pi(s,\nu,[s,s+t]): = F(t) \quad \text{for } (s,\nu)\in S\times A_1 \text{ and } t\geq 0\\
&p(s,o,\{0\}): = 1 \quad \text{for } s\in S\\
&f(s,t): = s+\sigma t \quad \text{for } (s,t)\in S\times[0,\infty).
\end{aligned}$$

The cost rate $c_1(.,.)$ and the lump cost $c_2(.,.)$ are not specified, since they play no role in theorem 2.4.3.
One easily checks that assumption 2.2.2. holds. The k-th approximating decision process is given by $(S,A_1,A_2,p_1^{(k)},p_2^{(k)},c_1^{(k)},c_2^{(k)},f,k)$, where

$$(2.5.2)\qquad p_1^{(k)}(s,\nu,\{s\}) = 1-k^{-1}\nu, \quad (s,\nu)\in S\times A_1$$

(2.5.3) $\quad p_1^{(k)}(s,\nu,[s,s+t]) = 1-k^{-1}\nu+k^{-1}\nu F(t), \quad (s,\nu)\in S\times A_1,\ t\geq 0$

(2.5.4) $\quad p_2^{(k)}(s,0,\{0\}) = 1, \quad s\in S.$

Choose a policy (V,R) for the CTMDP where $V:=[D,\infty)$, with $D>0$ and $R=(R_1,R_2)$, with R_i a non-anticipating transition probability from $J[0,\infty)\times[0,\infty)$ to A_i. Since D>0 it follows that part (ii) of definition 2.2.3. is fulfilled (choose $\delta=D\sigma^{-1}$).

Consider for the k-th approximating decision process the policy (V_k,R), where $V_k:=[D_k,\infty)$ with $D_k>0$, $k\geq 1$.

THEOREM 2.5.2. Consider $(S,A_1,A_2,q,\Pi,p,c_1,c_2,f)$ defined by (2.5.1). For the policies (V,R) and (V_k,R) defined above and any initial distribution P_0 on S the conditions of theorem 2.4.3 hold if:

(2.5.5) $\quad D_k \to D \quad \text{as } k\to\infty$

(2.5.6) $\quad (P_{(V,R)}\times\lambda)(\mathrm{Disc}(R_1)) = 0$

and if $\sigma=0$

(2.5.7) $\quad P^{(n)}_{(V,R)}\{z=(t_j,i_j)_{j=1}^n : i_n=D\} = 0, \quad n\geq 1.$

PROOF. Since F(0)=0 condition (2.4.1) holds. Condition (2.4.4) follows from the fact that $A_2=\{0\}$ and hence $R_2(x,t)$ is degenerated in $\{0\}$ for all $(x,t)\in J[0,\infty)\times[0,\infty)$.
Next we note that $\tau_V(i)=\sigma^{-1}(D-i)^+$ and $\sigma^{-1}(D_k-i)^+\leq\tau^{(k)}_{V_k}(i)\leq\sigma^{-1}(D_k-i)^++k^{-1}$, which implies with (2.5.5) condition (2.4.3) for $\sigma>0$. When $\sigma=0$ we have $\tau_V(i)=\tau_{V_k}(i)=0$ if $i>D\vee D_k$ and $\tau_V(i)=\tau_{V_k}(i)=\infty$ when $i<D\wedge D_k$ and hence (2.4.3) follows from (2.5.5) and (2.5.7). Moreover, with theorem 2.3.11 follows that $g^{(k)}_{V_k}(.) \overset{c}{\to} g_V(.)$, $P^{(n)}_{(V,R)}$-a.e.. Hence (2.5.6) implies (2.4.5) □

EXAMPLE 2.5.3. Consider the *inventory model* of example 1.3.3. This inventory model can be formulated as a CTMDP, in which the state of the system denotes the stock on hand, where a negative value of the state variable indicates a backlog. The only control available is the fixed arrival rate ν, while the impulsive controls are the different order sizes. Choosing the impulsive control a means that an amount a is ordered. We put for this example

$$
\begin{aligned}
&S: = \mathbb{R} \\
&A_1: = \{\nu\} \\
&A_2: = [0,\infty) \\
&q(s,\nu): = \nu \quad \text{for } s\in S \\
&\Pi(s,\nu,[s,s-t]): = F(t) \quad \text{for } s\in S \text{ and } t\geq 0 \\
&p(s,a,\{s+a\}): = 1 \quad \text{for } (s,a)\in S\times A_2 \\
&f(s,t): = s-\sigma t \quad \text{for } (s,t)\in S\times[0,\infty)
\end{aligned}
\tag{2.5.8}
$$

Here the demands of the customers are assumed to be distributed according to the distribution function F, with F(0)=0. Again the cost rate $c_1(.,.)$ and lump cost $c_2(.,.)$ are not specified.
For k≥1 the k-th approximating decision process is given by $(S,A_1,A_2,p_1^{(k)},p_2^{(k)},c_1^{(k)},c_2^{(k)},f,k)$ with

$$p_1^{(k)}(s,\nu,\{s\}) = 1-k^{-1}\nu, \quad s\in S. \tag{2.5.9}$$

$$p_1^{(k)}(s,\nu,[s,s-t]) = 1-k^{-1}\nu+k^{-1}\nu F(t), \quad s\in S,\ t\geq 0. \tag{2.5.10}$$

$$p_2^{(k)}(s,a,\{s+a\}) = 1, \quad (s,a)\in S\times A_2. \tag{2.5.11}$$

Suppose the CTMDP is controlled by the (s^*,S^*)-policy (V,R) i.e. $V=(-\infty,s^*]$ and $R=(R_1,R_2)$ with

$$R_2(x,t): =[S^*-\pi_t x]^{+\,*)} \text{for } (x,t)\in J[0,\infty)\times[0,\infty),$$

where we assume that $S^*>s^*$.

For the k-th approximating DTMDP we define a policy $(V^{(k)},R^{(k)})$ with

*) $[x]^+:=\max(x,0)$.

$$V^{(k)} := (-\infty, s_k^*]$$

$$R_2^{(k)}(x,t) := S_k^* - \pi_t x \quad \text{for } (x,t) \in J_k [0,\infty) \times L_k.$$

THEOREM 2.5.4. Consider $(S, A_1, A_2, q, \Pi, p, c_1, c_2, f)$ defined by (2.5.8). For the policies (V,R) and $(V^{(k)}, R^{(k)})$ defined above and any initial distribution P_0 on S the conditions of theorem 2.4.3. hold if:

(2.5.12) $\quad S_k^* \to S^* \quad$ as $k \to \infty$

(2.5.13) $\quad s_k^* \to s^* \quad$ as $k \to \infty$.

PROOF. Since F(0)=0 condition (2.4.1) holds. Using (2.5.13) the validity of (2.4.3) can be shown as in the proof of theorem 2.5.2. Condition (2.4.5) is trivially fulfilled because $A_1 = \{\nu\}$. Finally to prove that condition (2.4.4) holds note that for $P_{(V,R)}$-almost all $z = (t_j, i_j)_{j=1}^n$ we have:

(2.5.14) $\quad R_2(g_V(z), t_n + \tau_V(i_n)) = S^* - f(i_n, \tau_V(i_n))$

and

(2.5.15) $\quad R_2^{(k)}(g_{V^{(k)}}^{(k)}(z), t_n + \tau_{V^{(k)}}^{(k)}(i_n)) = S_k^* - f(i_n, \tau_{V^{(k)}}^{(k)}(i_n)).$

From (2.5.13) follows that

(2.5.16) $\quad \tau_{V^{(k)}}^{(k)}(.) \xrightarrow{c} \tau_V(.).$

Combining (2.5.12), (2.5.16) with (2.5.14) and (2.5.15) completes the proof. □

CHAPTER 3

EQUIVALENT POLICIES

3.1. INTRODUCTION.

In this chapter we will use the convergence results from the previous chapter in order to transpose a well-known result for discrete time processes to CTMDP's. It was shown by DERMAN and STRAUCH (1966) that for every history-remembering policy for a discrete time Markov decision process one can construct a memoryless policy such that the one-dimensional marginal distributions of state and action are the same under both policies. This implies that under a cost structure which is uniquely determined by those one-dimensional marginal distributions one can restrict attention to the memoryless policies in the search for the optimal policy. However, as we will show in the sequel of this introduction, for a CTMDP most cost structures are not uniquely determined by the one dimensional marginal distributions of state and action. This implies that a continuous version of the Derman and Strauch theorem is less powerful in the CTMDP case as it is for a DTMDP. However, when we restrict our attention to CTMDP's in which no impulsive controls are available it turns out that a theorem of this kind can be applied in the CTMDP case like it is used for DTMDP's. For that reason we will restrict our attention in section 2 of this chapter to those CTMDP's for which the set of impulsive controls is empty. For ease of presentation we will also assume that the state space is countable.

In the previous chapter we introduced for any DTMDP and CTMDP under a fixed policy probability measures on the sequence space $J[0,\infty)$. When we want to compare two policies for a given decision process we have to take into account the cost functions (cost rate, lump cost or one step cost) of the process. With these cost functions we can define several cost functionals on $J[0,\infty)$. An important example of such a functional is the α-discounted cost functional (other functionals will be defined in chapter 4).

DEFINITION 3.1.1.

(i) Let $(S,A_1,A_2,p_1,p_2,c_1,c_2,f,k)$ be a DTMDP and (V,R) a policy for this process with $R=(R_1,R_2)$. Suppose that $c_i(.,.)$ is bounded from below, $i=1,2$. For $\alpha > 0$ the *α-discounted cost functional under* (V,R) is a function $c^{(k)}_{\alpha,(V,R)}$ on $J_k[0,\infty)$, defined for $x\in J_k[0,\infty)$ by

$$(3.1.1) \qquad c^{(k)}_{\alpha,(V,R)}(x) := \sum_{n=1}^{\infty} \alpha^n \int_{A_1} c_1(\pi_{nk^{-1}}x,a)\, dR_1(x,nk^{-1})(a) + \sum_{n=1}^{\infty} \alpha^n 1_V(\pi_{nk^{-1}}x) \int_{A_2} c_2(\pi_{nk^{-1}}x,a)\, dR_2(x,nk^{-1})(a)$$

(ii) Let $(S,A_1,A_2,q,\Pi,p,c_1,c_2,f)$ be a CTMDP and (V,R) a policy for this process with $R=(R_1,R_2)$. Suppose that $c_1(.,.)$ is bounded from below and $c_2(.,.)$ non-negative. For $\alpha > 0$ the *α-discounted cost functional under (V,R)* is a function $c_{\alpha,(V,R)}(.)$ on $J[0,\infty)$, defined by

$$(3.1.2) \qquad c_{\alpha,(V,R)}(x) := \int_0^{\infty} e^{-\alpha t} \int_{A_1} c_1(\pi_t x,a)\, dR_1(x,t)(a)dt + \sum_{n=1}^{\infty} e^{-\alpha\tau_n(x)} \int_{A_2} c_2(\pi_{\tau_n(x)}x,a)\, dR_2(x,\tau_n(x))(a),$$

where $\tau_n(x)$ represents the epoch of the n-th entrance of x into the set V.

DEFINITION 3.1.2. Let $(S,A_1,A_2,p_1,p_2,c_1,c_2,f,k)$ be a DTMDP and $(S,A_1,A_2,q,\Pi,p,c_1,c_2,f)$ a CTMDP and let (V,R) be a policy for one of these decision processes. Assume that $c_1(.,.)$ is bounded from below and $c_2(.,.)$ non-negative. For any initial distribution P_0 on S *the expected α-discounted costs under (V,R)* are defined for the DTMDP by

$$(3.1.3) \qquad c_k((V,R),\alpha) := \int_{J[0,\infty)} c^{(k)}_{\alpha,(V,R)}(x)\, dP^{(k)}_{(V,R)}(x)$$

and for the CTMDP by

$$(3.1.4) \qquad c((V,R),\alpha) := \int_{J[0,\infty)} c_{\alpha,(V,R)}(x)\, dP_{(V,R)}(x).$$

From Fubini's theorem and (3.1.1) respectively (3.1.2) we find

(3.1.5) $$c_k((V,R),\alpha) = \sum_{n=1}^{\infty} \alpha^n \int_{S\times P(A_1)} \int_{A_1} c_1(s,a)\,d\nu(a)\,dP^{(k)}_{(V,R)}(\pi_{nk^{-1}}, R_1(.,nk^{-1}))^{-1}(s,\nu) + \sum_{n=1}^{\infty} \alpha^n \int_{V\times P(A_2)} \int_{A_2} c_2(s,a)\,d\nu(a)\,dP^{(k)}_{(V,R)}(\pi_{nk^{-1}}, R_2(.,nk^{-1}))^{-1}(s,\nu)$$

and

(3.1.6) $$c((V,R),\alpha) = \int_0^{\infty} e^{-\alpha t} \int_{S\times P(A_1)} \int_{A_1} c_1(s,a)\,d\nu(a)\,dP_{(V,R)}(\pi_t, R_1(.,t))^{-1}(s,\nu)\,dt + \sum_{n=1}^{\infty} \int_{[0,\infty)\times S\times P(A_2)} e^{-\alpha t} \int_{A_2} c_2(s,a)\,d\nu(a)\,dP_{(V,R)}(\tau_n, \pi_{\tau_n}, R_2(.,\tau_n))^{-1}(t,s,\nu).$$

Comparing the expressions (3.1.5) and (3.1.6) reveals an interesting difference between CTMDP's and DTMDP's. From (3.1.5) follows that for the DTMDP the expected α-discounted costs under policy (V,R) are uniquely determined by the one dimensional marginal distributions of state and randomized action i.e. $c_k((V,R),\alpha)$ is determined by the sequence of probability measures on $S \times P(A_i)$ defined by

(3.1.7) $$P^{(k)}_{(V,R)}(\pi_{nk^{-1}}, R_i(.,nk^{-1}))^{-1}, \quad n\geq 1;\ i=1,2.$$

This assertion can be strengthened in the following sense: $c_k((V,R),\alpha)$ is uniquely determined by the sequence of probability measures $(\Theta_n^{(i)})_{n=1}^{\infty}$ on $S \times A_i$, defined by

(3.1.8) $$\Theta_n^{(i)}(B \times F) := \int_{B\times P(A_i)} \nu(F)\,dP^{(k)}_{(V,R)}(\pi_{nk^{-1}}, R_i(.,nk^{-1}))^{-1}(s,\nu)$$

for $B \in S$ and $F \in A_i$, i=1,2.

<u>LEMMA 3.1.3.</u> $\Theta_n^{(i)}$ is a well-defined probability measure on $S \times A_i$ for $n \geq 1$ and i=1,2 and the $(\Theta_n^{(i)})_{n=1}^{\infty}$, i=1,2 determine uniquely $c_k((V,R),\alpha)$.

PROOF. The proof of the first statement is straightforward and the second follows from the equality

$$(3.1.9)\qquad c_k((V,R),\alpha) = \sum_{n=1}^{\infty} \alpha^n \{ \int_{S\times A_1} c_1(s,a)\,d\Theta_n^{(1)}(s,a) + \int_{V\times A_2} c_2(s,a)\,d\Theta_n^{(2)}(s,a) \}.$$

To prove (3.1.9) we first consider the case where

$$c_i(s,a) = 1_{B_i\times F_i}(s,a), \quad \text{for some } B_i\in\mathcal{S} \text{ and } F_i\in\mathcal{A}_i, \quad i=1,2.$$

Then

$$c_k((V,R),\alpha) = \sum_{n=1}^{\infty} \alpha^n \int_{S\times P(A_1)} 1_{B_1}(s)\nu(F_1)\,dP^{(k)}_{(V,R)}(\pi_{nk^{-1}},R_1(.,nk^{-1}))^{-1}(s,\nu) +$$

$$+ \sum_{n=1}^{\infty} \alpha^n \int_{V\times P(A_2)} 1_{B_2}(s)\nu(F_2)\,dP^{(k)}_{(V,R)}(\pi_{nk^{-1}},R_2(.,nk^{-1}))^{-1}(s,\nu) =$$

$$= \sum_{n=1}^{\infty} \alpha^n \Theta_n^{(1)}(B_1\times F_1) + \sum_{n=1}^{\infty} \alpha^n \Theta_n^{(2)}(B_2\cap V\times F_2) =$$

$$= \sum_{n=1}^{\infty} \alpha^n \{ \int_{S\times A_1} c_1(s,a)\,d\Theta_n^{(1)}(s,a) + \int_{V\times A_2} c_2(s,a)\,d\Theta_n^{(2)}(s,a) \}.$$

Since every non-negative measurable function can be approximated by a monotone sequence of step functions the lemma follows by application of the monotone convergence theorem. □

Note that the probability measure $\Theta_n^{(i)}$ can be interpreted as the simultaneous distribution of state and (impulsive) control at epoch nk^{-1}, $n\geq 1$. Lemma 3.1.3 tells us that for a DTMDP under policy (V,R) knowledge of the probability distributions (3.1.8) is enough to compute $c_k((V,R),\alpha)$. However, from (3.1.6) we learn that for a CTMDP under policy (V,R) the one dimensional marginal distributions of state and randomized control *together with* the simultaneous distributions of the n-th entrance time into V, the entrance state and the randomized impulsive control uniquely determine $c((V,R),\alpha)$. In other words $c((V,R),\alpha)$ is determined by the probability measures

$$(3.1.10)\qquad P_{(V,R)}(\pi_t,R(.,t))^{-1} \quad , \; t\geq 0$$

and

$$(3.1.11) \qquad P_{(V,R)}(\tau_n, \pi_{\tau_n}, R_2(\cdot,\tau_n))^{-1}, \quad n \geq 1.$$

As for the DTMDP we can strengthen this assertion. For a CTMDP the expected α-discounted costs under policy (V,R) is uniquely determined by the class of probability measures $(\Psi_t)_{t\geq 0}$ on $S\times A_1$ and the sequence of defective probability measures $(\Phi_n)_{n=1}^{\infty}$ on $[0,\infty)\times S\times A_2$, defined respectively by

$$(3.1.12) \qquad \Psi_t(B \times F) := \int_{B \times P(A_1)} \nu(F)\, dP_{(V,R)}(\pi_t, R_1(\cdot,t))^{-1}(s,\nu)$$

for $B \in S$ and $F \in A_1$; $t \geq 0$

and

$$(3.1.13) \qquad \Phi_n(C \times B \times F) := \int_{C\times B\times P(A_2)} \nu(F)\, dP_{(V,R)}(\tau_n, \pi_{\tau_n}, R_2(\cdot,\tau_n))^{-1}(t,s,\nu)$$

for all Borelsets C in $[0,\infty)$, $B \in S$, $F \in A_2$; $n \geq 1$.

3.2. GENERALIZATION OF A THEOREM OF DERMAN AND STRAUCH.

DEFINITION 3.2.1.

(i) Let $(S,A_1,A_2,p_1,p_2,c_1,c_2,f,k)$ be a DTMDP and P_0 an initial distribution on S. Two policies (V_1,R_1) and (V_2,R_2) are called *equivalent w.r.t. P_0*, if they generate the same sequence of probability measures $(\Theta_n^{(i)})_{n=1}^{\infty}$ on $S \times A_i$, where $\Theta_n^{(i)}$ is defined by (3.1.8) for $n \geq 1$ and $i = 1,2$.

(ii) Let $(S,A_1,A_2,q,\Pi,p,c_1,c_2,f)$ be a CTMDP and P_0 an initial distribution on S. Two policies (V_1,R_1) and (V_2,R_2) are called *equivalent w.r.t. P_0*, if they generate the same class of probability measures $(\Psi_t)_{t\geq 0}$ on $S \times A_1$ and the same sequence of defective probability measures $(\Phi_n)_{n=1}^{\infty}$ on $[0,\infty) \times S \times A_2$, where Ψ_t is defined by (3.1.12) for $t \geq 0$ and Φ_n by (3.1.13) for $n \geq 1$.

A well known theorem by Derman and Strauch states that for a DTMDP one can construct for every policy (V,R) an equivalent policy which is memoryless. Using this result and our convergence theorem 2.4.3 of the previous chapter, we want to prove a similar result for CTMDP's. For every regular policy

(V,R) for the CTMDP there exists by definition a policy (V_k,R_k) for the k-th approximating DTMDP such that $P^{(k)}_{(V_k,R_k)} \overset{w}{\to} P_{(V,R)}$. From Derman and Strauch follows the existence of memoryless policies (V_k',R_k') such that the $\Theta_n^{(i)}$ measures induced by (V_k,R_k) and (V_k',R_k') are the same for all $n \geq 1$ and $i = 1,2$. Now the following question arises. If there exists a memoryless policy (V',R') for the CTMDP such that $P^{(k)}_{(V_k',R_k')} \to P_{(V',R')}$, are the policies (V,R) and (V',R') then equivalent?

Unfortunately, we are not able to prove this in general. The problem is that the equivalence of (V_k,R_k) and (V_k',R_k') does not yield information about the relation between the Φ_n measures generated by (V,R) and (V',R').

However, when we restrict our attention to CTMDP's with empty set of impulsive controls we have for any policy (V,R)

$$\Phi_n([0,\infty) \times S \times A_2) = 0, \quad n \geq 1,$$

which implies that the sequence of defective probability measures $(\Phi_n)_{n=1}^{\infty}$ is the same for all policies. In this case we are able to obtain equivalence results for a CTMDP by combining the theorem of Derman and Strauch with our convergence theorem. For ease of presentation we restrict our attention to a countable state space.

Hence we make in this chapter the following assumptions on the component parts $(S,A_1,A_2,q,\Pi,p,c_1,c_2,f)$ of a CTMDP.

<u>ASSUMPTION 3.2.2</u>. (i) S is countable with the discrete metric.
(ii) $A_2 = \emptyset$.

Note that by assumption 3.2.2 (i) the drift function f is necessarily equal to $f(i,t)=i$, $t \geq 0$, $i \in S$.

Moreover, by assumption 3.2.2 (ii) the transition probability p and the lump cost c_2 are undefined. Hence we will simplify in this chapter the notation of a CTMDP, a DTMDP and a policy.

<u>NOTATION 3.2.3</u>. In this chapter we use the following notation.

(i) A CTMDP is denoted by a five tuple (S,A,q,Π,c).

(ii) A DTMDP is denoted by a five typle (S,A,p,c,k).

(iii) A policy for a CTMDP and DTMDP is denoted by R, which represents a transition probability from $S \times A$ to S.

(iv) The probability measures induced by a policy R are denoted by $P_R^{(k)}$ for a DTMDP with time parameter k and by P_R for a CTMDP.

THEOREM 3.2.4. Let (S,A,p,c,k) be a DTMDP and P_0 an initial distribution on S. For any policy R there exists a memoryless policy R', which is equivalent to R w.r.t. P_0.

PROOF. For every $i \in S$ and $t \in L_k$, for which

(3.2.1) $$P_R(\{x: \pi_t x = i\}) > 0$$

we define a probability measure $Q(i,t)$ on $\mathcal{P}(A)$, by

$$Q(i,t)(C) := P_R(\{x: R(x,t) \in C\} \mid \{x: \pi_t x = i\})$$

for all Borelsets C in $\mathcal{P}(A)$.
Next we define a probability measure $r(i,t)$ on A, by

$$r(i,t)(F) := \int_{\mathcal{P}(A)} \nu(F)\ dQ(i,t)(\nu) \qquad \text{for } F \in \mathcal{A}.$$

For those $(i,t) \in S \times L_k$ for which (3.2.1) does not hold, we define

$$r(i,t) := \nu_0$$

for some arbitrary but fixed $\nu_0 \in \mathcal{P}(A)$.
Finally we define a function R' from $J_k[0,\infty) \times L_k$ into $\mathcal{P}(A)$, by

$$R'(x,t) := r(\pi_t x,t).$$

It is straightforward to check that R' is a well-defined memoryless policy for (S,A,p,c,k). We shall show that R and R' are equivalent w.r.t. P_0. By the definition of equivalent policies we have to show that

(3.2.2) $$\Theta_n = \Theta_n', \quad n \geq 1$$

where Θ_n and Θ'_n denote the probability measures $\Theta_n^{(1)}$ defined by (3.1.8) for policy R and R'. Note that

$$\Theta_n(\{i\} \times F) = \int_{\{i\} \times P(A)} \nu(F)\, dP_R^{(k)}(\pi_{nk^{-1}}, R(.,nk^{-1}))^{-1}(s,\nu) =$$

$$= r(i,nk^{-1})(F) \; . \; P_R^{(k)}\{\pi_{nk^{-1}} = i\},$$

while

$$\Theta'_n(\{i\} \times F) = r(i,nk^{-1})(F) \; . \; P_{R'}^{(k)}\{\pi_{nk^{-1}} = i\} .$$

Hence it is sufficient to prove that

$$(3.2.3) \qquad P_R^{(k)}\pi^{-1}_{nk^{-1}} = P_{R'}^{(k)}\pi^{-1}_{nk^{-1}}, \qquad n \geq 1.$$

We will prove (3.2.3) by induction on n.
For n = 1 we have

$$P_R^{(k)}\pi^{-1}_{nk^{-1}}(B) = P_0(B) = P_{R'}^{(k)}\pi^{-1}_{nk^{-1}}(B).$$

Assume that (3.2.3) holds for n = m. Then

$$(3.2.4) \qquad P_R^{(k)}(\{\pi_{(m+1)k^{-1}} = i\}) = \sum_{j\in S} P_R^{(k)}(\{\pi_{mk^{-1}} = j\}).$$

$$\int_{P(A)} P_R^{(k)}(\{\pi_{(m+1)k^{-1}}=i\} \,|\, \{\pi_{mk^{-1}}=j;\ R(.,mk^{-1})=\nu\})\ dQ(j,mk^{-1})(\nu) =$$

$$= \sum_{j\in S} \int_{P(A)} \int_A p(j,a,\{i\})\, d\nu(a)\, dQ(j,mk^{-1})(\nu) . P_R^{(k)}(\{\pi_{mk^{-1}}=j\}).$$

The last equality follows from (2.3.1). From (3.2.4) and the induction hypothesis we find

$$P_R^{(k)}(\{\pi_{(m+1)k^{-1}} = i\}) =$$

$$= \sum_{j\in S} P_{R'}^{(k)}(\{\pi_{mk^{-1}} = j\}) \int_{P(A)} \int_A p(j,a,\{i\})d\nu(a)dQ(j,mk^{-1})(\nu) =$$

$$= \sum_{j\in S} P_{R'}^{(k)}(\{\pi_{mk^{-1}} = j\}) \int_A p(j,a,\{i\})dr(j,mk^{-1})(a) =$$

$$= P_{R'}^{(k)}(\{\pi_{(m+1)k^{-1}} = i\}).$$

The last equality follows again from (2.3.1) and the last but one equality follows from

$$\int_A f(a)\ dr(i,t)(a) = \int_{P(A)} \int_A f(a)\ d\nu(a)\ dQ(i,t)(\nu),$$

which holds for every measurable real-valued function f on A. □

This theorem has been proved for the first time by DERMAN and STRAUCH (1966) for a finite state space. Generalization to the countable state space was given by HORDÍJK (1974). DERMAN and STRAUCH as well as HORDIJK considered policies for which the decision at time n may depend on all past states *and* actions, while in our approach there is only dependency on the past states.

THEOREM 3.2.5. Let (S,A,q,Π,c) be a CTMDP and P_0 an initial distribution on S. Assume that $\Pi(s,a\{s\}) = 0$ for all $(s,a)\in S\times A$. For every policy R for which

$$(3.2.5) \qquad (P_R \times \lambda)\{\mathrm{Disc}(R)\} = 0$$

there exists a regular memoryless policy R', which is equivalent to R w.r.t. P_0.

PROOF. Let R be a policy for which (3.2.5) holds and define policies R_k for the k-th approximating DTMDP by

$$R_k(x,t) := R(x,t), \quad \text{for } (x,t) \in J_k[0,\infty) \times L_k;\ k \geq 1.$$

From (3.2.5) follows with theorem 2.4.3. that

$$(3.2.6) \qquad P_{R_k} \xrightarrow{w} P_R.$$

From (2.2.6) and assumption 3.2.2.(ii) follows that

$$(3.2.7) \qquad P_{\tilde{R}}(\{x: T_n(x) = t\}) = 0, \quad \text{for } n \geq 1 \text{ and } t > 0$$

for every policy $\tilde{R}$.
Since $\delta\{x: \pi_t x = i\} \subset \bigcup_{n=1}^{\infty} \{x: T_n(x) = t\}$ for all $i \in S$ and $t > 0$ we conclude from (3.2.6) and (3.2.7) that

$$(3.2.8) \qquad P_{R_k}(\{x: \pi_t x = i\}) \to P_R(\{x: \pi_t x = i\}), \text{ for all } i \in S \text{ and } t \geq 0.$$

For $i \in S$ and $t \in L_k$ for which $P_{R_k}(\{x: \pi_t x = i\}) > 0$ we define a probability measure $Q_k(i,t)$ on $P(A)$ by

$$(3.2.9) \qquad Q_k(i,t)(C) := P_{R_k}(\{x: R_k(x,t) \in C\} \mid \{x: \pi_t x = i\})$$

for all Borelsets C in $P(A)$,
and a probability measure $r_k(i,t)$ on A by

$$(3.2.10) \qquad r_k(i,t)(F) := \int_{P(A)} \nu(F)\, dQ_k(i,t)(\nu), \quad \text{for } F \in A.$$

For those $(i,t) \in S \times L_k$ for which $P_{R_k}(\{x: \pi_t x = i\}) = 0$ we put

$$(3.2.11) \qquad r_k(i.t) := \nu_0$$

for some arbitrary but fixed $\nu_0 \in P(A)$.

Next we define memoryless policies R_k' for the k-th approximating DTMDP by

$$(3.2.12) \qquad R_k' \ (x,t) := r_k(\pi_t x, t) \qquad \text{for } (x,t) \in J_k[0,\infty) \times L_k.$$

From theorem 3.2.4 and its proof it follows that

$$(3.2.13) \qquad P_{R_k} \ (\{x: \pi_t x = i\}) = P_{R_k'} \ (\{x: \pi_t x = i\}), \text{ for } i \in S \text{ and } t \geq 0.$$

For all $(i,t) \in S \times [0,\infty)$ for which P_R $(\{x: \pi_t x = i\}) > 0$ we define a probability measure $Q(i,t)$ on $P(A)$ by

$$(3.2.14) \qquad Q(i,t)(C) := P_R \ (\{x: R\ (x,t) \in C\} \,|\, \{x: \pi_t x = i\}),$$

for all Borelsets C in $P(A)$ and a probability measure $r(i,t)$ on A by

$$(3.2.15) \qquad r(i,t)(F) := \int_{P(A)} \nu(F) \ dQ(i,t)(\nu),$$

while we put

$$(3.2.16) \qquad r(i,t) := \nu_0$$

for those $(i,t) \in S \times [0,\infty)$ for which P_R $(\{x: \pi_t x = i\}) = 0$. Finally we put

$$(3.2.17) \qquad R'(x,t) := r(\pi_t x, t), \qquad \text{for } (x,t) \in J[0,\infty) \times [0,\infty).$$

From lemma 3.2.7 below it follows that $R'(.,.)$ is a well-defined regular memoryless policy for (S,A,q,Π,c) and we prove in lemma 3.2.8 that

$$(3.2.18) \qquad P_{R_k'} \overset{w}{\to} P_{R'}.$$

From (3.2.7) and (3.2.18) it follows that

$$(3.2.19) \qquad P_{R_k'} \ (\{x: \pi_t x = i\}) \to P_{R'}(\{x: \pi_t x = i\}), \text{ for all } i \in S,\ t > 0.$$

Combination of (3.2.8), (3.2.13) and (3.2.19) yields

(3.2.20) $\quad P_R(\{x: \pi_t x = i\}) = P_{R'}(\{x: \pi_t x = i\})$, for all $i \in S$, $t \geq 0$.

Hence, (see formula (3.1.12) for definition of Ψ_t)

$$\Psi_t(\{i\} \times F) = \int_{\{i\} \times \mathcal{P}(A)} \nu(F)\, dP_R(\pi_t, R(.,t))^{-1}(s,\nu) =$$
$$= P_R(\{\pi_t = i\}) \cdot \int_{\mathcal{P}(A)} \nu(F)\, dQ(i,t)(\nu) =$$
$$= P_{R'}(\{\pi_t = i\}) \cdot r(i,t)(F) =$$
$$= \int_{\{i\} \times \mathcal{P}(A)} \nu(F)\, dP_{R'}(\pi_t, R'(.,t))^{-1}(s,\nu) = \Psi'_t(\{i\} \times F).$$

This yields the theorem. □

The notation introduced in the proof of theorem 3.2.5 will also be used in the next three lemma's.

LEMMA 3.2.6. Define

(3.2.21) $\quad t^*(i) := \inf\{t: P_R(\{x: \pi_t x = i\}) > 0\}$ for $i \in S$.

Then

(i) $\quad P_R(\{\pi_t = i\}) = 0$ iff $0 \leq t \leq t^*(i)$, for all $i \in S$ with $t^*(i) > 0$

(ii) $\quad Q_k(i,.) \overset{c}{\to} Q(i,.)$, λ-a.e. on $[t^*(i),\infty)$; $i \in S$

(iii) $r_k(i,.) \overset{c}{\to} r(i,.)$, λ-a.e. on $[t^*(i),\infty)$; $i \in S$.

PROOF.

(i) From (2.2.6) follows that $P_R(\{\pi_t = i\}) > 0$ for all $t > t^*(i)$. Since $P_R(\{\pi_t = i\})$ is continuous in t (which follows easily from theorem 1.2.26.) we conclude from the definition of $t^*(i)$ that $P_R(\{\pi_{t^*(i)} = i\}) = 0$ if $t^*(i) > 0$.

(ii) Put

$G(t) := \{x: R(.,.)$ is discontinuous at $(x,t)\}$.

From (3.2.5) follows that

$$P_R\ (G(t)) = 0,\ \lambda\text{-a.e.}$$

Let $i \in S$, $t > t^*(i)$ such that $P_R\ (G(t)) = 0$. If $t_k \in L_k$, $k \geq 1$ such that $t_k \to t$, then

$$R_k\ (x_k, t_k) \to R\ (x,t), \qquad \text{if } d(x_k,x) \to 0,\ P_R\text{-a.e.}$$

Application of (3.2.6) and (3.2.8) and corollary 1.2.21. yields the proof.

(iii) Choose $i \in S$, $t > t^*(i)$ and $t_k \in L_k$, $k \geq 1$ such that $t_k \to t$ and $Q_k(i,t_k) \overset{w}{\to} Q(i,t)$. Let F be a closed subset of A. Then

$$\limsup_{k\to\infty} r_k(i,t_k)(F) = \limsup_{k\to\infty} \int_{P(A)} \nu(F)\ dQ_k(i,t_k)(\nu) \leq$$
$$\leq \int_{P(A)} \nu(F)\ dQ(i,t)(\nu) = r(i,t)(F).$$

The inequality is justified by the upper semicontinuity of the mapping from $P(A)$ to $[0,1]$ defined by

$$\nu \to \nu(F).$$

(cf. BILLINGSLEY (1968), page 17). □

<u>LEMMA 3.2.7</u>. R' is a well-defined, regular, memoryless policy for (S,A,q,Π,c).

<u>PROOF</u>. Since the limit of a continuously convergent sequence is continuous it follows from lemma 3.2.6.(iii) and (3.2.16) that $r(i,.)$ is continuous λ-a.e. This implies that R' is a well-defined transition probability from $J[0,\infty) \times [0,\infty)$ to A and that R' is regular. Finally $R'(.,t)$ is π_t-measurable, since for all $t \geq 0$ and all Borelsets C in $P(A)$

$$\{x\colon R'(x,t) \in C\} = \pi_t^{-1}\{i \in S\colon r(i,t) \in C\}. \qquad \square$$

<u>LEMMA 3.2.8</u>. $P_{R'_k} \overset{w}{\to} P_{R'}$.

PROOF. If $t^*(i) = 0$ for all $i \in S$ the proof of this lemma follows from lemma 3.2.6.(iii) and theorem 2.4.3. Now suppose that $t^*(i) > 0$ for some $i \in S$. Condition (2.4.5) of theorem 2.4.3. is now not always satisfied. Instead of (2.4.5) we have

$$(P_{R'}^{(n)} \times \lambda)\{(z,t)\colon z=(t_j,i_j)_{j=1}^{n}, \quad t \geq t_n \vee t^*(i_n);$$

$$r_k(.,.) \not\xrightarrow{c} r(.,.) \text{ at } (i_n,t)\} = 0.$$

Proceeding as in the proof of theorem 2.4.3. it is sufficient to show that

$$(3.2.22) \qquad P_{R'_k}(F) \to P_{R'}(F).$$

for all $F = \bigcap_{n=1}^{m} T_n^{-1}(B_n) \cap S_n^{-1}(F_n) \in \mathcal{F}$ with $P_{R'}(\delta F) = 0$.

For m=1 (3.2.22) follows straightforward from the definitions. Suppose that (3.2.22) holds for all $F = \bigcap_{n=1}^{m} T_n^{-1}(B_n) \cap S_n^{-1}(F_n) \in \mathcal{F}$ with $P_{R'}(\delta F) = 0$ for some fixed m.

Moreover, we assume that for all $s > 0$ and $j \in S$

$$(3.2.23) \qquad P_{R'}(\{x\colon T_m(x) < s \wedge t^*(j);\ S_m(x) = j\}) = 0.$$

Choose $F_0 = \bigcap_{n=1}^{m} T_n^{-1}(B_n) \cap S_n^{-1}(F_n) \cap T_{m+1}^{-1}(B) \cap S_{m+1}^{-1}(F) \in \mathcal{F}$

with $P_{R'}(\delta F_0) = 0$.

The proof of (3.2.22) for F_0 proceeds similar as in theorem 2.4.3.
The only difference is that we additionally use (3.2.23) to show that $P_{R'}^{(m)}(E_3 \cap E_1^c) = 0$ (see for notation pages 56 and 57).
To complete the inductive argument we have to show that (3.2.23) holds when m is replaced by m+1, for all $s > 0$ and $j \in S$.
Since (3.2.22) holds for F_0 it is sufficient to show that

$$(3.2.24) \qquad P_{R'_k}(\{x\colon T_{m+1}(x) < s \wedge t^*(j);\ S_{m+1}(x) = j\}) \to 0.$$

From lemma 3.2.6.(i) it follows that

(3.2.25) $$P_R\{x: \pi_{s\wedge t^*(j)}(x) = j\}) = 0$$

for all $j \in S$ with $t^*(j) > 0$ and for all $s > 0$.

Combination of (3.2.8), (3.2.13) and (3.2.25) yields

(3.2.26) $$P_{R'_k}(\{x: \pi_{s\wedge t^*(j)}(x) = j\}) \to 0$$

for all $j \in S$ with $t^*(j) > 0$ and for all $s > 0$.

On the other hand we have

(3.2.27) $$P_{R'_k}\{x: \pi_{s\wedge t^*(j)}(x) = j\}) \geq$$

$$\geq \sum_{n=1}^{[(s\wedge t^*(j))k]} P_{R'_k}(\{x: S_{m+1}(x)=j;\ T_{m+1}(x)=nk^{-1};T_{m+2}(x)>s\wedge t^*(j)\}) =$$

$$= \sum_{n=1}^{[(s\wedge t^*(j))k]} \prod_{\ell=n+1}^{[(s\wedge t^*(j))k]} \int_A (1-k^{-1}q(j,a))\,dr_k(j,\ell k^{-1})(a).$$

$$\cdot\ P_{R'_k}(\{x: T_{m+1}(x) = nk^{-1};\ S_{m+1}(x) = j\}) \geq$$

$$\geq \exp(-2b(s\wedge t^*(j)))\ P_{R'_k}(\{x: T_{m+1}(x)<s\wedge t^*(j);S_{m+1}(x) = j\}).$$

In the right hand side of the last inequality b denotes $\sup\{q(s,a):\ (s,a) \in S \times A\}$.

Combination of (3.2.27) with (3.2.26) yields (3.2.24). □

CHAPTER 4

COST FUNCTIONALS ON $J[0,\infty)$

4.1. INTRODUCTION.

In this chapter we show how the results of chapter 2 can be applied in the study of continuous time stochastic optimization problems. Let $(S,A_1,A_2,q,\Pi,p,c_1,c_2,f)$ be a CTMDP, P_0 an initial distribution on S, (V,R) a policy with $R=(R_1,R_2)$ and $(V^{(k)},R^{(k)})$ a policy for the k-th approximating DTMDP, $k \geq 1$ with $R^{(k)}=(R_1^{(k)},R_2^{(k)})$. Finally let c(.) and $c^{(k)}(.)$ be cost functionals on $J[0,\infty)$ depending on policy (V,R) and $(V^{(k)},R^{(k)})$ respectively. That is, c(x) represents the total costs incurred when the stochastic process governed by policy (V,R) evolves along sample path x. In order to apply the discretization procedure to optimization problems the convergence result of theorem 2.4.3 is not sufficient. Additionally we need sufficient conditions for

$$(4.1.1) \qquad \int_{J_k[0,\infty)} c^{(k)}(x)\ dP^{(k)}_{(V^{(k)},R^{(k)})}(x) \to \int_{J[0,\infty)} c(x)\ dP_{(V,R)}(x), \quad \text{as } k\to\infty.$$

In this chapter we will give several sets of sufficient conditions for (4.1.1). First we introduce three possible cost functionals, for which we shall investigate relation (4.1.1) (see also definition 3.1.1).

<u>DEFINITION 4.1.1</u>. Let $(S,A_1,A_2,p_1,p_2,c_1,c_2,f,k)$ be a DTMDP and (V,R) a policy with $R=(R_1,R_2)$. Assume that $c_i(.,.)$ is bounded from below, i=1,2.

(i) For $0 < \alpha < 1$ the *α-discounted cost functional under (V,R)* is a function $c^{(k)}_{\alpha,(V,R)}(.)$ on $J_k[0,\infty)$ defined by

$$(4.1.2) \qquad c^{(k)}_{\alpha,(V,R)}(x) := \sum_{n=1}^{\infty} \alpha^n \int_{A_1} c_1(\pi_{nk^{-1}}x,a)\ dR_1(x,nk^{-1})(a) + \\ + \sum_{n=1}^{\infty} \alpha^n 1_V(\pi_{nk^{-1}}x) . \int_{A_2} c_2(\pi_{nk^{-1}}x,a)\,dR_2(x,nk^{-1})(a).$$

(ii) For $0 < \alpha \le 1$ the *α-discounted, n-horizon cost functional under (V,R)* is a function $c^{(k)}_{\alpha,n,(V,R)}(.)$ on $J_k[0,\infty)$ defined by

(4.1.3) $$c^{(k)}_{\alpha,n,(V,R)}(x) := \sum_{\ell=1}^{n} \alpha^{\ell} \int_{A_1} c_1(\pi_{\ell k^{-1}}x,a)\ dR_1(x,\ell k^{-1})(a) + \sum_{\ell=1}^{n} \alpha^{\ell}\, 1_V(\pi_{\ell k^{-1}}x) \int_{A_2} c_2(\pi_{\ell k^{-1}}x,a)\, dR_2(x,\ell k^{-1})(a)$$

(iii) The *average cost functional under (V,R)* is a function $c^{(k)}_{(V,R)}(.)$ on $J_k[0,\infty)$ defined by

(4.1.4) $$c^{(k)}_{(V,R)}(x) := \limsup_{n\to\infty} n^{-1} c^{(k)}_{1,n,(V,R)}(x).$$

<u>DEFINITION 4.1.2</u>. Let $(S,A_1,A_2,q,\Pi,p,c_1,c_2,f)$ be a CTMDP and (V,R) a policy. Assume that $c_1(.,.)$ is bounded from below and $c_2(.,.)$ is non-negative.

(i) For $\alpha > 0$ the *α-discounted cost functional under (V,R)* is a function $c_{\alpha,(V,R)}(.)$ on $J[0,\infty)$ defined by

(4.1.5) $$c_{\alpha,(V,R)}(x) := \int_0^{\infty} e^{-\alpha t} \int_{A_1} c_1(\pi_t x,a)\ dR_1(x,t)(a)\ dt + \sum_{n=1}^{\infty} e^{-\alpha T_{n+1}(x)}\, 1_{\{\tau_V(S_n)+T_n \le T_{n+1}\}}(x) \cdot \int_{A_2} c_2(\sigma_V(S_n(x)),a)\ dR_2(x,T_{n+1}(x))(a).$$

(ii) For $\alpha \ge 0$ the *α-discounted, T-horizon cost functional under (V,R)* is a function $c_{\alpha,T,(V,R)}(.)$ on $J[0,\infty)$ defined by

(4.1.6) $$c_{\alpha,T,(V,R)}(x) := \int_0^{T} e^{-\alpha t} \int_{A_1} c_1(\pi_t x,a)\ dR_1(x,t)(a)\ dt + \sum_{n=1}^{\infty} e^{-\alpha T_{n+1}(x)}\, 1_{\{\tau_V(S_n)+T_n \le T_{n+1} \le T\}}(x) \cdot \int_{A_2} c_2(\sigma_V(S_n(x)),a)\ dR_2(x,T_{n+1}(x))(a).$$

(iii) The *average cost functional under (V,R)* is a function $c_{(V,R)}(.)$ on $J[0,\infty)$ defined by

(4.1.7) $\quad c_{(V,R)}(x) := \limsup_{T\to\infty} T^{-1} c_{0,T,(V,R)}(x).$

Note that by remark 2.2.19.(i) the formula's (4.1.5) and (3.1.2) coincide $P_{(V,R)}$-a.e. We conclude this introduction with two theorems.

<u>THEOREM 4.1.3</u>. Let c(.) and $c^{(k)}(.)$ be cost functionals depending on (V,R) and $(V^{(k)},R^{(k)})$ respectively. If

(i) $\quad P^{(k)}_{(V^{(k)},R^{(k)})} \overset{w}{\to} P_{(V,R)}$

(ii) $\quad |c^{(k)}(x)| \le M \quad$ for all $x \in J_k[0,\infty)$ and all $k \ge 1$; for some $M > 0$

(iii) $\quad c^{(k)}(.) \overset{c}{\to} c(.)$, $P_{(V,R)}$-a.e.

then (4.1.1) holds.

<u>PROOF</u>. An immediate consequence of corollary 1.2.21. □

<u>THEOREM 4.1.4</u>. Suppose that (2.4.1) holds, $c_1(.,.)$ is bounded and $A_2=\emptyset$. Denote $\alpha_k:=\exp(-\alpha k^{-1})$, $k\ge1$ for some $\alpha>0$. Then

(4.1.8)
$$\int_{J_k[0,\infty)} c^{(k)}_{\alpha_k,(V(k),R(k))}(x)\, dP^{(k)}_{(V(k),R(k))}(x) \to \int_{J[0,\infty)} c_{\alpha,(V,R)}(x)\, dP_{(V,R)}(x), \quad \text{as } k\to\infty,$$

if condition (2.4.5) of theorem 2.4.3 is fulfilled.

<u>PROOF</u>. Choose M > 0 such that

$$|c_1(s,a)| \le M \qquad \text{for all } (s,a) \in S\times A_1.$$

Then

$$|c^{(k)}_{\alpha_k,(V^{(k)},R^{(k)})}(x)| \le Mk^{-1} \sum_{n=1}^{\infty} e^{-\alpha n k^{-1}} \le M\alpha^{-1}.$$

For $P_{(V,R)}$ almost all x we have: if $d(x_k,x) \to 0$ then

$$\lim_{k\to\infty} c^{(k)}_{\alpha_k,(V^{(k)},R^{(k)})}(x_k) =$$

$$= \lim_{k\to\infty} \sum_{n=1}^{\infty} \exp(-\alpha n k^{-1}) \int_{A_1} k^{-1} c_1(\pi_{nk^{-1}} x_k, a)\, dR_1^{(k)}(x_k, nk^{-1})(a) =$$

$$= \lim_{k\to\infty} \int_{t=0}^{\infty} \exp(-\alpha k^{-1}[tk]) \int_{A_1} c_1(\pi_{[tk]k^{-1}} x_k, a)\, dR_1^{(k)}(x_k,[tk]k^{-1})(a)\,dt=$$

$$= \int_{t=0}^{\infty} e^{-\alpha t} \int_{A_1} c_1(\pi_t x, a)\ dR_1(x,t)(a)\,dt = c_{\alpha,(V,R)}(x).$$

The last but one equality follows from (2.4.5), the continuity of $c_1(.,.)$, corollary 1.2.21. and the consideration that

$$\pi_{t_k}(.) \overset{c}{\to} \pi_t(.) \qquad \text{if } t_k \to t,$$

at all $x \in J[0,\infty)$ for which $\pi_t x = \pi_{t+} x = \pi_{t-} x$. Application of theorem 2.4.3 and 4.1.3 completes the proof. □

Theorem 2.4.3 cannot be applied to prove (4.1.8) in case of a general CTMDP with impulsive controls and/or unbounded cost rates. In this case we have to investigate relation (4.1.1) more closely.

4.2. THE α-DISCOUNTED COST FUNCTIONAL.

Let $(S,A_1,A_2,q,\Pi,p,c_1,c_2,f)$ be a CTMDP, P_0 an initial distribution on S and (V,R) a policy. Moreover, let $(V^{(k)},R^{(k)})$ be a policy for the k-th approximating DTMDP, $k \geq 1$. Assume that $c_1(.,.)$ is bounded from below and $c_2(.,.)$ non-negative. In this section we shall derive sufficient conditions for (4.1.8), which meet better the model assumptions of several applications than the conditions of theorem 4.1.4. For $\delta > 0$ and $k \geq 1$ we denote

$$W_k(\delta) := \{j \in S: \tau_{V^{(k)}}(j) > \delta\}.$$

Let $\{A_1(j): j \in S\}$ be a collection of measurable subsets of A_1, with the property that there exists a natural number k_0 (not depending on j), such that for all $k \geq k_0$

$$R_1^{(k)}(x,t)(A_1^c(j)) = 0 \qquad \text{if } \pi_t x = j,\ (P^{(k)}_{(V^{(k)},R^{(k)})} \times \lambda)\text{-a.e.}$$

Let $\{A_2(j): j \in S\}$ be a collection of measurable subsets of A_2, with the property that there exists a natural number k_0 (not depending on j), such that for all $k \geq k_0$ and $t \geq 0$ we have

$$R_2^{(k)}(x,t)(A_2^c(j)) = 0 \quad \text{if } \pi_t x = j \in V^{(k)},\ (P^{(k)}_{V^{(k)},R^{(k)}} \times \lambda)\text{-a.e.}$$

THEOREM 4.2.1. Suppose that (2.4.1) holds and that conditions (2.4.3), (2.4.4) and (2.4.5) of theorem 2.4.3 are satisfied. Choose $\alpha > 0$. Then (4.1.8) holds if the following conditions are fulfilled.

(i) There exist a $\delta_0 > 0$ and a natural number k_0 such that for all $k \geq k_0$ and $n \geq 1$

(4.2.1) $$\int_{A_2} p(\pi_{nk^{-1}} x, a, W_k(\delta_0))\, dR_2^{(k)}(x, nk^{-1})(a) = 1$$

if $\tau_{V^{(k)}}(\pi_{nk^{-1}} x) = k^{-1}$.

(ii) There exist a constant $\beta \geq 1$, real-valued non-negative continuous functions $h_1(.)$ and $h_2(.)$ on S and a measurable function ℓ from $[0,\infty)$ into $[1,\infty)$, such that

(4.2.2) $$\sup_{a \in A_i(j)} |c_i(j,a)| \leq h_i(j),\ j \in S;\ i = 1,2.$$

(4.2.3) $$h_i^2(f(j,s)) \leq h_i^2(j)\ell(s),\ (j,s) \in S \times [0,\infty);\ i = 1,2.$$

(4.2.4) $$e^{-\alpha s}\ell(s) \leq 1,\ s \geq 0$$

(4.2.5) $$\int_0^\infty e^{-\alpha s}\ell(s)\, ds < \infty.$$

(4.2.6) $$\sup_{a \in A_1(j)} \int_S h_i^2(j_0)\, d\Pi(j,a,j_0) \leq \beta h_i^2(j),\ j \in S;\ i = 1,2.$$

(4.2.7) $$\sup_{a \in A_2(j)} \int_S h_i^2(j_0)\, dp(j,a,j_0) \leq \beta h_i^2(j),\ j \in S;\ i = 1,2.$$

(4.2.8) $$\int_S h_i^2(j)\, dP_0(j) =: \gamma_i < \infty, \qquad i = 1,2.$$

(4.2.9) $\quad q(.,.) \int_S h_i(j)\, d\Pi(.,.,j)$ is continuous on $S \times A_1$; $i = 1,2$.

(4.2.10) $\quad \int_S h_i(j)\, dp(.,.,j)$ is continuous on $S \times A_2$; $i = 1,2$.

(4.2.11) $\quad \beta^2(p(\delta)e^{-\alpha\delta} + 1 - p(\delta)) < 1$ and $p(\delta) \geq 0$ for some $0 < \delta < \delta_0$

where

$$p(\delta) := 2\exp(-b\delta - b^2\delta) - 1$$

$$b := \sup\{q(s,a) : (s,a) \in S \times A_1\}.$$

PROOF

In the proof of this theorem we put for abbreviation: $P := P_{(V,R)}$, $P_k := P^{(k)}_{(V^{(k)},R^{(k)})}$, $P^{(n)} := P^{(n)}_{(V,R)}$, $P_k^{(n)} := P^{(n)}_{k,(V^{(k)},R^{(k)})}$, $\tau_k := \tau^{(k)}_{V^{(k)}}$, $\tau := \tau_V$, $g_k := g^{(k)}_{V^{(k)}}$, $g := g_V$, $\sigma_k := \sigma^{(k)}_{V^{(k)}}$ and $\sigma := \sigma_V$.

From theorem 2.4.3 follows that

$$P_k \overset{w}{\to} P.$$

With the continuous mapping theorem this implies for $n \geq 1$

(4.2.12) $\quad P_k^{(n)} \overset{w}{\to} P^{(n)} \quad$ on J_n.

We consider two cases separately, which together yield the proof of the theorem.

Case (i). Assume that $c_2(s,a) = 0$ for all $(s,a) \in S \times A_2$. Put for $n \geq 1$ and $k \geq 1$:

$$f_k(n) := \int_{J_k[0,\infty)} \sum_{\ell=T_n(x)k+1}^{T_{n+1}(x)k} \alpha_k^{\ell} k^{-1} \int_{A_1} c_1(\pi_{\ell k^{-1}}x,a)\, dR_1^{(k)}(x,\ell k^{-1})(a)\, dP_k(x)$$

and for $n \geq 1$:

$$f(n) := \int_{J[0,\infty)} \int_{t=T_n(x)}^{T_{n+1}(x)} e^{-\alpha t} \int_{A_1} c_1(\pi_t x,a)\, dR_1(x,t)(a)\, dt\, dP(x).$$

Then (4.1.8) holds if there exist sequences of non-negative real numbers $\{b(n)\}$, $\{b_k(n)\}$, $k \geq 1$ such that

(4.2.13) $\quad f_k(n) \to f(n) \qquad$ as $k \to \infty$, $n \geq 1$

(4.2.14) $\quad |f_k(n)| \leq b_k(n) \qquad$ for $n \geq 1$ and $k \geq 1$

(4.2.15) $\quad b_k(n) \to b(n) \qquad$ as $k \to \infty$, $n \geq 1$

(4.2.16) $\quad \sum_{n=1}^{\infty} b_k(n) \to \sum_{n=1}^{\infty} b(n) < \infty \qquad$ as $k \to \infty$.

Next we define for $z = (t_1, j_1, \ldots, t_n, j_n) \in J_n(k)$ respectively $\in J_n$

$$u_k(z) := 1_{[0,\infty)}(t_n) \sum_{m=t_n k+1}^{\infty} Q_k^{(n)}(z)(\{mk^{-1}\} \times S^+) .$$
$$\cdot \sum_{\ell=t_n k+1}^{m} \alpha_k^{\ell} \int_{A_1} k^{-1} c_1(f(j_n, \ell k^{-1} - t_n), a)\, dR_1^{(k)}(g_k(z), \ell k^{-1})(a)$$

(the value $m=\infty$ is supposed to be included in the first summation) and

$$u(z) := 1_{[0,\infty)}(t_n) \int_{(s,j)\in[t_n,\infty]\times S^+} \int_{t=t_n}^{s} e^{-\alpha t} \cdot$$
$$\cdot \int_{A_1} c_1(f(j_n, t-t_n), a)\, dR_1(g(z), t)(a)\, dt\; dQ^{(n)}(z)(s,j).$$

Then

$$f_k(n) = \int_{J_n(k)} u_k(z)\; dP_k^{(n)}(z)$$

and

$$f(n) = \int_{J_n} u(z)\; dP^{(n)}(z).$$

Hence we conclude from (4.2.12) and corollary 1.2.21 that sufficient for (4.2.13) is the existence of non-negative functions $v(.)$ and $v_k(.)$ on J_n and $J_n(k)$ respectively, such that

(4.2.17) $\quad u_k(.) \overset{c}{\to} u(.), \qquad P^{(n)}$-a.e.

(4.2.18) $\quad |u_k(.)| \le v_k(.) \qquad$ for $k \ge 1$

(4.2.19) $\quad v_k(.) \overset{c}{\to} v(.), \qquad P^{(n)}$-a.e.

(4.2.20) $$\int_{J_n(k)} v_k(z)\, dP_k^{(n)}(z) \to \int_{J_n} v(z)\, dP^{(n)}(z), \qquad \text{as } k \to \infty.$$

From lemma 4.2.2.(i) below follows that (4.2.17) holds. Next we define for $k \ge 1$ and for $z = (t_1,j_1,\dots,t_n,j_n) \in J_n(k)$ and $\in J_n$ respectively

$$v_k(z) := 1_{[0,\infty)}(t_n)\; h_1(j_n)\; e^{-\alpha t_n}\, k^{-1} \sum_{m=1}^{\infty} e^{-\alpha mk^{-1}}\, \ell(mk^{-1})$$

and

$$v(z) := 1_{[0,\infty)}(t_n)\; h_1(j_n) e^{-\alpha t_n} \int_0^\infty e^{-\alpha s}\ell(s)\,ds.$$

Then (4.2.18) holds by virtue of lemma 4.2.3.(i) below; (4.2.19) follows from (4.2.5) and the continuity of $h_1(.)$ and (4.2.20) follows from lemma 4.2.5 below.

What remains to prove is the existence of sequences $\{b(n)\}$, $\{b_k(n)\}$, $k \ge 1$ such that (4.2.14), (4.2.15) and (4.2.16) hold.

By Schwarz' inequality we have

(4.2.21) $$\int_{J_n(k)} h_1(j_n) e^{-\alpha t_n} dP_k^{(n)}(z) \le$$
$$\le \Big[\int_{J_n(k)} h_1^2(j_n)\; e^{-\alpha t_n} dP_k^{(n)}(z) \Big]^{\frac{1}{2}} \Big[\int_{J_n(k)} e^{-\alpha t_n} dP_k^{(n)}(z) \Big]^{\frac{1}{2}}.$$

Combining (4.2.21) with (4.2.8) and lemmas 4.2.4 and lemma 4.2.6 below yields for all $\delta \le \delta_0$ for which $p(\delta) \ge 0$ and for all k large enough

(4.2.22) $$\int_{J_n(k)} h_1(j_n) e^{-\alpha t_n} dP_k^{(n)}(z) \le (\beta^n \gamma_1)^{\frac{1}{2}} . (e^{-\alpha\delta} p(\delta) + 1 - p(\delta))^{\frac{n-1}{4}}.$$

Put

$$b_k(n) := (\beta^n \gamma_1)^{\frac{1}{2}} (e^{-\alpha\delta} p(\delta) + 1 - p(\delta))^{\frac{n-1}{4}}\; k^{-1} \sum_{m=1}^{\infty} e^{-\alpha mk^{-1}}\, \ell(mk^{-1})$$

and

$$b(n) := (\beta^n \gamma_1)^{\frac{1}{2}} (e^{-\alpha\delta} p(\delta)+1-p(\delta))^{\frac{n-1}{4}} \int_0^\infty e^{-\alpha s} \ell(s) ds$$

with $0 < \delta \leq \delta_0$ such that $\beta^2 (e^{-\alpha\delta} p(\delta)+1-p(\delta)) < 1$, which is possible by virtue of (4.2.11).

One easily verifies with (4.2.22) that these choices for $\{b_k(n)\}$ and $\{b(n)\}$ satisfy (4.2.14), (4.2.15) and (4.2.16). This completes the proof of case (i).

Case (ii). Assume that $c_1(s,a) = 0$ for all $(s,a) \in S \times A_1$.

For $z = (t_1, j_1, \ldots, t_n, j_n) \in J_n(k)$ and $\in J_n$ respectively, let $\tilde{u}_k(z)$ and $\tilde{u}(z)$ represent the expected discounted costs upto the next jump epoch under policy $(V^{(k)}, R^{(k)})$ and (V,R) respectively, i.e.

$$\tilde{u}_k(z) := Q_k^{(n)}(z)(\{\tau_k(j_n)+t_n\} \times S) \cdot \alpha_k^{t_n k+\tau_k(j_n)k} \cdot \int_{A_2} c_2(\sigma_k'(j_n),a) dR_2^{(k)}(g_k(z), t_n+\tau_k(j_n))(a)$$

and

$$\tilde{u}(z) := Q^{(n)}(z)(\{\tau(j_n)+t_n\} \times S) \cdot e^{-\alpha(t_n+\tau(j_n))} \cdot \int_{A_2} c_2(\sigma(j_n),a) dR_2(g(z), t_n+\tau(j_n))(a).$$

When $\tau_k(j_n)=\infty$ or $\tau(j_n)=\infty$ the right hand sides are interpreted as zero. Put

$$\tilde{f}_k(n) := \int_{J_n(k)} \tilde{u}_k(z) dP_k^{(n)}(z)$$

and

$$\tilde{f}(n) := \int_{J_n} \tilde{u}(z) dP^{(n)}(z).$$

Then (4.1.8) holds if there exist sequences of non-negative real numbers $\{\tilde{b}(n)\}$, $\{\tilde{b}_k(n)\}$, $k \geq 1$, such that

(4.2.23) $\quad \tilde{f}_k(n) \to \tilde{f}(n), \qquad$ as $k \to \infty$; $n \geq 1$

(4.2.24) $\quad |\tilde{f}_k(n)| \leq \tilde{b}_k(n) \qquad$ for $n \geq 1$ and $k \geq 1$

(4.2.25) $\quad \tilde{b}_k(n) \to \tilde{b}(n) \qquad$ as $k \to \infty$; $n \geq 1$

(4.2.26) $\quad \sum_{n=1}^{\infty} \tilde{b}_k(n) \to \sum_{n=1}^{\infty} \tilde{b}(n) < \infty \qquad$ as $k \to \infty$.

Sufficient for (4.2.23) is the existence of non-negative functions $\tilde{v}(.)$ and $\tilde{v}_k(.)$ on J_n and $J_n(k)$ respectively, such that

(4.2.27) $\quad \tilde{u}_k(.) \overset{c}{\to} \tilde{u}(.)$, $P^{(n)}$-a.e.

(4.2.28) $\quad |\tilde{u}_k(.)| \leq \tilde{v}_k(.), \qquad k \geq 1$

(4.2.29) $\quad \tilde{v}_k(.) \overset{c}{\to} \tilde{v}(.), \qquad P^{(n)}$-a.e.

(4.2.30) $\quad \int_{J_n(k)} \tilde{v}_k(z)\, dP_k^{(n)}(z) \to \int_{J_n} \tilde{v}(z)\, dP^{(n)}(z), \qquad$ as $k \to \infty$.

From lemma 4.2.2.(ii) below follows that (4.2.27) holds.

Define for $k \geq 1$ and for $z = (t_1, j_1, \ldots, t_n, j_n) \in J_n(k)$ and $\in J_n$ respectively

$$\tilde{v}_k(z) := \tilde{v}(z) := h_2(j_n)e^{-\alpha t_n}.$$

The rest of the proof proceeds similarly as in case (i). □

We assume that all quantities introduced in theorem 4.2.1. preserve their meaning in the next five lemma's.

LEMMA 4.2.2. (i) $u_k(.) \overset{c}{\to} u(.)$, $P^{(n)}$-a.e.
(ii) $\tilde{u}_k(.) \overset{c}{\to} \tilde{u}(.)$, $P^{(n)}$-a.e.

PROOF. Let $z = (t_1, j_1, \ldots, t_n, j_n) \in J_n$, with $t_n < \infty$ and $z(k) = (t_1(k), j_1(k), \ldots, t_n(k), j_n(k)) \in J_n(k)$, $k \geq 1$ such that $z(k) \to z$.

(i) By interchanging the summation in the definition of $u_k(.)$ we have

$$(4.2.31) \qquad u_k(z(k)) = \sum_{\ell=t_n(k)k+1}^{\infty} \alpha_k^{\ell} Q_k^{(n)}(z(k))([\ell k^{-1}, \infty] \times S^+) \,. \\ \cdot \int_{A_1} k^{-1} c_1(f(j_n(k), \ell k^{-1} - t_n(k)), a)\, dR_1^{(k)}(g_k(z(k)), \ell k^{-1})(a).$$

From lemmas 2.4.4 and 2.4.6 follows that for λ-almost all s

$$(4.2.32) \qquad Q_k^{(n)}(z(k))([s_k, \infty] \times S^+) \to Q^{(n)}(z)([s, \infty] \times S^+)$$

for any sequence $s_k \to s$.
Combining (4.2.31), (4.2.32), the continuity of $c_1(.,.)$, (2.4.5), (4.2.2), (4.2.3) and (4.2.5) yields

$$\lim_{k\to\infty} u_k(z(k)) = \\ = \int_{t=t_n}^{\infty} e^{-\alpha t} Q^{(n)}(z)([t,\infty]\times S^+) \int_{A_1} c_1(f(j_n, t-t_n), a)\, dR_1(g(z), t)(a)\, dt = \\ = u(z), \; P^{(n)}\text{-a.e.}$$

(ii) From lemma 2.4.4 follows with (2.4.3) that

$$(4.2.33) \qquad Q_k^{(n)}(z(k))(\{\tau_k(j_n(k)) + t_n(k)\} \times S) \to Q^{(n)}(z)(\{\tau(j_n) + t_n\} \times S).$$

Together (4.2.33), the continuity of $c_2(.,.)$, (2.4.3), (2.4.4) and the definition of $\tilde{u}_k(.)$ and $\tilde{u}(.)$ yield

$$\lim_{k\to\infty} \tilde{u}_k(z(k)) = Q^{(n)}(z)(\{\tau(j_n) + t_n\} \times S) e^{-\alpha(t_n + \tau(j_n))} \,. \\ \cdot \int_{A_2} c_2(\sigma(j_n), a)\, dR_2(g(z), t_n + \tau(j_n))(a) = \\ = \tilde{u}(z), \; P^{(n)}\text{-a.e.}$$

LEMMA 4.2.3. (i) $|u_k(z)| \le v_k(z)$ for all $z \in J_n(k)$, $k \ge 1$
(ii) $|\tilde{u}_k(z)| \le \tilde{v}_k(z)$ for all $z \in J_n(k)$, $k \ge 1$.

PROOF. From (4.2.31), (4.2.2) and (4.2.3) follows (i), while (ii) is a consequence of the definition of $\tilde{u}_k(z)$, (4.2.2), (4.2.3) and (4.2.4). □

LEMMA 4.2.4.

$$\int_{J_{n+1}(k)} e^{-\alpha t_{n+1}} h_i^{\ell}(j_{n+1})\, dP_k^{(n+1)}(z) \le$$

$$\le \beta \int_{J_n(k)} e^{-\alpha t_n} h_i^{\ell}(j_n)\, dP_k^{(n)}(z)$$

for all $k \ge k_0$, $n \ge 1$; $i = 1,2$; $\ell=1,2$.

We prove the lemma for $\ell=1$. The proof for $\ell=2$ proceeds similarly.

PROOF. From Schwarz' inequality, (4.2.6) and (4.2.7) follows

$$\text{(4.2.34)} \qquad \sup_{a\in A_1(j)} \int_S h_i(j_0)\, d\Pi(j,a,j_0) \le \beta h_i(j), \qquad j \in S;\ i = 1,2.$$

and

$$\text{(4.2.35)} \qquad \sup_{a\in A_2(j)} \int_S h_i(j_0)\, dp(j,a,j_0) \le \beta h_i(j), \qquad j \in S;\ i = 1,2.$$

Next we observe that

$$\text{(4.2.36)} \qquad \int_{J_{n+1}(k)} e^{-\alpha t_{n+1}} h_i(j_{n+1})\, dP_k^{(n+1)}(z) =$$

$$= \int_{J_n(k)} \int_{[0,\infty]\times S^+} e^{-\alpha t} h_i(j)\, dQ_k^{(n)}(z)(t,j)\, dP_k^{(n)}(z).$$

By the definition of $Q_k^{(n)}$ (see definition 2.3.7) we find for $z = (t_1, j_1, \ldots t_n, j_n)$

(4.2.37)
$$\int_{[0,\infty]\times S^+} e^{-\alpha t} h_i(j)\, dQ_k^{(n)}(z)(t,j) =$$

$$= \sum_{m=1}^{\tau_k(j_n)k-1} \prod_{\ell=1}^{m-1} \int_{A_1} \{1-k^{-1}q(f(j_n,\ell k^{-1}),a)+k^{-1}o(1)\}$$
$$dR_1^{(k)}(g_k(z),\ell k^{-1}+t_n)(a).$$
$$. e^{-\alpha(t_n+mk^{-1})} \int_{A_1} k^{-1}\, q(f(j_n,mk^{-1}),a)\ .$$
$$. \int_S h_i(j)\, d\Pi(f(j_n,mk^{-1}),a,j)\, dR_1^{(k)}(g_k(z),mk^{-1}+t_n)(a)\ +$$

$$+ 1_{\{\tau_k(j)<\infty\}}(j_n)\ .$$

$$. \prod_{\ell=1}^{\tau_k(j_n)k-1} \int_{A_1} \{1-k^{-1}q(f(j_n,\ell k^{-1}),a)+k^{-1}o(1)\}$$
$$dR_1^{(k)}(g_k(z),\ell k^{-1}+t_n)(a)\ .$$
$$. e^{-\alpha(t_n+\tau_k(j_n))} \int_{A_2}\int_S h_i(j)\, dp(\sigma_k(j_n),a,j)$$

$$dR_2^{(k)}(g_k(z),\tau_k(j_n)+t_n)(a).$$

From (4.2.37) follows with (4.2.33), (4.2.34) and (4.2.3) that

(4.2.38)
$$\int_{[0,\infty]\times S^+} e^{-\alpha t} h_i(j)\, dQ_k^{(n)}(z)(t,j) \le$$
$$\le \beta e^{-\alpha t_n}\, h_i(j_n) \int_{[0,\infty]\times S^+} e^{-\alpha(t-t_n)}\, \ell(t-t_n)\, dQ_k^{(n)}(z)(t,j).$$

Now (4.2.36) and (4.2.38) together with (4.2.4) yield the proof. □

<u>LEMMA 4.2.5</u>. For all $n \ge 1$

$$\int_{J_n(k)} e^{-\alpha t_n}\, h_i(j_n)\, dP_k^{(n)}(z) \to \int_{J_n} e^{-\alpha t_n}\, h_i(j_n)\, dP^{(n)}(z),\ \text{as } k \to \infty.$$

PROOF. The proof proceeds by induction on n. For n = 1 the lemma follows from the equality

$$\int_{J_1(k)} e^{-\alpha t_1} h_i(j_1)\, dP_k^{(1)}(z) = \int_S h_i(j)\, dP_0(j) =$$
$$= \int_{J_1} e^{-\alpha t_1} h_i(j_1)\, dP^{(1)}(z).$$

Suppose the lemma is true for n.
By (4.2.36) and (4.2.12) it is sufficient to show that there exist non-negative functions $w(.)$, $w_k(.)$ on J_n and $J_n(k)$, $k \geq 1$ respectively, such that

$$(4.2.39)\qquad \int_{[0,\infty]\times S^+} e^{-\alpha t} h_i(j)\, dQ_k^{(n)}(.)(t,j) \overset{c}{\to} \int_{[0,\infty]\times S^+} e^{-\alpha t} h_i(j)\, dQ^{(n)}(.)(t,j), \qquad \text{as } k \to \infty.$$

$$(4.2.40)\qquad \int_{[0,\infty]\times S^+} e^{-\alpha t} h_i(j)\, dQ_k^{(n)}(z)(t,j) \leq w_k(z), \qquad \text{for } z \in J_n(k)$$

$$(4.2.41)\qquad w_k(.) \overset{c}{\to} w(.), \qquad P^{(n)}\text{-a.e.}$$

$$(4.2.42)\qquad \int_{J_n(k)} w_k(z)\, dP_k^{(n)}(z) \to \int_{J_n} w(z)\, dP^{(n)}(z), \qquad \text{as } k \to \infty.$$

From (4.2.37) we find with application of (4.2.9) and (4.2.10) that (4.2.39) holds.
Put for $z = (t_1, j_1, \ldots, t_n, j_n) \in J_n$ and $\in J_n(k)$ respectively

$$w(z) := w_k(z) = \beta e^{-\alpha t_n} h_i(j_n).$$

Then (4.2.40) follows from (4.2.38); (4.2.41) follows from the continuity of $h_i(.)$ and the induction hypothesis yields (4.2.42). □

<u>LEMMA 4.2.6.</u> $\int_{J_{2n}(k)} e^{-\alpha t_{2n}} dP_k^{(2n)}(z) \le (e^{-\alpha\delta}p(\delta)+1-p(\delta))^n$

for all $0 < \delta < \delta_0$ with $p(\delta) \ge 0$, all $n \ge 1$ and all k large enough.

<u>PROOF</u>. From (4.2.1) follows, with the same arguments as used in the proof of theorem 2.2.18. that for all $0 \le \delta < \delta_0$, all $s \ge t$ and all k large enough

$$P_k^{(n+2)}(\{z=(t_j,i_j)_{j=1}^{n+2}:t_{n+2}\le t+\delta\}\,|\,\{z=(t_j,i_j)_{j=1}^{n+2}:t_n=s\}) \le \tag{4.2.43}$$

$$\le 2(1-(1-k^{-1}b)^{[\delta k]}) \le 1-p(\delta),$$

where $p(\delta) = 2\exp(-b\delta-b^2\delta) - 1$.

Choose $0 < \delta < \delta_0$ such that $p(\delta) > 0$ and define a sequence $(X_m)_{m=1}^{\infty}$ of independent random variables on a common probability space $(\Omega, A, \mathbb{P})$, such that for all $m \ge 1$

$$\mathbb{P}(X_m = \delta) = p(\delta)$$

and

$$\mathbb{P}(X_m = 0) = 1 - p(\delta).$$

Then we get from (4.2.43) by induction on n for all k large enough

$$P_k^{(2n)}\{z=(t_j,i_j)_{j=1}^{2n}:t_{2n} \le t\} \le \mathbb{P}(\sum_{m=1}^{n} X_m \le t), \qquad n \ge 1.$$

This in turn implies

$$\int_{J_{2n}(k)} e^{-\alpha t_{2n}} dP_k^{(2n)}(z) \le \int_{\Omega} e^{-\alpha \sum_{m=1}^{n} X_m(\omega)} d\mathbb{P}(\omega). \tag{4.2.44}$$

Since the right hand side of (4.2.44) equals

$$\sum_{k=0}^{n} \binom{n}{k} e^{-\alpha k\delta}(p(\delta))^k(1-p(\delta))^{n-k} = (e^{-\alpha\delta}p(\delta)+1-p(\delta))^n$$

the lemma follows. □

Theorem 4.2.1 can be successfully applied to prove structural properties of optimal policies for CTMDP's. To make this explicit we need some definitions.

DEFINITION 4.2.7.

(i) Let $(S,A_1,A_2,p_1,p_2,c_1,c_2,f,k)$ be a DTMDP and P_0 an initial distribution on S. Assume that $c_i(.,.)$ is bounded from below, $i = 1,2$. Let $\mathcal{R}$ denote a certain class of policies for this process. A policy (V^*,R^*) is called *α-discounted optimal in* $\mathcal{R}$ if for every policy $(V,R) \in \mathcal{R}$

$$\int_{J_k[0,\infty)} c^{(k)}_{\alpha,(V^*,R^*)}(x)\,dP^{(k)}_{(V^*,R^*)}(x) \le \int_{J_k[0,\infty)} c^{(k)}_{\alpha,(V,R)}\,dP^{(k)}_{(V,R)}(x).$$

(ii) Let $(S,A_1,A_2,q,\Pi,p,c_1,c_2,f)$ be a CTMDP and P_0 an initial distribution on S. Assume that $c_1(.,.)$ is bounded from below and $c_2(.,.)$ non-negative. Let $\mathcal{R}$ denote a certain class of policies for this process. A policy (V^*,R^*) is called *α-discounted optimal in* $\mathcal{R}$ if for every policy $(V,R) \in \mathcal{R}$

$$\int_{J[0,\infty)} c_{\alpha,(V^*,R^*)}(x)\,dP_{(V^*,R^*)}(x) \le \int_{J[0,\infty)} c_{\alpha,(V,R)}(x)\,dP_{(V,R)}(x).$$

Similar definitions can be given for optimality with respect to finite horizon cost functionals.

DEFINITION 4.2.8. Let $(S,A_1,A_2,q,\Pi,p,c_1,c_2,f)$ be a CTMDP and P_0 an initial distribution on S. Assume that $c_1(.,.)$ is bounded from below and $c_2(.,.)$ non-negative. A regular policy (V,R) for this process is called *strong regular* if there exists for all $\alpha > 0$ a policy $(V^{(k)},R^{(k)})$ for the k-th approximating DTMDP, $k \ge 1$ such that (4.1.8) holds.

THEOREM 4.2.9. Let $(S,A_1,A_2,q,\Pi,p,c_1,c_2,f)$ be a CTMDP and P_0 an initial distribution on S. Assume that (2.4.1) holds, that $c_1(.,.)$ is bounded from below and $c_2(.,.)$ non-negative. Put $\alpha_k := \exp(-\alpha k^{-1})$ for some $\alpha > 0$. Let (V_k^*,R_k^*) be a α_k-discounted optimal policy for the k-th approximating DTMDP in the class of all policies, $k \ge 1$. If there exists a policy

(V^*,R^*) for the CTMDP such that (V^*,R^*) and $(V^*_{k_j},R^*_{k_j})_{j=1}^{\infty}$ satisfy (4.1.8) for some subsequence $(k_j)_{j=1}^{\infty}$ then (V^*,R^*) is a α-discounted optimal policy in the class of strong regular policies.

PROOF. Assume the contrary i.e. suppose there exists a strong regular policy (V,R) such that

$$(4.2.45)\qquad \int_{J[0,\infty)} c_{\alpha,(V,R)}(x)\,dP_{(V,R)}(x) < \int_{J[0,\infty)} c_{\alpha,(V^*,R^*)}(x)\,dP_{(V^*,R^*)}(x).$$

Since (V,R) is strong regular there exists for $k \geq 1$ a policy $(V^{(k)},R^{(k)})$ for the k-th approximating DTMDP, such that

$$(4.2.46)\qquad \int_{J_k[0,\infty)} c^{(k)}_{\alpha_k,(V^{(k)},R^{(k)})}(x)\,dP^{(k)}_{(V^{(k)},R^{(k)})}(x) \to$$

$$\to \int_{J[0,\infty)} c_{\alpha,(V,R)}(x)\,dP_{(V,R)}(x),\ \text{as } k \to \infty.$$

From the assumptions of the theorem follows

$$(4.2.47)\qquad \int_{J_{k_j}[0,\infty)} c^{(k_j)}_{\alpha_{k_j},(V^*_{k_j},R^*_{k_j})}(x)\,dP^{(k_j)}_{(V^*_{k_j},R^*_{k_j})}(x) \to$$

$$\to \int_{J[0,\infty)} c_{\alpha,(V^*,R^*)}(x)\,dP_{(V^*,R^*)}.$$

Together (4.2.45), (4.2.46) and (4.2.47) contradict the optimality of $(V^*_{k_j},R^*_{k_j})$ for j large enough. □

THEOREM 4.2.10. Let $(S,A_1,A_2,q,\Pi,p,c_1,c_2,f)$ be a CTMDP and P_0 an initial distribution on S. Assume that (2.4.1) holds, that $c_1(.,.)$ is bounded from below and $c_2(.,.)$ non-negative. Put $\alpha_k := \exp(-\alpha k^{-1})$ for some $\alpha > 0$. Let $(V^{(k)},R^{(k)})$ be a α_k-discounted optimal policy for the k-th approximating DTMDP in the class of all policies, with $R^{(k)}=(R_1^{(k)},R_2^{(k)})$. A policy (V,R) for the CTMDP is α-discounted optimal in the class of strong regular policies if at least one of the following three sets of conditions is fulfilled:

(i)

$$\liminf_{k\to\infty} \int_{J_k[0,\infty)} c^{(k)}_{\alpha_k,(V^{(k)},R^{(k)})}(x)\,dP^{(k)}_{(V^{(k)},R^{(k)})}(x) \geq \tag{4.2.48}$$

$$\geq \int_{J[0,\infty)} c_{\alpha,(V,R)}(x)\,dP_{(V,R)}(x)$$

(ii) $c_1(.,.)$ is non-negative; (2.4.3), (2.4.4) and (2.4.5) hold and

$$\liminf_{k\to\infty} c^{(k)}_{\alpha_k,(V^{(k)},R^{(k)})}(x_k) \geq c_{\alpha,(V,R)}(x)$$

for any sequence $(x_k)_{k=1}^{\infty}$ with $d(x_k,x) \to 0$, $P_{(V,R)}$-a.e.

(iii) $c_1(.,.)$ is non-negative; $V=\emptyset$ and (2.4.5) holds.

PROOF.

(i) The proof of theorem 4.2.9 goes through with replacement of (4.2.47) by (4.2.48).

(ii) An immediate consequence of theorem 2.4.3, corollary 1.2.23 and part (i).

(iii) Since $V = \emptyset$ we have

$$c_{\alpha,(V,R)}(x) = \int_0^\infty e^{-\alpha t} \int_{A_1} c_1(\pi_t x,a)\ dR_1(x,t)(a)dt$$

while

$$c^{(k)}_{\alpha_k,(V^{(k)},R^{(k)})}(x)=k^{-1}\sum_{n=1}^{\infty}\alpha_k^n \int_{A_1} c_1(\pi_{nk^{-1}}x,a)\,dR_1^{(k)}(x,nk^{-1})(a).$$

Hence for any sequence (x_k) for which $d(x_k,x) \to 0$

$$\liminf c^{(k)}_{\alpha_k,(V^{(k)},R^{(k)})}(x_k) =$$

$$= \liminf_{k\to\infty} \int_{t=0}^{\infty} e^{-\alpha k^{-1}[tk]} \int_{A_1} c_1(\pi_{[tk]k^{-1}}\, x_k,a)\ dR_1^{(k)}(x_k,[tk]k^{-1})(a)dt \geq$$

$$\geq \int_0^\infty e^{-\alpha t} \liminf_{k\to\infty} \int_{A_1} c_1(\pi_{[tk]k^{-1}}\, x_k, a)\, dR_1^{(k)}(x_k, [tk]k^{-1})(a)dt =$$

$$= c_{\alpha,(V,R)}(x).$$

Application of part (ii) completes the proof. □

For applications of the theorems 4.2.9 and 4.2.10 the reader is referred to chapters 5, 6 and 7.

4.3. THE α-DISCOUNTED FINITE HORIZON COST FUNCTIONAL.

For the α-discounted finite horizon cost functional we have an analogy of theorem 4.2.1.

THEOREM 4.3.1. Let $(S,A_1,A_2,q,\Pi,p,c_1,c_2,f)$ be a CTMDP, P_0 an initial distribution on S, (V,R) a policy and $(V^{(k)},R^{(k)})$ a policy for the k-th approximating DTMDP, $k \geq 1$. Assume that $c_1(.,.)$ is bounded from below and $c_2(.,.)$ non-negative. Choose $\alpha > 0$, put $\alpha_k := \exp(-\alpha k^{-1})$ and let $T > 0$. If the conditions of theorem 4.2.1 hold, then

$$\int_{J_k[0,\infty)} c_{\alpha_k,T_k,(V^{(k)},R^{(k)})}(x)\ dP^{(k)}_{(V^{(k)},R^{(k)})}(x) \to$$

$$\to \int_{J[0,\infty)} c_{\alpha,T,(V,R)}(x)\ dP_{(V,R)}(x)$$

as $k \to \infty$, for any sequence (T_k) of natural numbers for which $k^{-1}T_k \to T$ and for any T for which

$$(4.3.1) \qquad P^{(n)}_{(V,R)}\{z = (t_j,i_j)_{j=1}^n : \tau_V(i_n) = T\} = 0, \qquad n \geq 1.$$

PROOF. The proof proceeds along the same lines as the proof of theorem 4.2.1 and hence it will be omitted. □

From theorem 4.3.1 we can easily derive a finite horizon analogy of theorem 4.2.9. However, this analogy is not as useful as theorem 4.2.9 itself, since in most finite horizon applications with impulsive controls there does not exist an optimal policy with time independent impulsive control set.

4.4. THE AVERAGE COST FUNCTIONAL.

The analysis given in section 4.2 for the α-discounted cost functional fails for the average cost functional. Even for a CTMDP $(S,A_1,A_2,q,\Pi,p,c_1,c_2,f)$ with $A_2 = \emptyset$ and $|c_1(.,.)|$ bounded, the conditions of theorem 2.4.3 do not ensure that

$$(4.4.1) \qquad \int_{J_k[0,\infty)} c^{(k)}_{(V^{(k)},R^{(k)})}(x)\,dP^{(k)}_{(V^{(k)},R^{(k)})}(x) \to$$

$$\to \int_{J[0,\infty)} c_{(V,R)}(x)\,dP_{(V,R)}(x), \qquad \text{as } k \to \infty.$$

<u>EXAMPLE 4.4.1</u>. Put $S = \{1,2\}$; $A_1 = [0,1]$; $A_2 = \emptyset$; $q(i,a) = a$ for $(i,a) \in S \times A_1$; $\Pi(1,a,\{2\}) = \Pi(2,a,\{1\}) = 1$ for all $a \in A_1$; $c_1(1,a) = c_1 \neq c_2 = c_1(2,a)$, $a \in A_1$; $f(i,t) = i$ for $(i,t) \in S \times [0,\infty)$. Suppose the initial distribution is given by $P_0(\{1\}) = 1$. Let (V,R) be a policy for this process defined by $V = \emptyset$ and $R_1(x,t) = 0$ for all $(x,t) \in J[0,\infty) \times [0,\infty)$.
Define for $k \geq 1$ a policy $(V^{(k)},R^{(k)})$ for the k-th approximating DTMDP by $V^{(k)}=\emptyset$ and $R_1^{(k)}(x,t) = k^{-1}$ for all $(x,t) \in J_k[0,\infty) \times L_k$.
Note that condition (2.4.1) holds, (2.4.3) and (2.4.4) are trivially satisfied and (2.4.5) follows from the definitions of R_1 and $R_1^{(k)}$. However,

$$\int_{J[0,\infty)} c_{(V,R)}(x)\; dP_{(V,R)}(x) = c_1$$

and

$$\int_{J_k[0,\infty)} c^{(k)}_{(V^{(k)},R^{(k)})}(x)\,dP^{(k)}_{(V^{(k)},R^{(k)})}(x) = \tfrac{1}{2}(c_1+c_2), \qquad k \geq 1.$$

The problem is that for the most simple cost structure the average cost functional is not necessarily continuous on $J[0,\infty)$. To provide an analysis for the average cost functional similar to that given in section 4.2 for the α-discounted cost functional we have to endow $J[0,\infty)$ with another metric under which the average cost functional becomes continuous.

Although it seems worthwhile to deduce general conditions which ensure that (4.4.1) holds, we have chosen another way in dealing with the average cost functional. In order to obtain sufficient conditions for the existence of (structured) stationary average optimal policies for a CTMDP we derive sufficient conditions which guarantee that a "limit point" of a sequence of α_k-discounted optimal policies with α_k converging to 0 is an average optimal policy for the CTMDP. When these conditions are satisfied we are able to carry forward structural results of optimal policies for a CTMDP from the discounted cost case to the average cost case. In this transition no discretization procedure is involved. Hence in obtaining structural results for an average optimal policy for a CTMDP we propose the following path: first of all analyse the discounted optimal policy for the k-th approximating DTMDP, then use theorem 4.2.9 to carry over structural results to the discounted optimal policy for the CTMDP and finally use the analysis of this section (theorem 4.4.6) to carry over structural results to the average optimal policy for the CTMDP.

DEFINITION 4.4.2. Let $(S,A_1,A_2,q,\Pi,p,c_1,c_2,f)$ be a CTMDP and P_0 an initial distribution on S. Assume that $c_1(.,.)$ is bounded from below and $c_2(.,.)$ non-negative. Let $\mathcal{R}$ denote a certain class of policies for this process. A policy (V^*,R^*) is called *average optimal in* $\mathcal{R}$ if for every policy $(V,R) \in \mathcal{R}$

$$\limsup_{t\to\infty} t^{-1} \int_{J[0,\infty)} c_{0,t,(V^*,R^*)}(x)\, dP_{(V^*,R^*)}(x) \le$$

$$\le \limsup_{t\to\infty} t^{-1} \int_{J[0,\infty)} c_{0,t,(V,R)}(x)\, dP_{(V,R)}(x).$$

We need the following well-known Abelian theorem.

THEOREM 4.4.3. Let w(.) be a non-negative, non-decreasing function on $[0,\infty)$ with $w(0) = 0$ and put

$$f(\alpha) := \int_0^\infty e^{-\alpha t} dw(t), \qquad \alpha > 0.$$

If $f(\alpha) < \infty$ for all $\alpha > 0$ then

$$\limsup_{\alpha\to 0^+} \alpha f(\alpha) \le \limsup_{t\to\infty} t^{-1} w(t) \tag{4.4.2}$$

and

$$\liminf_{\alpha\to 0^+} \alpha f(\alpha) \ge \liminf_{t\to\infty} t^{-1} w(t). \tag{4.4.3}$$

PROOF. (see also WIDDER (1946), page 181).
Since $f(\alpha) < \infty$ and $w(0) = 0$ we have

$$f(\alpha) = \alpha \int_0^\infty e^{-\alpha t} w(t)\,dt, \qquad \alpha > 0.$$

Hence

$$\alpha f(\alpha) \le \alpha^2 \int_0^T e^{-\alpha t} w(t)\,dt + \sup_{t\ge T} t^{-1} w(t) \qquad \text{for all } T > 0. \tag{4.4.4}$$

Since $f(\alpha) < \infty$ for all $\alpha > 0$ we can find a constant M such that for all $t>0$

$$w(t) \le Me^t$$

whence

$$\alpha^2 \int_0^T e^{-\alpha t} w(t)\;dt \le \alpha^2 M(e^{T(1-\alpha)} - 1)(1-\alpha)^{-1}. \tag{4.4.5}$$

From (4.4.4) and (4.4.5) follows that for all $T > 0$

$$\limsup_{\alpha\to 0^+} \alpha f(\alpha) \le \sup_{t\ge T} t^{-1} w(t)$$

By allowing T to become infinite we find (4.4.2). The proof of (4.4.3) is similar. □

REMARK 4.4.4. Relations (4.4.2) and (4.4.3) are also valid when $f(\alpha) = \infty$ for some $\alpha > 0$.

NOTATION 4.4.5. Let $(S,A_1,A_2,q,\Pi,p,c_1,c_2,f)$ be a CTMDP and P_0 an initial distribution on S. For any policy (V,R) we denote

$$c((V,R),\alpha) := \int_{J[0,\infty)} c_{\alpha,(V,R)}(x)\ dP_{(V,R)}(x)$$

(the *expected α-discounted costs under (V,R)*; see also definition 3.1.2);

$$c(V,R)(t) := \int_{J[0,\infty)} c_{0,t,(V,R)}(x)\ dP_{(V,R)}(x)$$

(the *expected total costs over* $[0,t]$ *under (V,R)*);

$$c(V,R) := \limsup_{t\to\infty} t^{-1} c(V,R)(t)$$

(the *average expected costs under (V,R)*).

THEOREM 4.4.6. Let $(S,A_1,A_2,q,\Pi,p,c_1,c_2,f)$ be a CTMDP, P_0 an initial distribution on S and $(\alpha_\ell)_{\ell=1}^{\infty}$ a sequence of positive real numbers converging to zero. Assume that $c_1(.,.)$ and $c_2(.,.)$ are non-negative. Suppose that for $\ell \geq 1$ the policy $(V^*_{\alpha_\ell}, R^*_{\alpha_\ell})$ is α_ℓ-discounted optimal in a given class $\mathcal{R}$. If there exists a policy (V^*,R^*) such that

$$(4.4.6) \qquad c(V^*_{\alpha_\ell}, R^*_{\alpha_\ell}) \to c(V^*,R^*) \qquad \text{as } \ell \to \infty$$

and if

(4.4.7) $\sup_{\ell} g_{\alpha_\ell}(T) \to 0 \quad$ as $T \to \infty$

where

$$g_{\alpha_\ell}(T) := c(V^*_{\alpha_\ell}, R^*_{\alpha_\ell}) - \inf_{t \geq T} t^{-1} c(V^*_{\alpha_\ell}, R^*_{\alpha_\ell})(t) \geq 0,$$

then (V^*, R^*) is average optimal in the class $\mathcal{R}$

PROOF. Let (V,R) be an arbitrary policy from $\mathcal{R}$, such that

(4.4.8) $c((V,R),\alpha) < \infty \quad$ for all $\alpha > 0$.

Since for all $x \in J[0,\infty)$ and all $\alpha > 0$

$$c_{\alpha,(V,R)}(x) = \int_{t=0}^{\infty} e^{-\alpha t} dc_{0,t,(V,R)}(x)$$

it follows that

$$c((V,R),\alpha) = \int_{t=0}^{\infty} e^{-\alpha t} dc(V,R)(t).$$

From theorem 4.4.3. we find with (4.4.8)

(4.4.9) $c(V,R) \geq \limsup_{\alpha \to 0^+} \alpha\, c((V,R),\alpha)$.

From the α_ℓ-discounted optimality of $(V^*_{\alpha_\ell}, R^*_{\alpha_\ell})$ in $\mathcal{R}$ follows

(4.4.10) $c((V,R),\alpha_\ell) \geq c((V^*_{\alpha_\ell}, R^*_{\alpha_\ell}),\alpha_\ell) \quad$ for $\ell \geq 1$.

Note that

$$\alpha_\ell c((V^*_{\alpha_\ell}, R^*_{\alpha_\ell}),\alpha_\ell) = \alpha_\ell \int_{t=0}^{\infty} e^{-\alpha_\ell t} dc(V^*_{\alpha_\ell}, R^*_{\alpha_\ell})(t) \geq$$

$$\geq \alpha_\ell^2 \int_{T}^{\infty} e^{-\alpha_\ell t} c(V^*_{\alpha_\ell}, R^*_{\alpha_\ell})(t)\,dt$$

for all $T > 0$.

Hence

$$\limsup_{\ell\to\infty} \alpha_\ell c((V^*_{\alpha_\ell},R^*_{\alpha_\ell}),\alpha_\ell) \geq$$
$$\geq \limsup_{\ell\to\infty} e^{-\alpha_\ell T}(1+\alpha_\ell T) \inf_{t\geq T} \{t^{-1}c(V^*_{\alpha_\ell},R^*_{\alpha_\ell})(t)\} =$$
$$= \limsup_{\ell\to\infty} \{c(V^*_{\alpha_\ell},R^*_{\alpha_\ell}) - g_{\alpha_\ell}(T)\} \geq$$
$$\geq \limsup_{\ell\to\infty} c(V^*_{\alpha_\ell},R^*_{\alpha_\ell}) - \sup_\ell g_{\alpha_\ell}(T).$$

Combining this with (4.4.7) we find by allowing T to become infinite

(4.4.11) $$\limsup_{\ell\to\infty} \alpha_\ell c((V^*_{\alpha_\ell},R^*_{\alpha_\ell}),\alpha_\ell) \geq \limsup_{\ell\to\infty} c(V^*_{\alpha_\ell},R^*_{\alpha_\ell}).$$

Combining (4.4.10) and (4.4.11) yields

(4.4.12) $$\limsup_{\ell\to\infty} \alpha_\ell c((V,R),\alpha_\ell) \geq \limsup_{\ell\to\infty} c(V^*_{\alpha_\ell},R^*_{\alpha_\ell}).$$

Finally (4.4.9), (4.4.12) and (4.4.6) yield

$$c(V,R) \geq c(V^*,R^*),$$

which proves the theorem. □

<u>REMARK 4.4.7 (i)</u>. From the proof of theorem 4.4.6 follows immediately that condition (4.4.6) can be relaxed to

$$\limsup_{\ell\to\infty} c(V^*_{\alpha_\ell},R^*_{\alpha_\ell}) \geq c(V^*,R^*)$$

(ii) It is worthwhile to notice that theorem 4.4.6 connects the α-discounted cost case with the average cost case without using any optimality equation.
(iii) The conditions of theorem 4.4.6 only concern the α_ℓ-discounted optimal policies which implies that all bad policies can be disregarded.

The verification of condition (4.4.6) is simple in those applications where the exact computation of $c(V^*_\alpha,R^*_\alpha)$ is possible. In chapters 5, 6 and 7 we give some examples of this situation.

Condition (4.4.7) however, is in general not easy to verify since the expected total costs over finite intervals are involved. In theorem 4.4.11 below we give sufficient conditions for (4.4.7) which are much easier to verify than (4.4.7) itself.
First we need some definitions.

DEFINITION 4.4.8. Let X be a random element of $J[0,\infty)$. A random element τ of $[0,\infty]$ is called a *stopping time for X* if

$$\{\omega\colon \tau(\omega) \le t\} \in \mathcal{B}_t \qquad \text{for all } t \ge 0,$$

where $\mathcal{B}_t$ is the product of the σ-algebra's $\{(\pi_s X)^{-1}\mathcal{S}\colon 0 \le s \le t\}$ (see also NEVEU (1965), page 99).

DEFINITION 4.4.9. Let X be a random element of $J[0,\infty)$, defined on $(\Omega,\mathcal{A},P)$. X is called a *regenerative stochastic process* if there exists a sequence $(\tau_n)_{n=0}^{\infty}$ of stopping times for X, such that

(i) $(\tau_n)_{n=0}^{\infty}$ is a renewal process

(ii) for any $n,m \in \mathbb{N}$, $t_1,\ldots,t_n \in [0,\infty)$ and any bounded function h defined on $(S^+)^n$:

$$E(h(\pi_{\tau_m+t_1}X,\ldots,\pi_{\tau_m+t_n}X)\,|\,\pi_t X,\ 0 \le t \le \tau_m) = E\,h(\pi_{t_1}X,\ldots,\pi_{t_n}X).$$

(see also ÇINLAR (1975), page 298).

REMARK 4.4.10.
(i) The sequence $(\tau_n)_{n=0}^{\infty}$ is called the *sequence of regeneration epochs*.
(ii) In definition 4.4.9. above the function $\pi_\infty(.)$ from $J[0,\infty)$ into S^+ is defined by $\pi_\infty(x) := \zeta$ for all $x \in J[0,\infty)$.
(iii) Since the function π from $J[0,\infty) \times [0,\infty]$ into S^+, defined by $\pi(x,t) := \pi_t x$ is measurable, it follows from proposition III.6.1 on page 101 of NEVEU (1965) that for any stopping time τ for X the mapping $\pi_\tau X$ from Ω into S^+ (defined by $\pi_\tau X(\omega) := \pi_{\tau(\omega)}X(\omega)$) is $\mathcal{B}_\tau$-measurable. Hence the expressions in definition 4.4.9 are well-defined.

THEOREM 4.4.11. Let $(S,A_1,A_2,q,\Pi,p,c_1,c_2,f)$ be a CTMDP and P_0 an initial distribution on S. Assume that $c_1(.,.)$ and $c_2(.,.)$ are non-negative. Suppose that for all $\alpha > 0$ the policy (V_α,R_α) induces on $J[0,\infty)$ a regenerative stochastic process X_α with sequence of regeneration epochs $(\tau_{n,\alpha})_{n=0}^{\infty}$, all defined on $(\Omega,A,\mathbb{P})$. Define for $n \geq 1$, $\alpha > 0$ and $\omega \in \Omega$

$$Y_{n,\alpha}(\omega) := c_{0,\tau_{n,\alpha}(\omega),(V_\alpha,R_\alpha)}(X_\alpha(\omega)) - c_{0,\tau_{n-1,\alpha}(\omega),(V_\alpha,R_\alpha)}(X_\alpha(\omega)).$$

Then $((V_\alpha,R_\alpha))_{\alpha>0}$ satisfies (4.4.7) for any sequence $(\alpha_\ell)_{\ell=1}^{\infty}$ converging to zero if the following three conditions are satisfied:

$$\text{(4.4.13)} \qquad \sup_\alpha \int_\Omega \tau_{1,\alpha}^2(\omega)\, d\mathbb{P}(\omega) < \infty$$

$$\text{(4.4.14)} \qquad \inf_\alpha \int_\Omega \tau_{1,\alpha}(\omega)\, d\mathbb{P}(\omega) > 0$$

$$\text{(4.4.15)} \qquad \sup_\alpha \int_\Omega Y_{1,\alpha}^2(\omega)\, d\mathbb{P}(\omega) < \infty.$$

PROOF. From definition 4.4.6. follows that $(Y_{n,\alpha})_{n=1}^{\infty}$ is a sequence of independent, identically distributed random elements of $[0,\infty]$ for all $\alpha > 0$. Denote for $t \geq 0$, $\alpha > 0$ and $\omega \in \Omega$

$$N_{t,\alpha}(\omega) := \sup\{n:\ \tau_{n,\alpha}(\omega) \leq t\}.$$

Then

$$\text{(4.4.16)} \qquad c(V_\alpha,R_\alpha)(t) \geq \int_\Omega \Big\{ \sum_{n=1}^{N_{t,\alpha}(\omega)+1} Y_{n,\alpha}(\omega) - Y_{N_{t,\alpha}(\omega)+1}(\omega) \Big\}\, d\mathbb{P}(\omega).$$

Since $N_{t,\alpha} + 1$ is a stopping time for $(Y_{n,\alpha})_{n=1}^{\infty}$ we find by Walds equation:

$$\text{(4.4.17)} \qquad \int_\Omega \sum_{n=1}^{N_{t,\alpha}(\omega)+1} Y_{n,\alpha}(\omega)\, d\mathbb{P}(\omega) = (M_\alpha(t)+1)\, EY_{1,\alpha}$$

where $M_\alpha(t) := EN_{t,\alpha}$, for $t \geq 0$ and $\alpha > 0$.

A well-known inequality in renewal theory states

$$M_\alpha(t) \geq \frac{t}{E\tau_{1,\alpha}} - 1 \tag{4.4.18}$$

and from the theory of regenerative processes we know by (4.4.13), (4.4.14) and (4.4.15) that

$$c(V_\alpha, R_\alpha) = \frac{EY_{1,\alpha}}{E\tau_{1,\alpha}} \tag{4.4.19}$$

Combining (4.4.16) upto (4.4.19) yields

$$t^{-1}c(V_\alpha, R_\alpha)(t) \geq c(V_\alpha, R_\alpha) - t^{-1}EY_{N_{t,\alpha}+1}. \tag{4.4.20}$$

Hence sufficient for (4.4.7) is

$$\sup_\alpha \sup_{t \geq T} t^{-1}EY_{N_{t,\alpha}+1} \to 0 \qquad \text{as } T \to \infty. \tag{4.4.21}$$

Put

$$f_\alpha(t) := EY_{N_{t,\alpha}+1}.$$

By conditioning on $\tau_{1,\alpha}$ we find

$$f_\alpha(t) = \int_0^t f_\alpha(t-s)\, d\mathbb{P}\tau_{1,\alpha}^{-1}(s) + \int_t^\infty E(Y_{1,\alpha} \mid \tau_{1,\alpha}=s)\, d\mathbb{P}\tau_{1,\alpha}^{-1}(s).$$

The solution of this renewal equation is

$$f_\alpha(t) = \int_0^t h_\alpha(t-s)\, dM_\alpha(s) + h_\alpha(t)$$

where

$$h_\alpha(t) := \int_t^\infty E(Y_{1,\alpha} \mid \tau_{1,\alpha} = s)\, d\mathbb{P}\tau_{1,\alpha}^{-1}(s).$$

Note that $h_\alpha(.)$ is a non-increasing function and by (4.4.15)

$$\sup_\alpha h_\alpha(0) = \sup_\alpha EY_{1,\alpha} < \infty.$$

Hence we conclude from

$$\sup_\alpha \sup_{t\geq T} t^{-1} EY_{N_{t,\alpha}+1} \leq T^{-1} \sup_\alpha h_\alpha(0) + T^{-1} \sup_\alpha \sup_{t\geq T} \int_0^t h_\alpha(t-s)\, dM_\alpha(s)$$

that sufficient for (4.4.21) is

$$(4.4.22) \qquad T^{-1} \sup_\alpha \sup_{t\geq T} \int_0^t h_\alpha(t-s)\ dM_\alpha(s) \to 0 \qquad \text{as } T \to \infty.$$

From the monotonicity of $h_\alpha(.)$ we find

$$(4.4.23) \qquad \int_0^t h_\alpha(t-s)\ dM_\alpha(s) \leq$$
$$\leq \sum_{n=0}^{\lceil t\rceil -1} h_\alpha(t-n-1)(M_\alpha(n+1)-M_\alpha(n)) + h_\alpha(0)(M_\alpha(t)-M_\alpha(\lceil t\rceil)).$$

From a result obtained by STONE (1972) follows that

$$(4.4.24) \qquad M_\alpha(t) \leq \frac{t}{E\tau_{1,\alpha}} + \frac{3E\tau_{1,\alpha}^2}{(E\tau_{1,\alpha})^2}.$$

Combining (4.4.18) and (4.4.24) yields

$$(4.4.25) \qquad M_\alpha(n+1) - M_\alpha(n) \leq 1 + \frac{1}{E\tau_{1,\alpha}} + \frac{3E\tau_{1,\alpha}^2}{(E\tau_{1,\alpha})^2}.$$

From (4.4.24) and (4.4.25) we conclude that for all $t \geq 0$

$$(4.4.26) \qquad \int_0^t h_\alpha(t-s)\ dM_\alpha(s) \leq$$
$$\leq \left(1 + \frac{1}{E\tau_{1,\alpha}} + \frac{3E\tau_{1,\alpha}^2}{(E\tau_{1,\alpha})^2}\right) \left(2h_\alpha(0) + \int_0^\infty h_\alpha(s)\ ds\right).$$

Finally we note that

$$\int_0^\infty h_\alpha(s)\ ds = E(Y_{1,\alpha}\tau_{1,\alpha})$$

which implies by Schwarz' inequality together with (4.4.13) and (4.4.15) that

$$(4.4.27) \qquad \sup_\alpha \int_0^\infty h_\alpha(s)\ ds < \infty.$$

Combining (4.4.26), (4.4.13), (4.4.14) and (4.4.27) yields (4.4.22) which completes the proof. □

Theorem 4.4.6 in combination with theorem 4.4.11 can be successfully applied to prove structural properties of average optimal policies for several CTMDP's. Examples will be given in the next three chapters.

CHAPTER 5

AN M/M/1 QUEUEING MODEL

5.1. INTRODUCTION AND ASSUMPTIONS.

Our first application is an M/M/1 queueing system with controllable arrival- and service rate and with infinite queue capacity. The decision maker can dynamically select the service rate in order to cope with random fluctuations in the arrival process. Moreover we assume that also the arrival rate of customers can be controlled by advertising or price adjustment.

The parameters of the process are specified as follows. In a service station with one server customers arrive according to a Poisson process with arrival rate ν, which can be varied within the interval $[\nu_1,\nu_2]$, where $0 \le \nu_1 \le \nu_2 < \infty$. The service times of the customers are independent random variables with negative exponential distribution with parameter μ when service rate μ is used. The service rate μ can be varied within the interval $[\mu_1,\mu_2]$ where $0 \le \mu_1 \le \mu_2 < \infty$. The cost structure consists of three parts: a holding cost rate $b(i)$ is incurred when i customers are in the system; there is an income rate $\tilde{b}_1(\nu)$ when arrival rate ν is maintained and a service cost rate $b_2(\mu)$ is incurred when service rate μ is used.

Semi-Markov versions of this model have been studied by CRABILL (1972), SABETI (1973), LOW (1974), LIPPMAN (1975) and SERFOZO (1981).

For this model we define the following CTMDP.

$$S: = \{0,1,2,\dots\}$$

$$A_1: = [\nu_1,\nu_2] \times [\mu_1,\mu_2]$$

$$A_2: = \phi$$

$$q(i,(\nu,\mu)): = \begin{cases} \nu+\mu, & i > 0 \\ & ; \ (\nu,\mu) \in A_1 \\ \nu, & i = 0 \end{cases}$$

(5.1.1)
$$\Pi(i,(\nu,\mu),j): = \begin{cases} \frac{\mu}{\nu+\mu}, & j = i - 1 \\ & ; \ 0 < i \in S; \ (\nu,\mu) \in A_1 \\ \frac{\nu}{\nu+\mu}, & j = i + 1 \end{cases}$$

$$\Pi(0,(\nu,\mu),1): = 1, \qquad (\nu,\mu) \in A_1$$

$$c_1(i,(\nu,\mu)): = b(i) - \tilde{b}_1(\nu) + b_2(\mu).$$

$$f(i,t): = i, \qquad t \geq 0, \ i \in S.$$

The state of ths system denotes the number of customers in the system. The set of controls is the cartesian product of the set of all possible arrival rates and the set of all possible service rates. Impulsive controls are not allowed. The cost rate consists of three additive parts: holding cost rate $b(i)$, income rate $\tilde{b}_1(\nu)$ and service cost rate $b_2(\mu)$. Finally the drift function is constant in time.

The following assumptions are made on the model parameters.

ASSUMPTION 5.1.1.

(i) $b(i): = bi, \ i \in S$ for some constant $b > 0$.

(ii) $\tilde{b}_1(.)$ and $b_2(.)$ are continuous and non-negative.

(iii) $\nu_2 \geq \mu_2 \geq \nu_1 \geq \mu_1$.

Assumptions 5.1.1.(i) and (iii) are made for ease of presentation but can easily be relaxed.

Note that by this assumption model (5.1.1) yields a well defined CTMDP, for which assumption 2.2.2. holds.

If we replace in (5.1.1) the income function $-\tilde{b}_1(.) \leq 0$ by

$$b_1(\nu): = -\tilde{b}_1(\nu) + \sup_{\nu\in[\nu_1,\nu_2]} \tilde{b}_1(\nu) \geq 0$$

then for every $x \in J[0,\infty)$ the α-discounted cost functional under (V,R) changes with the same amount $\alpha^{-1} \sup\{\tilde{b}_1(\nu): \nu_1 \leq \nu \leq \nu_2\}$. Similar assertions hold for the α-discounted, T-horizon and average cost functional. This implies that we can replace $-\tilde{b}_1(.)$ by $b_1(.)$, as far as the optimality of policies is concerned. Hence we will consider model (5.1.1) with $b_1(.) \geq 0$ instead of $-\tilde{b}_1(.) \leq 0$. For all $k > \nu_2 + \mu_2$ the k-th approximating DTMDP is defined by $(S, A_1, \emptyset, p_1^{(k)}, -, c_1^{(k)}, -, f, k)$, where

$$p_1^{(k)}(i,(\nu,\mu),j) := \begin{cases} k^{-1}\mu & j = i-1 \\ 1-k^{-1}(\nu+\mu), & j = i \\ k^{-1}\nu & j = i+1 \end{cases} \quad ; \; i > 0, \; (\nu,\mu) \in A_1$$

$$p_1^{(k)}(0,(\nu,\mu),j) := \begin{cases} 1 - k^{-1}\nu, & j = 0 \\ & \\ k^{-1}\nu \quad , & j = 1 \end{cases} \quad ; \; (\nu,\mu) \in A_1$$

$$c_1^{(k)}(i,(\nu,\mu)) := k^{-1}\{b(i) + b_1(\nu) + b_2(\mu)\}, \; i \in S, \; (\nu,\mu) \in A_1.$$

<u>REMARK 5.1.2</u>. Let (V,R) be a deterministic policy with $R = (R_1, R_2)$. Since $A_2 = \emptyset$ the impulsive control rule R_2 and the impulsive control set V are in fact irrelevant. Since A_1 is $[\nu_1,\nu_2] \times [\mu_1,\mu_2]$ any deterministic control rule R_1 is two-dimensional. Therefore we use in this chapter a slightly different notation. A deterministic policy will in this chapter be denoted by $R = (R_1, R_2)$, where R_i represents the i-th component of the control rule R, i = 1,2. In general a policy will be denoted by R (a transition probability from $J[0,\infty) \times [0,\infty)$ to A_1). The probability measure on $J[0,\infty)$ induced by a policy R is denoted by P_R (for the CTMDP) and by $P_R^{(k)}$ (for the k-th approximating DTMDP).

5.2. THE α-DISCOUNTED COST CASE.

<u>PROPOSITION 5.2.1</u>. Let P_0 be an arbitrary initial distribution on S. A policy R is regular if

(5.2.1) $\qquad (P_R \times \lambda)\{\mathrm{Disc}(R)\} = 0.$

<u>PROOF</u>. An immediate consequence of definition 2.4.7. and theorem 2.4.3. □

PROPOSITION 5.2.2. Let P_0 be an initial distribution on S with $\int_S i^2 dP_0(i) < \infty$. A policy $R = (R_1, R_2)$ is strong regular if (5.2.1) holds.

PROOF. Define for $k \geq \nu_2+\mu_2$ a policy $R^{(k)} = (R_1^{(k)}, R_2^{(k)})$ for the k-th approximating DTMDP by

$$R^{(k)}(x,t) := R(x,t), \qquad (x,t) \in J_k[0,\infty) \times L_k.$$

We will show that for R and $R^{(k)}$ the conditions of theorem 4.2.1. are fulfilled. Put for all $j \in S$

$$A_1(j) := A_1.$$

Then

$$R^{(k)}(x,t)(A_1^c(j)) = 0$$

for all $(x,t) \in J_k[0,\infty) \times L_k$, $k \geq 1$.
Note that condition (4.2.1) is trivially satisfied for all positive δ_0 and all k_0.
Choose $\alpha > 0$ and $\beta > 1$ such that (4.2.11) holds for some $\delta > 0$ (β may depend on α). Next we define

$$h_1(i) := \begin{cases} Mj_0, & 0 \leq i \leq j_0 \\ Mi, & i > j_0 \end{cases}$$

$$\ell(s) := 1, \qquad s \geq 0,$$

where $j_0 \in S$, such that $j_0 \geq (\sqrt{\beta} - 1)^{-1}$ and $M := b + \sup_\nu b_1(\nu) + \sup_\mu b_2(\mu)$.
We shall show that these choices satisfy the conditions (4.2.2) upto (4.2.10). The only condition that is not trivially satisfied is (4.2.6) for i=1.

For $j \geq 1$ we have

$$\int_S h_1^2(j_0)\, d\Pi(j,(\nu,\mu),j_0) = (\nu+\mu)^{-1}\{\nu h_1^2(j+1)+\mu h_1^2(j-1)\} =$$

$$= \begin{cases} M^2(j^2+1)+2M^2 j(\nu-\mu)(\nu+\mu)^{-1}, & j > j_0 \\ M^2 j_0^2 + (2j_0+1)\nu M^2(\nu+\mu)^{-1}, & j = j_0 \\ M^2 j_0^2, & 0 < j < j_0 \end{cases}$$

Hence

$$\sup_{(\nu,\mu)\in A_1(j)} \int_S h_1^2(j_0)\,d\Pi(j,(\nu,\mu),j_0) \le h_1^2(j)(1+j_0^{-1})^2 \le \beta h_1^2(j)$$

for all $j \in S$. □

DEFINITION 5.2.3. Let $R = (R_1,R_2)$ be a memoryless deterministic policy for the CTMDP or the k-th approximating DTMDP for some k. Hence there exist two functions $r_1(.,.)$ and $r_2(.,.)$ on $S \times [0,\infty)$ with values in $[\nu_1,\nu_2]$ and $[\mu_1,\mu_2]$ respectively such that $r_j(i,t) := R_j(x,t)$ for all $(x,t) \in J[0,\infty) \times [0,\infty)$ for which $\pi_t x = i$, $j=1,2$.

(i) The policy R is *monotone* if $r_1(.,t)$ is a monotone non-increasing function on S and $r_2(.,t)$ is a monotone non-decreasing function on S, for all $t \ge 0$.

(ii) The policy is of *bang-bang type* if it is monotone and $r_1(i,t) \in \{\nu_1,\nu_2\}$ and $r_2(i,t) \in \{\mu_1,\mu_2\}$ for all $(i,t) \in S \times [0,\infty)$.

THEOREM 5.2.4. Consider the k th approximating DTMDP with some initial distribution P_0 on S. Put $\alpha > 0$ and $\alpha_k := \exp(-\alpha k^{-1})$. For all $n \ge 1$ there exists a memoryless deterministic policy $R_n^{(k)} = (R_{n1}^{(k)},R_{n2}^{(k)})$ such that

(i) $R_n^{(k)}$ is α_k-discounted, n-horizon optimal in the class of all policies.

(ii) $R_n^{(k)}$ is monotone.

Moreover, if $b_1(.)$ and $b_2(.)$ are concave, then

(iii) $R_n^{(k)}$ is of bang-bang type.

PROOF. For $n \ge 1$ and $i \in S$ we denote

$$(5.2.2) \qquad f_{n,\alpha}^{(k)}(i) := \inf_{R\in\mathcal{R}} \int_{J_k[0,\infty)} c_{\alpha_k,n,R}^{(k)}(x)\, dP_R^{(k)}(x\,|\,\pi_0 x = i)$$

where $\mathcal{R}$ denotes the class of all policies.
Hence, $f_{n,\alpha}^{(k)}(i)$ is the minimal expected α_k-discounted costs over the n-horizon given that the initial state is i.
Define

$$f_{0,\alpha}^{(k)}(i) := 0 \qquad \text{for all } i \in S.$$

Then it follows by induction on n that for all $n \ge 1$

$$f_{n,\alpha}^{(k)}(i) = \inf_{(\nu,\mu)\in A_1} \{k^{-1}c_1(i,(\nu,\mu))+\alpha_k \int_S f_{n-1,\alpha}^{(k)}(j)\,dp_1^{(k)}(i,(\nu,\mu),j)\}.$$

For $i \in S$ and $n \geq 0$ we put

(5.2.3) $$v_{n,\alpha}^{(k)}(i) := f_{n,\alpha}^{(k)}(i) - f_{n,\alpha}^{(k)}((i-1) \vee 0)$$

and

$$w_n^{(1)}(i,\nu) := b_1(\nu) + \alpha_k \nu\, v_{n,\alpha}^{(k)}(i+1)$$

$$w_n^{(2)}(i,\mu) := b_2(\mu) - \alpha_k \mu\, v_{n,\alpha}^{(k)}(i).$$

Then

(5.2.4) $$f_{n,\alpha}^{(k)}(i) = $$

$$= k^{-1}b(i) + \alpha_k f_{n-1,\alpha}^{(k)}(i) + k^{-1}\inf_{\nu} w_{n-1}^{(1)}(i,\nu)+k^{-1}\inf_{\mu} w_{n-1}^{(2)}(i,\mu).$$

Define

(5.2.5) $$r_{n,\alpha,1}^{(k)}(i) := \inf\{\tilde{\nu}:\ w_{n-1}^{(1)}(i,\tilde{\nu}) = \inf_{\nu} w_{n-1}^{(1)}(i,\nu)\}$$

and

(5.2.6) $$r_{n,\alpha,2}^{(k)}(i) := \sup\{\tilde{\mu}:\ w_{n-1}^{(2)}(i,\tilde{\mu}) = \inf_{\mu} w_{n-1}^{(2)}(i,\mu)\}.$$

Finally we define for $(x,\ell k^{-1}) \in J_k[0,\infty) \times L_k$ and $j = 1,2$

$$R_{n,j}^{(k)}(x,\ell k^{-1}) := r_{n-\ell+1,\alpha,j}^{(k)}(\pi_{\ell k^{-1}}x), \quad 1 \leq \ell \leq n$$

$$R_{n,j}^{(k)}(x,\ell k^{-1}) := R_{n,j}^{(k)}(x,nk^{-1}), \quad \ell \geq n$$

Then $R_n^{(k)} = (R_{n,1}^{(k)}, R_{n,2}^{(k)})$ is a well defined memoryless deterministic policy which is α_k-discounted, n-horizon optimal in the class of all policies. This proves (i), while (ii) and (iii) follow from the next lemma. □

LEMMA 5.2.5. Let $f^{(k)}_{n,\alpha}(i)$, $r^{(k)}_{n,\alpha,1}(i)$ and $r^{(k)}_{n,\alpha,2}(i)$ be as given by (5.2.2), (5.2.5) and (5.2.6) in the proof of the previous theorem. Then

(i) $r^{(k)}_{n,\alpha,1}(.)$ is monotone non-increasing on S.

(ii) $r^{(k)}_{n,\alpha,2}(.)$ is monotone non-decreasing on S.

(iii) $f^{(k)}_{n,\alpha}(.)$ is convex on S.

Moreover, if $b_1(.)$ and $b_2(.)$ are concave then

(iv) $r^{(k)}_{n,\alpha,1}(.) \in \{\nu_1,\nu_2\}$ and $r^{(k)}_{n,\alpha,2}(.) \in \{\mu_1,\mu_2\}$.

PROOF. The proof proceeds by induction on n. For n = 1 we have

$$r^{(k)}_{1,\alpha,1}(i) = \inf\{\tilde{\nu}: b_1(\tilde{\nu}) = \inf_{\nu} b_1(\nu)\}$$

$$r^{(k)}_{1,\alpha,2}(i) = \sup\{\tilde{\mu}: b_2(\tilde{\mu}) = \inf_{\mu} b_2(\mu)\}$$

$$f^{(k)}_{1,\alpha}(i) = k^{-1}b(i) + k^{-1}\inf_{\nu} b_1(\nu) + k^{-1}\inf_{\mu} b_2(\mu)$$

which yields (i) upto (iv) for n = 1.
Suppose the statement is true for n-1. Then we have

$$w^{(1)}_{n-1}(i,\nu) - w^{(1)}_{n-1}(i,\tilde{\nu}) =$$

$$= b_1(\nu) - b_1(\tilde{\nu}) + \alpha_k(\nu-\tilde{\nu})\; v^{(k)}_{n-1,\alpha}(i+1).$$

Hence we conclude from the convexity of $f^{(k)}_{n-1,\alpha}(.)$ that

(5.2.7) $\qquad w^{(1)}_{n-1}(.,\nu) - w^{(1)}_{n-1}(.,\tilde{\nu})$ is a non-increasing function for $\nu < \tilde{\nu}$.

Let $i_1 < i_2$ and assume that $\nu := r^{(k)}_{n,\alpha,1}(i_1) < r^{(k)}_{n,\alpha,1}(i_2) =: \tilde{\nu}$. Then

$$w^{(1)}_{n-1}(i_1,\nu) - w^{(1)}_{n-1}(i_1,\tilde{\nu}) \le 0$$

and

$$w^{(1)}_{n-1}(i_2,\nu) - w^{(1)}_{n-1}(i_2,\tilde{\nu}) > 0$$

which contradicts (5.2.7). This proves (i) for n.

Similarly we have

$$w_{n-1}^{(2)}(i,\mu) - w_{n-1}^{(2)}(i,\tilde{\mu}) =$$

$$= b_2(\mu) - b_2(\tilde{\mu}) - \alpha_k(\mu-\tilde{\mu})\ v_{n-1,\alpha}^{(k)}(i),$$

which implies by the induction hypothesis that

$$(5.2.8) \qquad w_{n-1}^{(2)}(.,\mu) - w_{n-1}^{(2)}(.,\tilde{\mu}) \text{ is a non-increasing function for } \mu > \tilde{\mu}.$$

Let $i_1 < i_2$ and assume that $\mu := r_{n,\alpha,2}^{(k)}(i_1) > r_{n,\alpha,2}^{(k)}(i_2) =: \tilde{\mu}$. Then

$$w_{n-1}^{(2)}(i_1,\mu) - w_{n-1}^{(2)}(i_1,\tilde{\mu}) \le 0$$

and

$$w_{n-1}^{(2)}(i_2,\mu) - w_{n-1}^{(2)}(i_2,\tilde{\mu}) > 0,$$

which contradicts (5.2.8). This yields (ii) for n.
From (5.2.4) follows that sufficient for the convexity of $f_{n,\alpha}^{(k)}(.)$ is the convexity of

$$\alpha_k f_{n-1,\alpha}^{(k)}(.) + k^{-1} \inf_{\nu} w_{n-1}^{(1)}(.,\nu) + k^{-1} \inf_{\mu} w_{n-1}^{(2)}(.,\mu).$$

Put

$$g(i) = \tfrac{1}{2}\ \alpha_k\ f_{n-1,\alpha}^{(k)}(i) + k^{-1} \inf_{\nu} w_{n-1}^{(1)}(i,\nu)$$

and

$$\tilde{g}(i) = \tfrac{1}{2}\ \alpha_k\ f_{n-1,\alpha}^{(k)}(i) + k^{-1} \inf_{\mu} w_{n-1}^{(2)}(i,\mu).$$

We shall show that both functions $g(.)$ and $\tilde{g}(.)$ are convex.

$$g(i+1) - g(i) \ge \tfrac{1}{2}\alpha_k\ v_{n-1,\alpha}^{(k)}(i+1) + k^{-1}\alpha_k r_{n,\alpha,1}^{(k)}(i+1) v_{n-1,\alpha}^{(k)}(i+2) -$$

$$+ k^{-1}\alpha_k\ r_{n,\alpha,1}^{(k)}(i+1)\ v_{n-1,\alpha}^{(k)}(i+1).$$

$$g(i) - g(i-1) \le \tfrac{1}{2}\alpha_k\ v_{n-1,\alpha}^{(k)}(i) + k^{-1}\alpha_k r_{n,\alpha,1}^{(k)}(i-1) v_{n-1,\alpha}^{(k)}(i+1) -$$

$$+ k^{-1}\alpha_k\ r_{n,\alpha,1}^{(k)}(i-1)\ v_{n-1,\alpha}^{(k)}(i).$$

These two inequalities together yield

$$g(i+1) - 2g(i) + g(i-1) \geq$$

$$\geq (\tfrac{1}{2}\alpha_k - k^{-1}\alpha_k r^{(k)}_{n,\alpha,1}(i-1)\{v^{(k)}_{n-1,\alpha}(i+1)-v^{(k)}_{n-1,\alpha}(i)\} +$$

$$+ k^{-1}\alpha_k r^{(k)}_{n,\alpha,1}(i+1)\{v^{(k)}_{n-1,\alpha}(i+2)-v^{(k)}_{n-1,\alpha}(i+1)\} \geq 0 .$$

The last inequality follows from the induction hypothesis for k large enough. The convexity of $\tilde{g}(.)$ is established in a similar way. This completes the proof of (iii) for n.

Finally we assume that $b_1(.)$ and $b_2(.)$ are concave. Hence $w^{(1)}_{n-1}(i,.)$ and $w^{(2)}_{n-1}(i,.)$ are concave for all $i \in S$, which implies that $r^{(k)}_{n,\alpha,1}(i)\in\{\nu_1,\nu_2\}$ and $r^{(k)}_{n,\alpha,2}(i)\in\{\mu_1,\mu_2\}$ for all $i \in S$. □

<u>LEMMA 5.2.6</u>. If $k \geq \nu_2+\mu_2$ then $v^{(k)}_{n,\alpha}(i) \geq v^{(k)}_{n-1,\alpha}(i) \geq 0$ for $i \in S$ and $n \geq 1$.

<u>PROOF</u>. Since $v^{(k)}_{0,\alpha}(i) = 0$ and $v^{(k)}_{1,\alpha}(i) = k^{-1}\ b$ for $i > 0$ the statement is true for n = 1. Suppose it is true for n - 1.

For $i \in S$ we have from (5.2.3) and (5.2.4)

$$v^{(k)}_{n,\alpha}(i+1) \geq k^{-1}b + \alpha_k v^{(k)}_{n-1,\alpha}(i+1) + k^{-1}b_1(r^{(k)}_{n,\alpha,1}(i+1)) +$$

$$+ \alpha_k k^{-1} r^{(k)}_{n,\alpha,1}(i+1)\ v^{(k)}_{n-1,\alpha}(i+2)+k^{-1}b_2(r^{(k)}_{n,\alpha,2}(i+1)) -$$

$$+ \alpha_k k^{-1} r^{(k)}_{n,\alpha,2}(i+1)\ v^{(k)}_{n-1,\alpha}(i+1)-k^{-1}b_1(r^{(k)}_{n-1,\alpha,1}(i)) -$$

$$+ \alpha_k k^{-1} r^{(k)}_{n-1,\alpha,1}(i)\ v^{(k)}_{n-1,\alpha}(i+1)-k^{-1}b_2(r^{(k)}_{n-1,\alpha,2}(i)) +$$

$$+ \alpha_k k^{-1} r^{(k)}_{n-1,\alpha,2}(i)\ v^{(k)}_{n-1,\alpha}(i).$$

$$v^{(k)}_{n-1,\alpha}(i+1) \leq k^{-1}b + \alpha_k v^{(k)}_{n-2,\alpha}(i+1)+k^{-1}b_1(r^{(k)}_{n,\alpha,1}(i+1)) +$$

$$+ \alpha_k k^{-1} r^{(k)}_{n,\alpha,1}(i+1)\ v^{(k)}_{n-2,\alpha}(i+2)+k^{-1}b_2(r^{(k)}_{n,\alpha,2}(i+1)) -$$

$$+ \alpha_k k^{-1} r^{(k)}_{n,\alpha,2}(i+1)\ v^{(k)}_{n-2,\alpha}(i+1)-k^{-1}b_1(r^{(k)}_{n-1,\alpha,1}(i)) -$$

$$+ \alpha_k k^{-1} r^{(k)}_{n-1,\alpha,1}(i)\ v^{(k)}_{n-2,\alpha}(i+1)-k^{-1}b_2(r^{(k)}_{n-1,\alpha,2}(i)) +$$

$$+ \alpha_k k^{-1} r^{(k)}_{n-1,\alpha,2}(i)\; v^{(k)}_{n-2,\alpha}(i).$$

Hence

$$v^{(k)}_{n,\alpha}(i+1) - v^{(k)}_{n-1,\alpha}(i+1) \geq$$

$$\geq \alpha_k k^{-1} r^{(k)}_{n,\alpha,1}(i+1)\; \{v^{(k)}_{n-1,\alpha}(i+2) - v^{(k)}_{n-2,\alpha}(i+2)\} +$$

$$+ \alpha_k \{1 - k^{-1} r^{(k)}_{n,\alpha,2}(i+1) - k^{-1} r^{(k)}_{n-1,\alpha,1}(i)\}\{v^{(k)}_{n-1,\alpha}(i+1) - v^{(k)}_{n-2,\alpha}(i+1)\} +$$

$$+ \alpha_k k^{-1} r^{(k)}_{n-1,\alpha,2}(i)\; \{v^{(k)}_{n-1,\alpha}(i) - v^{(k)}_{n-2,\alpha}(i)\}.$$

Application of the induction hypothesis completes the proof. □

<u>THEOREM 5.2.7</u>. Consider the k th approximating DTMDP with some initial distribution P_0 on S. Put $\alpha > 0$ and $\alpha_k := \exp(-\alpha k^{-1})$. There exists a stationary deterministic policy $R^{(k)}_* = (R^{(k)}_{*1}, R^{(k)}_{*2})$ such that

(i) $R^{(k)}_*$ is α_k-discounted optimal in the class of all policies.

(ii) $R^{(k)}_*$ is monotone.

Moreover, if $b_1(.)$ and $b_2(.)$ are concave, then

(iii) $R^{(k)}_*$ is of bang-bang type.

<u>PROOF</u>. Consider the sequences of functions $(r^{(k)}_{n,\alpha,1}(.))^{\infty}_{n=1}$ and $(r^{(k)}_{n,\alpha,2}(.))^{\infty}_{n=1}$ defined by (5.2.5) and (5.2.6).

By the well known diagonal procedure we can construct a subsequence $(n_\ell)^{\infty}_{\ell=1}$ of $\mathbb{N}$, such that

$$(5.2.9) \qquad r^{(k)}_{\alpha,1}(i) := \lim_{\ell\to\infty} r^{(k)}_{n_\ell,\alpha,1}(i)$$

and

$$(5.2.10) \qquad r^{(k)}_{\alpha,2}(i) := \lim_{\ell\to\infty} r^{(k)}_{n_\ell,\alpha,2}(i)$$

exist for all $i \in S$.

Put for $i \in S$

$$r^{(k)}_{\alpha}(i) := (r^{(k)}_{\alpha,1}(i),\; r^{(k)}_{\alpha,2}(i))$$

and for $(x,t) \in J_k[0,\infty) \times L_k$

$$R_*^{(k)}(x,t) := r_\alpha^{(k)}(\pi_t x).$$

Then $R_*^{(k)}$ is a well defined stationary deterministic policy. Obviously $R_*^{(k)}$ is monotone and of bang-bang type when $b_1(.)$ and $b_2(.)$ are concave. We complete the proof by showing that $R_*^{(k)}$ is α_k-discounted optimal in the class $\mathcal{R}$ of all policies.

Define for $i \in S$

$$(5.2.11) \qquad f_\alpha^{(k)}(i) := \inf_{R \in \mathcal{R}} \int_{J_k[0,\infty)} c_{\alpha_k,R}^{(k)}(x) \; dP_R^{(k)}(x \mid \pi_0 x = i).$$

Since $(f_{n,\alpha}^{(k)}(.))_{n=1}^{\infty}$ is a non-decreasing sequence of non-negative functions on S

$$w(i) := \lim_{n\to\infty} f_{n,\alpha}^{(k)}(i) \geq 0$$

exists for all $i \in S$.
The inequality

$$f_{n,\alpha}^{(k)}(i) \leq f_\alpha^{(k)}(i) \qquad \text{for } i \in S$$

implies

$$(5.2.12) \qquad w(i) \leq f_\alpha^{(k)}(i) \qquad \text{for } i \in S.$$

On the other side we have by (5.2.4)

$$(5.2.13) \qquad f_{n,\alpha}^{(k)}(i) = k^{-1}b(i) + \alpha_k f_{n-1,\alpha}^{(k)}(i) + k^{-1}b_1(r_{n,\alpha,1}^{(k)}(i)) +$$
$$+ k^{-1}\alpha_k r_{n,\alpha,1}^{(k)}(i) \; v_{n-1,\alpha}^{(k)}(i+1) + k^{-1}b_2(r_{n,\alpha,2}^{(k)}(i)) -$$
$$+ k^{-1}\alpha_k r_{n,\alpha,2}^{(k)}(i) \; v_{n-1,\alpha}^{(k)}(i).$$

Taking limits in (5.2.13) through $(n_\ell)_{\ell=1}^{\infty}$ we conclude by the continuity of $b_1(.)$ and $b_2(.)$ that

$$(5.2.14)\qquad w(i) = k^{-1}c_1(i,r_\alpha^{(k)}(i)) + \alpha_k \int_S w(j)\ dp_1^{(k)}(i,r_\alpha^{(k)}(i),j).$$

Iteration of (5.2.14) yields for all $n \geq 1$ and all $i \in S$

$$w(i) \geq \int_{J_k[0,\infty)} c^{(k)}_{\alpha_k,n,R_*^{(k)}}(x)\ dP^{(k)}_{R_*^{(k)}}(x|\pi_0 x=i).$$

By letting $n \to \infty$, we get for all $i \in S$

$$(5.2.15)\qquad w(i) \geq \int_{J_k[0,\infty)} c^{(k)}_{\alpha_k,R_*^{(k)}}(x)\,dP^{(k)}_{R_*^{(k)}}(x|\pi_0 x=i) \geq f_\alpha^{(k)}(i).$$

Combining (5.2.12) and (5.2.15) yields

$$(5.2.16)\qquad w(i) = f_\alpha^{(k)}(i) \qquad \text{for all } i \in S,$$

which implies by (5.2.15) the α_k-discounted optimality of $R_*^{(k)}$ in $\mathcal{R}$. □

Before we consider the CTMDP we need two lemma's.

LEMMA 5.2.8. Let $r=(r_1,r_2)$ be a measurable function from $S\times L_k$ into A_1 and P_0 an initial distribution on S, defined by $P_0(\{i\}) = 1$ for some $i \in S$. Suppose there are given pure memoryless policies $R^{(k)}$, $R_*^{(k)}$ and $\tilde{R}_*^{(k)}$ for the k-th approximating DTMDP, defined for $(x,t) \in J_k[0,\infty) \times L_k$ by

$$R^{(k)}(x,t) := r(\pi_t x,t)$$
$$R_*^{(k)}(x,t) := (\nu_1,\mu_1)$$
$$\tilde{R}_*^{(k)}(x,t) := (\nu_2,\mu_2)$$

(i) For any non-decreasing function f on $[0,\infty]$

$$\int_{J_k[0,\infty)} f(T_2(x))\ dP^{(k)}_{R^{(k)}}(x) \geq \int_{J_k[0,\infty)} f(T_2(x))\ dP^{(k)}_{\tilde{R}_*^{(k)}}(x).$$

(ii) For any non-increasing function f on $[0,\infty]$

$$\int_{J_k[0,\infty)} f(T_2(x))\,dP^{(k)}_{R^{(k)}}(x) \geq \int_{J_k[0,\infty)} f(T_2(x))\,dP^{(k)}_{R_*^{(k)}}(x).$$

PROOF.

(i) $$P^{(k)}_{R^{(k)}}\{x: T_2(x)>nk^{-1}\} = \prod_{\ell=1}^{n}\{1-k^{-1}(r_1(i,\ell k^{-1})+r_2(i,\ell k^{-1}))\} \geq$$

$$\geq \{1-k^{-1}(\nu_2+\mu_2)\}^n =$$

$$= P^{(k)}_{\tilde{R}^{(k)}}\{x: T_2(x)>nk^{-1}\}.$$

The lemma now follows from a well-known theorem on stochastic ordering (see STOYAN (1977), page 5).

(ii) The proof of (ii) proceeds similarly. □

LEMMA 5.2.9. Let $b_1(.)$ and $b_2(.)$ be concave. Then there exist $j^{(k)}_{\alpha,1} \in S$ and $j^{(k)}_{\alpha,2} \in S$ such that the policy $R^{(k)}_* = (R^{(k)}_{*1}, R^{(k)}_{*2})$ defined by

$$R^{(k)}_{*1}(x,t) = \begin{cases} \nu_1 & \text{if } \pi_t x > j^{(k)}_{\alpha,1} \\ \nu_2 & \text{if } \pi_t x \leq j^{(k)}_{\alpha,1} \end{cases}$$

$$R^{(k)}_{*2}(x,t) = \begin{cases} \mu_1 & \text{if } \pi_t x \leq j^{(k)}_{\alpha,2} \\ \mu_2 & \text{if } \pi_t x > j^{(k)}_{\alpha,2} \end{cases}$$

is α_k-discounted optimal in the class of all policies. Moreover, there exist $\alpha_0 > 0$, $k_0 \in \mathbb{N}$ and $j \in S$ such that

$$j^{(k)}_{\alpha,\ell} \leq j$$

for all $0 < \alpha \leq \alpha_0$, all $k \geq k_0$; $\ell = 1,2$.

PROOF. The first part of the lemma follows from (iii) of theorem 5.2.7. Denote

$$v_\alpha^{(k)}(i) := f_\alpha^{(k)}(i) - f_\alpha^{(k)}((i-1) \vee 0),\ i \in S$$

$$w^{(1)}(i,\nu) := b_1(\nu) + \alpha_k \nu\, v_\alpha^{(k)}(i+1)$$

$$w^{(2)}(i,\mu) := b_2(\mu) - \alpha_k \mu\, v_\alpha^{(k)}(i).$$

Using the fact that $\lim_{n\to\infty} v_{n,\alpha}^{(k)}(i) = v_\alpha^{(k)}(i)$ and the continuity of $b_1(.)$ and $b_2(.)$ it follows from (5.2.5) and (5.2.6) that

$$w^{(1)}(i, r_{\alpha,1}^{(k)}(i)) = \inf_\nu w^{(1)}(i,\nu)$$

$$w^{(2)}(i, r_{\alpha,2}^{(k)}(i)) = \inf_\mu w^{(2)}(i,\mu)\ .$$

Hence

$$r_{\alpha,1}^{(k)}(i) = \nu_1 \quad \text{if} \quad v_\alpha^{(k)}(i+1) > \frac{b_1(\nu_1) - b_1(\nu_2)}{\alpha_k(\nu_2 - \nu_1)}$$

and

$$r_{\alpha,2}^{(k)}(i) = \mu_2 \quad \text{if} \quad v_\alpha^{(k)}(i) > \frac{b_2(\mu_2) - b_2(\mu_1)}{\alpha_k(\mu_2 - \mu_1)}\ .$$

This implies that it is sufficient to show that for every constant $c > 0$ there exists an $\alpha_0 > 0$, $k_0 \in \mathbb{N}$ and $j \in S$ such that

(5.2.17) $\quad v_\alpha^{(k)}(j) \geq c$

for all $0 < \alpha \leq \alpha_0$ and all $k \geq k_0$.

For $j \in S$ and $n \geq 1$ we find by conditioning on T_2 (the epoch of the first jump)

$$v_{nk,\alpha}^{(k)}(j+1) \geq$$

$$\geq k^{-1} b \sum_{m=1}^{nk} \alpha_k^m \prod_{\ell=1}^{nk} \{1-k^{-1}(r_{nk-\ell,\alpha,1}^{(k)}(j+1)+r_{nk-\ell,\alpha,2}^{(k)}(j+1))\} +$$

$$+ \sum_{m=1}^{nk} \prod_{\ell=1}^{m-1} \{1-k^{-1}(r_{nk-\ell,\alpha,1}^{(k)}(j+1)+r_{nk-\ell,\alpha,2}^{(k)}(j+1))\}\ .$$

$$.\ [k^{-1} r_{nk-m,\alpha,1}^{(k)}(j+1)\{k^{-1} b \sum_{\ell=1}^{m-1} \alpha_k^\ell + \alpha_k^m\, v_{nk-m,\alpha}^{(k)}(j+2)\} +$$

$$+ k^{-1} r^{(k)}_{nk-m,\alpha,2}(j+1)\ \{k^{-1} b \sum_{\ell=1}^{m-1} \alpha_k^{\ell} + \alpha_k^m\, v^{(k)}_{nk-m,\alpha}(j)\}].$$

This inequality yields with the convexity of $f^{(k)}_{n,\alpha}(.)$

$$v^{(k)}_{nk,\alpha}(j+1) \geq$$

$$\geq k^{-1} b \sum_{m=1}^{nk} \alpha_k^m \prod_{\ell=1}^{nk} \{1-k^{-1}(r^{(k)}_{nk-\ell,\alpha,1}(j+1)+r^{(k)}_{nk-\ell,\alpha,2}(j+1))\} +$$

$$+ \sum_{m=1}^{nk} \prod_{\ell=1}^{m-1} \{1-k^{-1}(r^{(k)}_{nk-\ell,\alpha,1}(j+1)+r^{(k)}_{nk-\ell,\alpha,2}(j+1))\}\ .$$

$$.\{k^{-1} b \sum_{\ell=1}^{m-1} \alpha_k^{\ell}+\alpha_k^m\, v^{(k)}_{nk-m,\alpha}(j)\}k^{-1}\{r^{(k)}_{nk-m,\alpha,1}(j+1)+r^{(k)}_{nk-m,\alpha,2}(j+1)\}.$$

Note that the function $f(.)$ from $\mathbb{N}$ into $[0,\infty)$, defined by

$$f(m) := \begin{cases} k^{-1} b \sum_{\ell=1}^{m-1} \alpha_k^{\ell} & \text{for } 1 \leq m \leq nk \\ k^{-1} b \sum_{\ell=1}^{nk} \alpha_k^{\ell} & \text{for } m > nk \end{cases}$$

is non-decreasing on $\mathbb{N}$, while the function $g(.)$ defined by

$$g(m) := \begin{cases} \alpha_k^m\, v^{(k)}_{nk-m,\alpha}(j) & \text{for } 1 \leq m \leq nk \\ 0 & \text{for } m > nk \end{cases}$$

is, by lemma 5.2.6., non-increasing on $\mathbb{N}$.

Hence application of lemma 5.2.8. on the last inequality yields

$$v^{(k)}_{nk,\alpha}(j+1) \geq k^{-1} b\ \{1 - k^{-1}(\nu_2+\mu_2)\}^{nk} \sum_{m=1}^{nk} \alpha_k^m +$$
$$+ k^{-1} b \sum_{m=1}^{nk} k^{-1}(\nu_2+\mu_2)\{1-k^{-1}(\nu_2+\mu_2)\}^{m-1} \sum_{\ell=1}^{m-1} \alpha_k^{\ell} +$$
$$+ \sum_{m=1}^{nk} \alpha_k^m\, v^{(k)}_{nk-m,\alpha}(j)\ k^{-1}(\nu_1+\mu_1)\ \{1-k^{-1}(\nu_1+\mu_1)\}^{m-1}. \tag{5.2.18}$$

After some algebra we find from (5.2.18) for all $0 < \delta \leq n$

$$(5.2.19)\qquad v^{(k)}_{nk,\alpha}(j+1) \ge (1 - e^{-\alpha n})\ b\alpha_k\ \{1 - k^{-1}(\nu_2+\mu_2)\}\ (\alpha + \nu_2 + \mu_2)^{-1} + \\ + \sum_{m=1}^{k\delta} \alpha_k^m\ v^{(k)}_{nk-m,\alpha}(j)\ k^{-1}(\nu_1+\mu_1)\ \{1 - k^{-1}(\nu_1+\mu_1)\}^{m-1}.$$

By letting $n \to \infty$ we find with lemma 5.2.6. for all $\delta > 0$

$$v^{(k)}_{\alpha}(j+1) \ge b\alpha_k\ \{1 - k^{-1}(\nu_2 + \mu_2)\}\ (\alpha + \nu_2 + \mu_2)^{-1} + \\ + v^{(k)}_{\alpha}(j) \sum_{m=1}^{k\delta} \alpha_k^m\ k^{-1}(\nu_1 + \mu_1)\ \{1 - k^{-1}(\nu_1 + \mu_1)\}^{m-1} \ge \\ \ge b\alpha_k\ \{1 - k^{-1}(\nu_2 + \mu_2)\}\ (\alpha + \nu_2 + \mu_2)^{-1} + \\ + v^{(k)}_{\alpha}(j)\ \alpha_k\ (\nu_1 + \mu_1)(1 - e^{-\alpha\delta})(\alpha + \nu_1 + \mu_1)^{-1}.$$

Hence there exists a number $k_0 \in \mathbb{N}$ such that

$$(5.2.20)\qquad v^{(k)}_{\alpha}(j+1) \ge \frac{b}{2(\alpha+\nu_2+\mu_2)} + v^{(k)}_{\alpha}(j)\ \alpha_k\ \frac{\nu_1+\mu_1}{\nu_1+\mu_1+\alpha}\ (1 - e^{-\alpha\delta})$$

for all $j \in S$, all $k \ge k_0$, all $\alpha > 0$ and all $\delta > 0$.

Iteration of (5.2.20) yields

$$v^{(k)}_{\alpha}(j+1) \ge \frac{b}{2(\alpha+\nu_2+\mu_2)} \sum_{\ell=0}^{j} \left\{\frac{\nu_1+\mu_1}{\nu_1+\mu_1+\alpha}\ \alpha_k\ (1 - e^{-\alpha\delta})\right\}^{\ell},$$

which implies for every constant $c > 0$ the existence of an $\alpha_0 > 0$, $k_0 \in \mathbb{N}$ and $j \in S$ such that (5.2.17) holds. □

<u>THEOREM 5.2.10</u>. Consider the CTMDP defined by (5.1.1) with some initial distribution P_0 on S. For all $\alpha > 0$ there exists a stationary deterministic policy R such that

(i) R is α-discounted optimal in the class of strong regular policies.

(ii) R is monotone.

Moreover if $b_1(.)$ and $b_2(.)$ are concave, then

(iii) R is of bang-bang type.

PROOF. Consider the sequences of functions $(r_{\alpha,1}^{(k)}(.))_{k=1}^{\infty}$ and $(r_{\alpha,2}^{(k)}(.))_{k=1}^{\infty}$ defined by (5.2.9) and (5.2.10).

By the diagonal procedure we construct a subsequence $(k_\ell)_{\ell=1}^{\infty}$ of $\mathbb{N}$ such that for all $i \in S$

(5.2.21) $$r_{\alpha,1}(i) := \lim_{\ell\to\infty} r_{\alpha,1}^{(k_\ell)}(i)$$

and

(5.2.22) $$r_{\alpha,2}(i) := \lim_{\ell\to\infty} r_{\alpha,2}^{(k_\ell)}(i)$$

exist.

Put for $i \in S$

$$r_\alpha(i) := (r_{\alpha,1}(i),\ r_{\alpha,2}(i))$$

and for $(x,t) \in J[0,\infty) \times [0,\infty)$

$$R(x,t) := r_\alpha(\pi_t x)$$

Then R is a well-defined stationary deterministic policy which is monotone and of bang-bang type if $b_1(.)$ and $b_2(.)$ are concave. The α-discounted optimality of R in the class of strong regular policies follows from theorem 4.2.10. since the third set of conditions of this theorem is fulfilled. □

5.3. THE AVERAGE COST CASE.

THEOREM 5.3.1. Consider the CTMDP defined by (5.1.1) with some initial distribution P_0 on S. Assume that $b_1(.)$ and $b_2(.)$ are concave. There exists a stationary deterministic policy R such that

(i) R is average optimal in the class of strong regular policies.

(ii) R is of bang-bang type.

PROOF. From theorem (5.2.10) follows for all $\alpha > 0$ the existence of $j_{\alpha,1} \in S$ and $j_{\alpha,2} \in S$ such that the policy $R_\alpha = (R_{\alpha 1}, R_{\alpha 2})$

$$R_{\alpha 1}(x,t) := \begin{cases} \nu_1 & \text{if} \quad \pi_t x > j_{\alpha,1} \\ \nu_2 & \text{if} \quad \pi_t x \le j_{\alpha,1} \end{cases}$$

and

$$R_{\alpha 2}(x,t) := \begin{cases} \mu_1 & \text{if} \quad \pi_t x \le j_{\alpha,2} \\ \mu_2 & \text{if} \quad \pi_t x > j_{\alpha,2} \end{cases}$$

is α-discounted optimal in the class of strong regular policies. From lemma 5.2.9. follows that $(j_{\beta_k,1})_{k=1}^{\infty}$ and $(j_{\beta_k,2})_{k=1}^{\infty}$ are bounded for any sequence $(\beta_k)_{k=1}^{\infty}$ with $\lim_{k\to\infty} \beta_k = 0$.
Hence there exist $j_1 \in S$ and $j_2 \in S$ such that

$$j_\ell = \lim_{k\to\infty} j_{\beta_k,\ell} \qquad ,\ell = 1,2$$

for some sequence $(\beta_k)_{k=1}^{\infty}$ with $\lim_{k\to\infty} \beta_k = 0$.
Define the policy $R = (R_1, R_2)$ by

$$(5.3.1) \qquad R_1(x,t) := \begin{cases} \nu_1 & \text{if} \quad \pi_t x > j_1 \\ \nu_2 & \text{if} \quad \pi_t x \le j_1 \end{cases}$$

and

$$(5.3.2) \qquad R_2(x,t) := \begin{cases} \mu_1 & \text{if} \quad \pi_t x \le j_2 \\ \mu_2 & \text{if} \quad \pi_t x > j_2 \end{cases}$$

then

$$c(R_{\beta_k}) \to c(R) \qquad \text{as } k \to \infty,$$

since the state space S is discrete.

The theorem now follows from theorems 4.4.6. and 4.4.11. and lemma 5.3.2. below. (Note that from lemma 5.3.2. follows that the conditions of theorem 4.4.11. are fulfilled, since the state space is discrete). □

<u>LEMMA 5.3.2.</u> Consider the bang-bang type policy R defined by (5.3.1) and (5.3.2). This policy induces on $J[0,\infty)$ a regenerative stochastic process X. Moreover

$$(5.3.3) \qquad \int_{J[0,\infty)} \tau_1^2(x)\, dP_R(x) < \infty,$$

$$(5.3.4)\qquad \int_{J[0,\infty)} \tau_1(x)\ dP_R(x) > 0$$

and

$$(5.3.5)\qquad \int_{J[0,\infty)} Y_1^2(x)\ dP_R(x) < \infty$$

where $(\tau_n)_{n=0}^{\infty}$ is the sequence of regeneration epochs of X and Y_n is defined in theorem 4.4.11.(Note that X, τ_n and Y_n are supposed to be defined on $J[0,\infty)$ for all $n \geq 1$).

PROOF. Assume without loss of generality that $j_2 > j_1$ and $P_0\{j_2\} = 1$. (The proof for the case in which $j_2 \leq j_1$ proceeds similarly).
Define for $n \geq 1$ and $x \in J[0,\infty)$

$$\tau_n(x) := \text{time of } n^{th} \text{ entrance of the path x into the state } j_2.$$

Then X is a regenerative stochastic process with sequence of regeneration epochs $(\tau_n)_{n=0}^{\infty}$

Obviously

$$\int_{J[0,\infty)} \tau_1(x)\ dP_R(x) \geq (\nu_1 + \mu_1)^{-1} > 0,$$

which implies (5.3.4). Moreover, we find by conditioning on $S_2(x)$, which denotes the state after the first jump

$$(5.3.6)\qquad \int_{J[0,\infty)} \tau_1^2(x)\,dP_R(x) = \frac{\nu_1}{\nu_1+\mu_1}\int_{J[0,\infty)} \tau_1^2(x)\,dP_R(x|S_2(x)=j_2+1) + \frac{\mu_1}{\nu_1+\mu_1}\int_{J[0,\infty)} \tau_1^2(x)\ dP_R(x|S_2(x)=j_2-1).$$

A well-known result from queueing theory states

(5.3.7)
$$\int_{J[0,\infty)} \tau_1^2(x)\,dP_R(x|S_2(x)=j_2+1) = \frac{2}{(\nu_1+\mu_2)^2} + \frac{2}{\nu_1+\mu_2} EB + EB^2 < \infty$$

where B denotes the length of a busy cycle in an M/M/1 queue with parameters ν_1 and μ_2.

Moreover,

(5.3.8)
$$\int_{J[0,\infty)} \tau_1^2(x)\,dP_R(x|S_2(x)=j_2-1) \leq \frac{2}{(\nu_1+\mu_1)^2} + \frac{2}{\nu_1+\mu_1} ET + ET^2 < \infty$$

where T denotes the length of the time interval in which a random walk with reflecting barrier 0 moves from j_2-1 to j_2.
Together (5.3.6), (5.3.7) and (5.3.8) yield (5.3.3).

Finally

(5.3.9)
$$\int_{J[0,\infty)} Y_1^2(x)\,dP_R(x) = \frac{\nu_1}{\nu_1+\mu_1} \int_{J[0,\infty)} Y_1^2(x)\,dP_R(x|S_2(x)=j_2+1) + \frac{\mu_1}{\nu_1+\mu_1} \int_{J[0,\infty)} Y_1^2(x)\,dP_R(x|S_2(x)=j_2-1),$$

while

(5.3.10)
$$\int_{J[0,\infty)} Y_1^2(x)\;dP_R(x|S_2(x)=j_2-1) \leq$$
$$\leq \gamma^2 \int_{J[0,\infty)} \tau_1^2(x)\;dP_R(x|S_2(x)=j_2-1) < \infty$$

and

(5.3.11)
$$\int_{J[0,\infty)} Y_1^2(x)\;dP_R(x|S_2(x)=j_2+1) \leq$$
$$\leq \gamma^2 \int_{J[0,\infty)} (N(x)+1)^2\,\tau_1^2(x)\;dP_R(x|S_2(x)=j_2+1) < \infty,$$

where

$$\gamma := bj_2 + \sup_\nu b_1(\nu) + \sup_\mu b_2(\mu)$$

and

$$N(x) := \text{number of upward jumps in the path } x \text{ during the time interval } [0,\tau_1(x)].$$

Note that

$$\int_{J[0,\infty)} N^k(x)\ dP_R\ (x|S_2(x)=j_2+1)$$

is the k-th moment of the number of customers served in a busy cycle in an M/M/1 queue with parameters ν_1 and μ_2.
Since the right hand sides of (5.3.10) and (5.3.11) are finite (5.3.5) follows from (5.3.9). □

CHAPTER 6

A MAINTENANCE REPLACEMENT MODEL

Our second application is a maintenance replacement model which has been used earlier in this monograph as illustration of the conditions of theorem 2.4.3 (see example 2.5.1). For completeness we describe the model here once more.

A device is subject to shocks which occur randomly in time according to a Poisson process with rate ν. This rate of the shock arrival process can be controlled by maintenance. Every shock causes independently of the others a certain amount of damage. The total damage incurred by the device accumulates additively. The amount of damage caused by a single shock is a random variable with known distribution function F. Besides the damage caused by shocks the device decays continuously at constant rate $\sigma > 0$, i.e. between shocks the total damage increases linearly at rate σ. The decision maker has control on the system in two different ways: the arrival rate ν of the shocks can be chosen arbitrarily in the interval $[\nu_1,\nu_2]$, where $0 < \nu_1 \leq \nu_2 < \infty$. On the other hand the device can be replaced at every moment by a new one without damage.

The cost structure consists of three parts: an operating cost rate b(s) is incurred when the total damage equals s, a maintenance cost rate $c(\nu)$ is incurred when the shock arrival rate is ν and a lump cost $M > 0$ is associated with every replacement.

Versions of this model have been treated by TAYLOR (1975) and ZUCKERMAN (1977). For this model we define the following CTMDP.

$$
\begin{aligned}
&S: = [0,\infty) \\
&A_1: = [\nu_1,\nu_2] \\
&A_2: = \{\mu\} \\
&q(s,\nu): = \nu, \qquad (s,\nu) \in S \times A_1 \\
(6.1.1)\qquad &\Pi(s,\nu,[s,s+t]): = F(t), \qquad (s,\nu) \in S \times A_1 \text{ and } t \geq 0 \\
&p(s,\mu,\{0\}): = 1, \qquad s \in S \\
&c_1(s,\nu): = b(s) + c(\nu), \qquad (s,\nu) \in S \times A_1 \\
&c_2(s,\mu): = M, \qquad s \in S \\
&f(s,t) = s + \sigma t, \qquad (s,t) \in S \times [0,\infty).
\end{aligned}
$$

The state of the system denotes the accumulated damage of the device under operation. The control set is the set of all possible shock arrival rates. The only possible impulsive control is replacement denoted by μ, where we choose $\mu > \nu_2$. We make the following assumptions on the model parameters.

ASSUMPTION 6.1.1.

(i) F has a density function.

(ii) b(.) is continuous, non-decreasing, non-negative and concave on S.

(iii) $b'(0) < \infty$.

(iv) $\gamma: = \int_0^\infty b(t)\ dF(t) < \infty$ and $\gamma^{(2)}: = \int_0^\infty b^2(t)\ dF(t) < \infty$.

(v) c(.) is continuous and non-negative on A_1.

For $k > \nu_2$ the k-th approximating DTMDP is defined by $(S,A_1,A_2,p_1^{(k)},p_2^{(k)},c_1^{(k)},c_2,f,k)$, with

$$
\begin{aligned}
&p_1^{(k)}(s,\nu,\{s\}) = 1 - k^{-1}\nu, \qquad (s,\nu) \in S \times A_1 \\
&p_1^{(k)}(s,\nu,[s,s+t]) = 1 - k^{-1}\nu + k^{-1}\nu F(t), \qquad (s,\nu) \in S \times A_1,\ t\geq 0 \\
&p_2^{(k)}(s,\mu,\{0\}) = 1, \qquad s \in S \\
&c_1^{(k)}(s,\nu) = k^{-1}\{b(s) + c(\nu)\}, \qquad (s,\nu) \in S \times A_1
\end{aligned}
$$

REMARK 6.1.2. (i) Since $A_2 = \{\mu\}$ it follows for every policy (V,R) for this CTMDP with $R = (R_1,R_2)$ that $R_2(x,t)$ is a probability measure degenerated in $\{\mu\}$ for all (x,t). Hence we denote in this chapter a policy for the CTMDP with (V,R), where V denotes a closed subset of S and R a transition probability from $J[0,\infty) \times [0,\infty)$ to A_1.

(ii) Let R be a transition probability from $J[0,\infty) \times [0,\infty)$ to A_1, such that $R(.,t)$ is r_t-measurable. From the definition of a policy (definition 2.2.3) follows that (V,R) is a policy iff V is closed and $0 \notin V$.

6.2. THE α-DISCOUNTED COST CASE.

PROPOSITION 6.2.1. Let P_0 be an arbitrary initial distribution on S. A policy (V,R) is regular if

(6.2.1) $\qquad \delta V$ is countable.

(6.2.2) $\qquad (P_{(V,R)} \times \lambda)\ \{\mathrm{Disc}(R)\} = 0.$

PROOF. A consequence of theorem 2.4.3. and assumption 6.1.1.(i). □

PROPOSITION 6.2.2. Let P_0 be an initial distribution on S with

(6.2.3) $$\int_0^{\infty} b^2(s)\ dP_0(s) < \infty.$$

Then the policy (V,R) is strong regular if (6.2.1) and (6.2.2) hold.

PROOF. Let (V,R) be a policy for which (6.2.1) and (6.2.2) hold. Define for $k > \nu_2$ policies $(V^{(k)},R^{(k)})$ for the k-th approximating DTMDP by $V^{(k)} = V$ and $R^{(k)} = R$. We will show that (V,R) and $(V^{(k)},R^{(k)})$ satisfy the conditions of theorem 4.2.1.

First of all we note that condition (2.4.3) follows from (6.2.1) and assumption 6.1.1.(i), that condition (2.4.4) is trivially satisfied, while (2.4.5) follows from (6.2.2). Condition (4.2.1) holds since $0 \notin V$. Put for all $j \in S$

$$A_1(j) := A_1$$

$$A_2(j) := A_2$$

Choose $\alpha > 0$ and $\beta > 1$ (depending on α) such that (4.2.11) holds for some $\delta < \delta_0$. Without loss of generality we assume that

$$b(0) = 0$$

and

$$\lim_{s\to\infty} b(s) = \infty. \tag{6.2.4}$$

(Note that adding a finite constant to the function $c_1(.,.)$ does not affect the validity of theorem 4.2.1. On the other hand the verification of the conditions for strong regularity becomes rather easy when $b(.)$ is bounded).

Define

$$h_1(s) := \begin{cases} a\, b(s_0), & \text{for } 0 \le s \le s_0 \\ a\, b(s), & \text{for } s > s_0 \end{cases}$$

$$h_2(s) := M, \qquad \text{for } s \in S$$

$$\ell(t) := (1 + \frac{b(\sigma t)}{b(s_0)})^2 \qquad \text{for } t \ge 0,$$

where $a := 1 + \sup_\nu c(\nu)$ and $s_0 \in S$ such that $b(s_0) \ge 1$,

$$\alpha b(s_0) \ge 2\sigma b'(0) \tag{6.2.5}$$

and

$$1 + \frac{2\gamma}{b(s_0)} + \frac{\gamma^{(2)}}{b^2(s_0)} \le \beta. \tag{6.2.6}$$

We shall show that these choices satisfy the conditions (4.2.2) upto (4.2.10).

The verification of (4.2.2) is straightforward. From the concavity of $b(.)$ and $b(0) = 0$ follows for $(j,t) \in S \times [0,\infty)$

$$b(f(j,t)) \le b(j) + b(\sigma t)$$

which implies

$$h_1^2(f(j,t)) \le h_1^2(j)\ (1 + \frac{b(\sigma t)}{b(s_0)})^2.$$

This yields with the definition of $\ell(.)$ condition (4.2.3).
Conditions (4.2.4) and (4.2.5) are implied by the concavity of b(.) and relation (6.2.5).
From the concavity of b(.) also follows

$$\sup_{\nu} \int_S h_1^2(j_0)\ d\Pi(s,\nu,j_0) \le h_1^2(s)\ \{1 + \frac{2\gamma}{b(s_0)} + \frac{\gamma^{(2)}}{b^2(s_0)}\},$$

which together with (6.2.6) yields condition 4.2.6. Condition 4.2.7. and (4.2.10) follow from the fact that

$$\int_S h_1^2(j_0)\ dp(s,\mu,j_0) = h_1^2(0) = 0.$$

Finally (4.2.8) follows from (6.2.3), while (4.2.9) follows from the continuity and concavity of b(.). □

DEFINITION 6.2.3. Let (V,R) be a deterministic memoryless policy for the CTMDP. Hence there exists a function r(.,.) on $S \times [0,\infty)$ with values in $[\nu_1,\nu_2]$ such that $r(s,t) = R(x,t)$ for all $(x,t) \in J[0,\infty) \times [0,\infty)$ with $\pi_t x = s$.

(i) The policy (V,R) is of *control limit type* if there exists a number $D > 0$ such that $V = [D,\infty)$.

(ii) The policy (V,R) is *monotone* if it is of control limit type and if r(.,t) is a monotone non-decreasing function on S with values in $[\nu_1,\nu_2]$, for all $t \ge 0$.

(iii) The policy (V,R) is of *bang-bang type* if it is monotone and $r(s,t) \in \{\nu_1,\nu_2\}$ for all $(s,t) \in S \times [0,\infty)$.

REMARK 6.2.4. In order to analyse the k-th approximating DTMDP for a finite horizon it is useful to consider these discrete time processes from a slightly different point of view than we did so far.
Let us consider the k-th approximating DTMDP as a classical discrete time Markov decision process (cf. remark 2.3.2) $(S,A,p^{(k)},c^{(k)})$ with

$$A = A_1 \cup A_2$$

$$p^{(k)}(s,a,B) = \begin{cases} p_1^{(k)}(s,a,B+\sigma k^{-1}), & a \in A_1 \\ p_2^{(k)}(s,a,B+\sigma k^{-1}), & a \in A_2 \end{cases}$$

$$c^{(k)}(s,a) = \begin{cases} c_1^{(k)}(s,a), & a \in A_1 \\ c_2(s,a), & a \in A_2. \end{cases}$$

A policy for this discrete time Markov decision process is denoted by R, a transition probability from $J_k[0,\infty) \times L_k$ to $A_1 \cup A_2$, such that $R(.,t)$ is r_t-measurable. Note that this definition of a policy is less restrictive than definition 2.3.5.

DEFINITION 6.2.5. Let R be a deterministic, memoryless policy for the k-th approximating DTMDP. Hence there exists a function $r(.,.)$ on $S \times [0,\infty)$ with values in $[\nu_1,\nu_2] \cup \{\mu\}$ such that $r(s,t) = R(x,t)$ for all $(x,t) \in J_k[0,\infty) \times L_k$ with $\pi_t x = s$.

(i) The policy R is of *control limit type* if there exists for all $t \geq 0$ a number $D(t) > 0$ such that $r(s,t) = \mu$ iff $s \geq D(t)$.

(ii) The policy R is *monotone* if $r(.,t)$ is monotone non-decreasing for all $t \geq 0$ as a function on S with values in $[\nu_1,\nu_2] \cup \{\mu\}$.

(iii) The policy R is of *bang-bang type* if it is monotone and $r(s,t) \in \{\nu_1,\nu_2,\mu\}$ for all $(s,t) \in S \times [0,\infty)$.

Note that a monotone policy is always of control limit type since we have chosen $\mu > \nu_2$.

THEOREM 6.2.6. Consider the k-th approximating DTMDP with some initial distribution P_0 on S. Put $\alpha > 0$ and $\alpha_k := \exp(-\alpha k^{-1})$. For all $n \geq 1$ there exists a memoryless deterministic policy $R_n^{(k)}$ such that

(i) $R_n^{(k)}$ is α_k-discounted, n-horizon optimal in the class of all policies.

(ii) $R_n^{(k)}$ is monotone.

Moreover, if c(.) is concave, then

(iii) $R_n^{(k)}$ is of bang-bang type.

PROOF. For $n \geq 1$ and $s \in S$ we denote

$$f_{n,\alpha}^{(k)}(s) := \inf_{R\in\mathcal{R}} \int_{J_k[0,\infty)} c_{\alpha_k,n,R}(x)\ dP_R(x|\pi_0 x = s),$$

where $\mathcal{R}$ denotes the class of all policies.

Define

$$f_{0,\alpha}^{(k)}(s) := 0 \qquad \text{for all } s \in S.$$

Then we find, by induction on n, for all $n \geq 1$

$$(6.2.7) \qquad f_{n,\alpha}^{(k)}(s) = \min \{M + \alpha_k f_{n-1,\alpha}^{(k)}(\sigma k^{-1}),\ \inf_{\nu} \{k^{-1}b(s) + k^{-1}c(\nu) + \alpha_k(1-k^{-1}\nu)\ f_{n-1,\alpha}^{(k)}(s+\sigma k^{-1}) + \alpha_k k^{-1}\nu \int_0^\infty f_{n-1,\alpha}^{(k)}(s + \sigma k^{-1} + t)\ dF(t)\}\}.$$

Put

$$w_n(s,\nu) := c(\nu) + \nu\alpha_k \int_0^\infty \{f_{n,\alpha}^{(k)}(s+\sigma k^{-1}+t) - f_{n,\alpha}^{(k)}(s+\sigma k^{-1})\}dF(t).$$

Hence

$$(6.2.8) \qquad f_{n,\alpha}^{(k)}(s) = \min \{M+\alpha_k f_{n-1,\alpha}^{(k)}(\sigma k^{-1}),\ k^{-1}b(s)+\alpha_k f_{n-1,\alpha}^{(k)}(s+\sigma k^{-1})+k^{-1} \inf_{\nu} w_{n-1}(s,\nu)\}.$$

Define

$$(6.2.9) \qquad r_{n,\alpha}^{(k)}(s) := \begin{cases} \mu & \text{if } M + \alpha_k f_{n-1,\alpha}^{(k)}(\sigma k^{-1}) \leq k^{-1}b(s) + \alpha_k f_{n-1,\alpha}^{(k)}(s+\sigma k^{-1})+k^{-1} \inf_{\nu} w_{n-1}(s,\nu) \\ \inf \{\tilde{\nu}: w_{n-1}(s,\tilde{\nu}) = \inf_{\nu} w_{n-1}(s,\nu)\} & \text{otherwise.} \end{cases}$$

Finally we define for $(x,\ell k^{-1}) \in J_k[0,\infty) \times L_k$

$$R_n^{(k)}(x,\ell k^{-1}) := r_{n-\ell+1,\alpha}^{(k)}(\pi_{\ell k^{-1}}x), \qquad 1 \le \ell \le n$$

$$R_n^{(k)}(x,\ell k^{-1}) := R_n^{(k)}(x,nk^{-1}), \qquad \ell \ge n.$$

From lemma 6.2.7.(i) below follows that $R_n^{(k)}(.,.)$ is a well-defined policy. Obviously $R_n^{(k)}(.,.)$ is memoryless, deterministic and α_k-discounted optimal in the class of all policies. The statements (ii) and (iii) follow from the next lemma. □

<u>LEMMA 6.2.7</u>. Let $f_{n,\alpha}^{(k)}(.)$ and $r_{n,\alpha}^{(k)}(.)$ be as given by (6.2.7) and (6.2.9) in the proof of the previous theorem. Then

(i) $r_{n,\alpha}^{(k)}(.)$ is non-decreasing on S.

(ii) $f_{n,\alpha}^{(k)}(.)$ is concave and non-decreasing on S.

Moreover, if c(.) is concave on A_1, then

(iii) $r_{n,\alpha}^{(k)}(s) \in \{\nu_1,\nu_2,\mu\}$ for all $s \in S$.

<u>PROOF</u>. The proof proceeds by induction on n. For n = 1 we have

$$r_{1,\alpha}^{(k)}(s) = \begin{cases} \mu & \text{if} \quad M<k^{-1}b(s)+k^{-1}\inf_\nu c(\nu) \\ \inf \{\tilde{\nu}: c(\tilde{\nu}) = \inf_\nu c(\nu)\} & \text{otherwise.} \end{cases}$$

$$f_{1,\alpha}^{(k)}(s) = \min \{M, k^{-1}b(s) + k^{-1} \inf_\nu c(\nu)\}.$$

This implies by the concavity and monotonicity of b(.), the statements (i), (ii) and (iii) for n = 1.

Suppose that the statements are true for n-1.

Note that

$$w_{n-1}(s,\nu) - w_{n-1}(s,\tilde{\nu}) = c(\nu) - c(\tilde{\nu}) +$$

$$+ \alpha_k(\nu-\tilde{\nu})\int_0^\infty \{f^{(k)}_{n-1,\alpha}(s+\sigma k^{-1}+t) - f^{(k)}_{n-1,\alpha}(s+\sigma k^{-1})\}\, dF(t).$$

Hence, by the concavity of $f^{(k)}_{n-1,\alpha}(.)$,

(6.2.10) $\quad w_{n-1}(.,\nu) - w_{n-1}(.,\tilde{\nu})$ is non-increasing if $\nu > \tilde{\nu}$.

Let $s_1 < s_2$ and assume that $\nu: = r^{(k)}_{n,\alpha}(s_1) > r^{(k)}_{n,\alpha}(s_2) =: \tilde{\nu}$ with $\nu \neq \mu$. Then

$$w_{n-1}(s_1,\nu) - w_{n-1}(s_1,\tilde{\nu}) < 0$$

and

$$w_{n-1}(s_2,\nu) - w_{n-1}(s_2,\tilde{\nu}) \geq 0$$

which contradicts (6.2.10).
Moreover, we conclude from the induction hypothesis, that

$$k^{-1}b(s) + \alpha_k\, f^{(k)}_{n-1,\alpha}(s+\sigma k^{-1}) + k^{-1} \inf_{\nu} w_{n-1}(s,\nu)$$

is non-decreasing in s.
Hence $r^{(k)}_{n,\alpha}(.)$ is non-decreasing on S.
Since the infimum of concave functions is concave, we conclude from (6.2.7) and the induction hypothesis that $f^{(k)}_{n,\alpha}(.)$ is non-decreasing and concave.
Finally, if c(.) is concave, then $w_{n-1}(s,.)$ is concave, which implies that $r^{(k)}_{n,\alpha}(s) \in \{\nu_1,\nu_2,\mu\}$ for all $s \in S$. □

For the analysis with respect to the α_k-discounted cost functional for the k th approximating DTMDP we need the following lemma.

LEMMA 6.2.8. Let B be a compact subset of $\mathbb{R}$. For any sequence $(f_k(.))_{k=1}^{\infty}$ of non-decreasing functions from $\mathbb{R}$ into B, there exists a non-decreasing function f(.) from $\mathbb{R}$ into B and a sequence of natural numbers $(k_j)_{j=1}^{\infty}$ such that

$$\lim_{j\to\infty} f_{k_j}(s) = f(s), \qquad \text{for all } s \in \mathbb{R}.$$

PROOF. Let E be a countable dense subset of $\mathbb{R}$. By the diagonal procedure we can construct a sequence $(\ell_j)_{j=1}^{\infty}$ of natural numbers, such that

$$f_1(s) := \lim_{j\to\infty} f_{\ell_j}(s)$$

exists, for all $s \in E$.

Define

$$f_2(s) := \inf_{\substack{t\in E\\ t>s}} f_1(t), \qquad s \in \mathbb{R}.$$

Then $f_2(.)$ is obviously non-decreasing on $\mathbb{R}$.

Moreover, choose $s \notin \mathrm{Disc}(f_2)$ and $\varepsilon > 0$. Then there exist $t_1, t_2 \in E$, such that

$$t_1 < s < t_2$$

and

$$f_2(s) - \varepsilon < f_1(t_1) \le f_1(t_2) < f_2(s) + \varepsilon.$$

Since

$$f_{\ell_j}(t_1) \le f_{\ell_j}(s) \le f_{\ell_j}(t_2), \qquad \text{for all } j \ge 1$$

we conclude that

$$\limsup_{j\to\infty} f_{\ell_j}(s) \le f_2(s) + \varepsilon$$

and

$$\liminf_{j\to\infty} f_{\ell_j}(s) \ge f_2(s) - \varepsilon.$$

Hence

$$\lim_{j\to\infty} f_{\ell_j}(s) = f_2(s), \qquad \text{for all } s \notin \mathrm{Disc}(f_2).$$

Since $\text{Disc}(f_2)$ is countable we can construct a subsequence $(k_j)_{j=1}^{\infty}$ of $(\ell_j)_{j=1}^{\infty}$ such that $\lim_{j\to\infty} f_{k_j}(s)$ exists for all $s \in \text{Disc}(f_2)$. Hence

$$f(s) := \lim_{j\to\infty} f_{k_j}(s),$$

exists for all $s \in \mathbb{R}$ and obviously $f(.)$ is non-decreasing. □

<u>THEOREM 6.2.9</u>. Consider the k th approximating DTMDP with some initial distribution P_0 on S. Put $\alpha > 0$ and $\alpha_k := \exp(-\alpha k^{-1})$. There exists a stationary deterministic policy $R_*^{(k)}$ such that
(i) $R_*^{(k)}$ is α_k-discounted optimal in the class of all policies.
(ii) $R_*^{(k)}$ is monotone.
Moreover, if $c(.)$ is concave, then
(iii) $R_*^{(k)}$ is of bang-bang type.

<u>PROOF</u>. Consider the sequence of functions $(r_{n,\alpha}^{(k)}(.))_{n=1}^{\infty}$ on S defined by (6.2.9). From part (i) of lemma 6.2.7. follows that these functions are non-decreasing on S with values in $[\nu_1,\nu_2]\cup\{\mu\}$. By lemma 6.2.8. there exists a sequence $(n_j)_{j=1}^{\infty}$ of natural numbers, such that

$$r_\alpha^{(k)}(s) := \lim_{j\to\infty} r_{n_j,\alpha}^{(k)}(s) \tag{6.2.11}$$

exists for all $s \in S$.
Put

$$R_*^{(k)}(x,t) := r_\alpha^{(k)}(\pi_t x), \qquad \text{for } (x,t) \in J_k[0,\infty) \times L_k. \tag{6.2.12}$$

Then $R_*^{(k)}(.,.)$ is a well-defined stationary deterministic policy, which is obviously monotone and of bang-bang type if $c(.)$ is concave.
What remains to prove is that $R_*^{(k)}(.,.)$ is α_k-discounted optimal.
Define for $s \in S$

$$f_\alpha^{(k)}(s) := \inf_{R\in\mathcal{R}} \int_{J_k[0,\infty)} c_{\alpha_k,R}(x)\, dP_R(x|\pi_0 x = s).$$

Since $(f_{n,\alpha}^{(k)}(.))_{n=1}^{\infty}$ is a non-decreasing sequence of non-negative functions on S,

$$w(s) := \lim_{n\to\infty} f_{n,\alpha}^{(k)}(s) \geq 0$$

exists for all $s \in S$.

The inequality

$$f_{n,\alpha}^{(k)}(s) \leq f_{\alpha}^{(k)}(s), \qquad s \in S$$

implies

$$(6.2.13) \qquad w(s) \leq f_{\alpha}^{(k)}(s) \qquad \text{for all } s \in S.$$

On the other hand we have

$$(6.2.14) \qquad f_{n,\alpha}^{(k)}(s) = c^{(k)}(s, r_{n,\alpha}^{(k)}(s)) + \alpha_k \int_S f_{n-1,\alpha}^{(k)}(u)\, dp^{(k)}(s, r_{n,\alpha}^{(k)}(s), u).$$

By plugging in the definition of $p^{(k)}(.,.,.)$ in (6.2.14) and taking limits through $(n_j)_{j=1}^{\infty}$ we find with the monotone convergence theorem

$$(6.2.15) \qquad w(s) = c^{(k)}(s, r_{\alpha}^{(k)}(s)) + \alpha_k \int_S w(u)\, dp^{(k)}(s, r_{\alpha}^{(k)}(s), u).$$

Iteration of (6.2.15) yields for all $n \geq 1$ and $s \in S$

$$w(s) \geq \int_{J[0,\infty)} c_{\alpha_k, n, R_*^{(k)}}(x)\, dP_{R_*^{(k)}}(x \mid \pi_0 x = s).$$

By letting $n \to \infty$ we get for all $s \in S$

$$(6.2.16) \qquad w(s) \geq \int_{J[0,\infty)} c_{\alpha_k, R_*^{(k)}}(x)\, dP_{R_*^{(k)}}(x \mid \pi_0 x = s) \geq f_{\alpha}^{(k)}(s).$$

Combining (6.2.13) and (6.2.16) yields the α_k-discounted optimality of $R_*^{(k)}$ in $\mathcal{R}$. □

<u>THEOREM 6.2.10</u>. Let $r_{\alpha}^{(k)}(s)$ be defined by (6.2.11) and put

$$(6.2.17) \qquad D_{\alpha}^{(k)} := \inf\{s \in S : r_{\alpha}^{(k)}(s) = \mu\}.$$

There exist an $\alpha_0 > 0$, a natural number k_0 and a number $D_0 > 0$ such that

$$D_\alpha^{(k)} \geq D_0,$$

for all $0 < \alpha \leq \alpha_0$ and all $k \geq k_0$.

PROOF. Without loss of generality we assume that $P_0(\{0\}) = 1$. Consider the policy $R^{(k)}$ defined for $t \in L_k$ by

$$R^{(k)}(x,t) := \begin{cases} \mu & \text{if } \pi_t x \neq \pi_{t-}x \\ \nu_2 & \text{otherwise.} \end{cases}$$

Denote

$$c_0 := \int_{J_k[0,\infty)} c_{\alpha_k, R^{(k)}}(x)\ dP_{R^{(k)}}(x).$$

By conditioning on the arrival time of the first shock we find

$$(6.2.18) \qquad c_0 = \sum_{\ell=1}^{\infty} (1-k^{-1}\nu_2)^{\ell-1} k^{-1}\nu_2 [k^{-1} \sum_{m=1}^{\ell} \alpha_k^m \{c(\nu_2)+b(m\sigma k^{-1})\} + \alpha_k^{\ell+1}(M+c_0)].$$

From (6.2.18) we find after some algebra

$$(6.2.19) \qquad c_0 \leq e^{\nu_2+\alpha} \int_0^\infty e^{-\nu_2 t} \{c(\nu_2)+b(\sigma t)\}dt + M + \frac{c_0\ \alpha_k k^{-1}\nu_2}{1-\alpha_k(1-k^{-1}\nu_2)}.$$

Put

$$c_1 := e^{\nu_2+1} \int_0^\infty e^{-\nu_2 t} \{c(\nu_2)+b(\sigma t)\}dt + M,$$

then we conclude from the concavity of b(.) and assumption 6.1.1.(iii) that $c_1 < \infty$. Moreover (6.2.19) yields

$$(6.2.20) \qquad c_0 \le (1 + \frac{2\nu_2}{\alpha})c_1$$

for all $\alpha \le 1$ and all k.
On the other hand we have for the α_k-discounted optimal policy $R_*^{(k)}$, defined by (6.2.12)

$$(6.2.21) \qquad \int_{J_k[0,\infty)} c_{\alpha_k, R_*^{(k)}}(x)\,dP_{R_*^{(k)}}(x) \ge M \sum_{\ell=0}^{\infty} \alpha_k^{\ell D_\alpha^{(k)} k\sigma^{-1}} \ge \frac{M\sigma}{\alpha D_\alpha^{(k)}} .$$

Since c_0 is at least as large as the left hand side of (6.2.21) we conclude

$$(6.2.22) \qquad D_\alpha^{(k)} \ge \frac{M\sigma}{(2\nu_2+1)c_1}$$

for all $\alpha \le 1$ and all $k \ge 1$. □

<u>THEOREM 6.2.11</u>. Consider the CTMDP with an initial distribution P_0 on S for which condition (6.2.3) holds. There exists for all $\alpha > 0$ a stationary deterministic policy (V_α, R_α) such that
(i) (V_α, R_α) is α-discounted optimal in the class of strong regular policies.
(ii) (V_α, R_α) is monotone.
Moreover, if c(.) is concave, then
(iii) (V_α, R_α) is of bang-bang type.

<u>PROOF</u>. Consider the sequence of functions $(r_\alpha^{(k)}(.))_{k=1}^{\infty}$ on S defined by (6.2.11). By part (ii) of theorem 6.2.9. and lemma 6.2.8. there exists a sequence $(k_j)_{j=1}^{\infty}$ of natural numbers such that

$$(6.2.23) \qquad r_\alpha(s) := \lim_{j\to\infty} r_\alpha^{(k_j)}(s)$$

exists for all $s \in S$.
Put

$$(6.2.24) \qquad D_\alpha := \inf \{s \in S: r_\alpha(s) = \mu\}.$$

From theorem 6.2.10. follows the existence of a $0 < D_0$ and $\alpha_0 > 0$ such that $D_\alpha \ge D_0 > 0$ for all $0 < \alpha < \alpha_0$. Put

$$V_\alpha := [D_\alpha, \infty)$$

(6.2.25)
$$R_\alpha(x,t) := \begin{cases} \nu_2 & \text{if } \pi_t x \in V_\alpha \\ r_\alpha(\pi_t x) & \text{otherwise} \end{cases}$$

Since $0 \notin V_\alpha$ and V_α is closed we know that (V_α, R_α) is a well-defined policy for the CTMDP and from proposition 6.2.2 follows that (V_α, R_α) is strong regular. Obviously (V_α, R_α) is stationary, deterministic and monotone. Moreover, if c(.) is concave then (V_α, R_α) is of bang-bang type. Finally, we find from theorem 4.2.9 that (V_α, R_α) is α-discounted optimal in the class of strong regular policies. □

6.3. THE AVERAGE COST CASE.

In order to make the transition from the α-discounted cost case to the average cost case we need the following theorem.

THEOREM 6.3.1. Let $r_\alpha^{(k)}(s)$ be defined by (6.2.11) and $D_\alpha^{(k)}$ by (6.2.17). If

(6.3.1)
$$\sup_{s \in S} \int_0^\infty e^{-\nu_2 t} \{b(s+\sigma t) - b(\sigma t)\}\, dt > M$$

then there exist an $\alpha_0 > 0$, a natural number k_0 and a number $\tilde{D}_0$ such that

$$D_\alpha^{(k)} \le \tilde{D}_0$$

for all $\alpha \le \alpha_0$ and $k \ge k_0$.

PROOF. From (6.2.7) and the fact that $\lim_{n\to\infty} f_{n,\alpha}^{(k)}(s) = f_\alpha^{(k)}(s)$ for all $s \in S$ it follows that

(6.3.2)
$$\alpha_k (f_\alpha^{(k)}(s) - f_\alpha^{(k)}(\sigma k^{-1})) \le M.$$

Choose, for fixed $\alpha > 0$ and fixed k a state $0 < s < D_\alpha^{(k)}$. From (6.2.11) follows that for all $n \in \{n_j : j \ge 1\}$ with n large enough $r_{n,\alpha}^{(k)}(s) \in [\nu_1, \nu_2]$ and

$$f_{n,\alpha}^{(k)}(s) - f_{n,\alpha}^{(k)}(\sigma k^{-1}) \geq k^{-1}(b(s)-b(\sigma k^{-1})) +$$

$$+ \alpha_k(1-k^{-1}r_{n,\alpha}^{(k)}(s)) \quad \{f_{n-1,\alpha}^{(k)}(s+\sigma k^{-1})-f_{n-1,\alpha}^{(k)}(2\sigma k^{-1})\} \geq$$

$$\geq k^{-1}(b(s)-b(\sigma k^{-1}))+\alpha_k(1-k^{-1}\nu_2)\{f_{n-1,\alpha}^{(k)}(s+\sigma k^{-1})-f_{n-1,\alpha}^{(k)}(2\sigma k^{-1})\}.$$

By letting $n \to \infty$ we find

$$f_{\alpha}^{(k)}(s) - f_{\alpha}^{(k)}(\sigma k^{-1}) \geq$$

$$\geq k^{-1}(b(s)-b(\sigma k^{-1}))+\alpha_k(1-k^{-1}\nu_2)\{f_{\alpha}^{(k)}(s+\sigma k^{-1})-f_{\alpha}^{(k)}(2\sigma k^{-1})\}.$$

Iteration yields

$$(6.3.3) \qquad f_{\alpha}^{(k)}(s) - f_{\alpha}^{(k)}(\sigma k^{-1}) \geq$$

$$\geq \sum_{n=0}^{[(D_{\alpha}^{(k)}-s)k\sigma^{-1}]} \alpha_k^n (1-k^{-1}\nu_2)^n k^{-1}\{b(s+n\sigma k^{-1})-b((n+1)\sigma k^{-1})\} \geq$$

$$\geq \int_0^{(D_{\alpha}^{(k)}-s)\sigma^{-1}} e^{-\alpha t} e^{-t(\nu_2+\frac{1}{2}\nu_2^2k^{-1})} \{b(s+\sigma t)-b(\sigma t)\}dt.$$

From (6.3.1) follows that we can find an $\varepsilon > 0$ and $\alpha_0 > 0$ such that for all $\alpha \leq \alpha_0$

$$\sup_{s\in S} \int_0^{\infty} e^{-(\nu_2+\alpha)t} \{b(s+\sigma t)-b(\sigma t)\}dt \geq M+4\varepsilon.$$

Next we choose $s_1 \in S$ such that for all $s \geq s_1$ and $\alpha \leq \alpha_0$

$$(6.3.4) \qquad \int_0^{\infty} e^{-(\nu_2+\alpha)t} \{b(s+\sigma t)-b(\sigma t)\}dt \geq M+3\varepsilon.$$

Since the left hand side of (6.3.4) is non-decreasing in s and non-increasing in α we can choose $s_2 \in S$ such that

$$\int_0^{s_2} e^{-(\nu_2+\alpha)t} \{b(s+\sigma t)-b(\sigma t)\}dt \geq M+2\varepsilon$$

for all $s \geq s_1$ and $\alpha \leq \alpha_0$.

Finally, choose k_0 such that for all $s \ge s_1$, $\alpha \le \alpha_0$ and $k \ge k_0$

$$(6.3.5) \qquad \alpha_k \int_0^{s_2} e^{-(\nu_2+\alpha)t} \, e^{-\frac{1}{2}\nu_2^2 t k^{-1}} \{b(s+\sigma t)-b(\sigma t)\}dt \ge M+\varepsilon.$$

Combination of (6.3.2), (6.3.3) and (6.3.5) yields

$$(D_\alpha^{(k)} - s)\sigma^{-1} \le s_2$$

for all $\alpha \le \alpha_0$, $k \ge k_0$ and $s \ge s_1$.
Hence we have for all $\alpha \le \alpha_0$ and $k \ge k_0$

$$D_\alpha^{(k)} \le \sigma s_2 + s_1 =: \tilde{D}_0. \qquad \square$$

<u>THEOREM 6.3.2</u>. Consider the CTMDP with an initial distribution P_0 on S, for which (6.2.3) holds. Assume that c(.) is concave and that condition (6.3.1) is satisfied.
Then there exists a stationary deterministic policy (V,R) such that
(i) (V,R) is average optimal in the class of strong regular policies.
(ii) (V,R) is of bang-bang type.

<u>PROOF</u>. Consider for $\alpha > 0$ the functions $r_\alpha(.)$ on S defined by (6.2.23). By part (iii) of theorem 6.2.11. we know that $r_\alpha(.)$ is a non-decreasing function with values in $\{\nu_1,\nu_2,\mu\}$ for all $\alpha > 0$. By lemma 6.2.8. there exists for any sequence $(\beta_k)_{k=1}^\infty$ with $\lim_{k\to\infty} \beta_k = 0$ a subsequence $(\beta_{k_j})_{j=1}^\infty$ such that

$$r(s) := \lim_{j\to\infty} r_{\beta_{k_j}}(s)$$

exists for all $s \in S$.
Put

$$D := \inf \{s: r(s) = \mu\}.$$

Then we know by theorems 6.2.10. and 6.3.1 that $0 < D < \infty$.

Put

$$V := [D,\infty)$$

(6.3.6) $$R(x,t) := \begin{cases} \nu_2 & \text{if } \pi_t x \in V \\ r(\pi_t x) & \text{otherwise} \end{cases}$$

Obviously (V,R) is a well-defined, deterministic and stationary policy, which is strong regular and of bang-bang type.
In order to apply theorem 4.4.6. we have to show that the conditions (4.4.6) and (4.4.7) of this theorem hold for the policies (V,R) and (V_α,R_α), defined by (6.3.6) and (6.2.25) resp. In lemma 6.3.4 below it is shown that (4.4.6) holds. Note that for all $\alpha > 0$ the policy (V_α,R_α) induces on $J[0,\infty)$ a regenerative stochastic process X_α, with sequence of regeneration epochs $(\tau_n(X_\alpha))_{n=0}^{\infty}$, where for $x \in J[0,\infty)$

(6.3.7) $\tau_n(x)$: = the epoch of the n-th entrance of x into state 0.

Define

(6.3.8) $$Y_{n,\alpha}(x) := c_{0,\tau_n(x),(V_\alpha,R_\alpha)}(x) - c_{0,\tau_{n-1}(x),(V_\alpha,R_\alpha)}(x).$$

Assume without loss of generality that the initial distribution P_0 is degenerated in state 0. Then we have for $P_{(V_\alpha,R_\alpha)}$-almost all $x = (t_j,i_j)_{j=1}^{\infty} \in J[0,\infty)$:

(6.3.9) $$\tau_1(x) \le D_\alpha\sigma^{-1}$$

(6.3.10) $$t_2 \ge D_\alpha\sigma^{-1} \rightarrow \tau_1(x) = D_\alpha\sigma^{-1}.$$

From (6.3.9) we conclude that

$$\int_{J[0,\infty)} \tau_1^2(x)\ dP_{(V_\alpha,R_\alpha)}(x) \le D_\alpha^2\sigma^{-2}$$

and

$$\int_{J[0,\infty)} Y_{1,\alpha}^2(x)\ dP_{(V_\alpha,R_\alpha)}(x) \le (\{b(D_\alpha)+\max_\nu c(\nu)\}D_\alpha\sigma^{-1} + M)^2.$$

From (6.3.10) we find

$$\int_{J[0,\infty)} \tau_1(x)\, dP_{(V_\alpha,R_\alpha)}(x) \geq D_\alpha \sigma^{-1} \exp\{-\nu_2 D_\alpha \sigma^{-1}\}.$$

From these inequalities together with theorem 6.2.10 and 6.3.1 follows that the conditions (4.4.13), (4.4.14) and (4.4.15) of theorem 4.4.11 are fulfilled, which implies that (4.4.7) holds.

REMARK 6.3.3. A stationary bang-bang type policy (V,R) can be characterized by two control parameters D_1 (the bang-bang parameter of R) and D_2 (the control limit) i.e. $R(x,t) = \nu_1$ iff $0 \leq \pi_t(x) < D_1$, $R(x,t) = \nu_2$ iff $\pi_t(x) \geq D_1$ and $V = [D_2,\infty)$. In the sequel we will assume that $D_1 \leq D_2$ and $D_2 > 0$.

In the next lemma we show that the average expected costs under stationary bang-bang type policies depend continuously on the control parameters of these policies.

LEMMA 6.3.4. Let (V,R) and (V_n,R_n), $n \geq 1$ be stationary bang-bang type policies for the CTMDP with control parameters (D_1,D_2) and $(D_1^{(n)},D_2^{(n)})$ respectively. If $\lim_{n\to\infty} D_1^{(n)} = D_1$ and $\lim_{n\to\infty} D_2^{(n)} = D_2$ then $\lim_{n\to\infty} c(V_n,R_n) = c(V,R)$.

PROOF. Let the initial distribution be degenerated in state 0. Since (V,R) and (V_n,R_n) induce regenerative stochastic processes on $J[0,\infty)$, it is by (4.4.19) sufficient to show that

$$\int_{J[0,\infty)} \tau_1(x)\, dP_{(V_n,R_n)}(x) \to \int_{J[0,\infty)} \tau_1(x)\, dP_{(V,R)}(x) \tag{6.3.11}$$

and

$$\int_{J[0,\infty)} Y_1^{(n)}(x)\, dP_{(V_n,R_n)}(x) \to \int_{J[0,\infty)} Y_1(x)\, dP_{(V,R)}(x), \tag{6.3.12}$$

where $\tau_1(x)$ is defined by (6.3.7), $Y_1(x) = c_{0,\tau_1(x),(V,R)}(x)$ and $Y_1^{(n)}(x) = c_{0,\tau_1(x),(V_n,R_n)}(x)$.

With a minor modification of the proof of theorem 2.4.3 it follows that

$$(6.3.13) \qquad P_{(V_n,R_n)} \overset{w}{\to} P_{(V,R)}.$$

Since $\tau_1(.)$ is bounded and continuous $P_{(V,R)}$-a.e. on $J[0,\infty)$ relation (6.3.11) follows from (6.3.13). To prove the validity of (6.3.12) we note that

$$(6.3.14) \qquad \left| \int Y_1^{(n)}(x)\ dP_{(V_n,R_n)}(x) - \int Y_1(x)\ dP_{(V,R)}(x) \right| \leq$$

$$\leq \int |Y_1^{(n)}(x)-Y_1(x)|\,dP_{(V_n,R_n)}(x) + \left|\int Y_1(x)\,dP_{(V_n,R_n)}(x) - \int Y_1(x)\,dP_{(V,R)}(x)\right|.$$

The second term on the right hand side of (6.3.14) converges to zero by (6.3.13) and the continuity and boundedness of $Y_1(.)$ on $J[0,\infty)$, $P_{(V,R)}$-a.e. Moreover, we have for $P_{(V_n,R_n)}$-almost all x:

$$|Y_1^{(n)}(x)-Y_1(x)| \leq |D_1^{(n)}-D_1|\sigma^{-1} \max_{\nu} c(\nu).$$

This inequality implies that also the first term on the right hand side of (6.3.14) converges to zero. □

The remainder of this section is devoted to the computation of the average expected costs under a stationary bang-bang type policy. Computation of this quantity is important for the development of algorithms to determine the optimal policy within the class of stationary bang-bang type policies. These algorithms, which make use of the special structure of the optimal policy, are more efficient than the general methods of successive approximation or standard policy iteration.

Consider the stationary bang-bang type policy (V,R) with control parameters D_1 and D_2, defined by

$$V := [D_2,\infty)$$

$$R(x,t) := \begin{cases} \nu_1 & \text{if } 0 \leq \pi_t(x) < D_1 \\ \nu_2 & \text{if } \pi_t(x) \geq D_1. \end{cases}$$

For any $s \in S$ denote with $P^{(s)}_{(V,R)}$ the probability measure on $J[0,\infty)$ induced by (V,R) and the initial distribution degenerated in s.
From the theory of regenerative stochastic processes we have for any initial distribution

$$(6.3.15) \qquad c(V,R) = \frac{\int_{J[0,\infty)} Y_1(x)\,dP^{(0)}_{(V,R)}(x)}{\int_{J[0,\infty)} \tau_1(x)\,dP^{(0)}_{(V,R)}(x)},$$

where $\tau_1(x)$ is defined by (6.3.7) and $Y_1(x) := c_{0,\tau_1(x),(V,R)}(x)$ for all $x \in J[0,\infty)$.
In view of the computation of the right hand side of (6.3.15) we introduce some auxiliary functions.
For $x \in J[0,\infty)$ and $i = 1,2$ define

$Z_i(x)$: = epoch of the first entrance of x into $[D_i,\infty)$

$E_i(x)$: = first entry state of x in $[D_i,\infty)$.

Next we denote for any Borelset B in $[D_1,\infty)$

$$(6.3.16) \qquad \psi(B) := P^{(0)}_{(V,R)}\{x:\ E_1(x) \in B\}$$

and for any $s \in S$ and $i \in \{1,2\}$

$$(6.3.17) \qquad t_i(s) := \int_{J[0,\infty)} Z_i(x)\ dP^{(s)}_{(V,R)}(x)$$

and

$$(6.3.18) \qquad k_i(s) := \int_{J[0,\infty)} \tilde{c}_{0,Z_i(x),(V,R)}(x)\ dP^{(s)}_{(V,R)}(x),$$

where $\tilde{c}_{0,T,(V,R)}(x)$ denotes the total *operating* costs incurred along the path x in the time interval $[0,T]$ when policy (V,R) is used (so replacement- and maintenance costs are not included in $\tilde{c}$).
We have the following obvious result.

THEOREM 6.3.5.

$$\int_{J[0,\infty)} \tau_1(x)\ dP^{(0)}_{(V,R)}(x) = t_1(0) + \int_{[D_1,D_2)} t_2(s)\ d\psi(s)$$

and

$$\int_{J[0,\infty)} Y_1(x)\ dP^{(0)}_{(V,R)}(x) = k_1(0) + M + t_1(0)c(\nu_1) + \\ + \int_{[D_1,D_2)} \{k_2(s) + t_2(s)c(\nu_2)\}\ d\psi(s). \quad \square$$

From theorem 6.3.5. and relation (6.3.15) follows that sufficient for the computation of c(V,R) is knowledge of the functions $t_i(.)$ and $k_i(.)$, i=1,2 and the probability distribution ψ. To determine these quantities we need the following lemma (cf. page 77 of COHEN (1976) and page 215 of TIJMS and VAN DER DUYN SCHOUTEN (1978)).

LEMMA 6.3.6. Let a(.) be a bounded measurable function defined on a finite interval (a,b) and let κ be a positive constant. Assume that F(.) is a probability distribution function with F(0) = 0 and with finite first moment. Let u(.) be a continuous function on (a,b), which satisfies for all except countable many $x \in (a,b)$ the integro-differential equation

$$\frac{du(x)}{dx} = a(x) + \kappa\{u(x) - \int_0^{b-x} u(x+y)\ dF(y)\} \qquad (6.3.19)$$

and the boundary condition

$$\lim_{x\uparrow b} u(x) = 0.$$

Then for all $x \in (a,b)$

$$u(x) = B(x) + \int_0^{b-x} e^{\delta y} B(x+y)\, dM(y),$$

where

$$B(x) = A(x) - \int_0^{b-x} A(x+y)\, dH(y)$$

with

$$A(x) = -\int_x^b a(y)\, dy$$

and

$$H(x) = \begin{cases} 0 & \text{for } x < 0 \\ \kappa \int_0^x (1-F(y))dy & \text{for } x \geq 0. \end{cases} \tag{6.3.20}$$

Furthermore δ is uniquely determined by

$$\int_0^\infty e^{-\delta y}\, dH(y) = 1 \tag{6.3.21}$$

and

$$M(x) := \sum_{n=1}^{\infty} G^{(2n)}(x) \tag{6.3.22}$$

with

$$G(x) := \int_0^x e^{-\delta y}\, dH(y), \tag{6.3.23}$$

while $G^{(n)}(.)$ denotes the n-fold convolution of G with itself, $n \geq 1$.

<u>PROOF</u>. First we note that by partial integration

$$\frac{d}{dx}\int_0^{b-x} u(x+y)\ (1-F(y))\ dy = -u(x) + \int_0^{b-x} u(x+y)\ dF(y).$$

Combining this with (6.3.19) yields

$$u(x) = -\int_x^b a(y)dy - \kappa \int_0^{b-x} u(x+y)(1-F(y))dy.$$

Hence, by the definition of H(.) and A(.)

$$(6.3.24) \qquad u(x) = A(x) - \int_0^{b-x} u(x+y)\ dH(y).$$

Iterating (6.3.24) yields

$$(6.3.25) \qquad u(x) = B(x) + \int_0^{b-x} u(x+y)\ dH^{(2)}(y).$$

It is straightforward to verify from (6.3.21) and (6.3.23) that

$$G^{(2)}(x) = \int_0^{x} e^{-\delta y}\ dH^{(2)}(y) \qquad \text{for } x \geq 0$$

and

$$\int_0^{\infty} e^{-\delta y}\ dH^{(2)}(y) = 1.$$

Equation (6.3.25) can now equivalently be written as (cf. page 362 of FELLER (1966))

$$(6.3.26) \qquad e^{\delta x}\ u(x) = e^{\delta x}\ B(x) + \int_0^{b-x} e^{\delta(x+y)}\ u(x+y)\ dG^{(2)}(y).$$

Put

$$u^{*}(x) = e^{\delta x} u(x).$$

Then (6.3.26) is equivalent with

$$(6.3.27) \qquad u^{*}(x) = e^{\delta x}\ B(x) + \int_0^{b-x} u^{*}(x+y)\ dG^{(2)}(y).$$

It is well-known (see for example FELLER (1966)) that

$$u^{*}(x) = e^{\delta x}\ B(x) + \int_0^{b-x} e^{\delta(x+y)}\ B(x+y)\ dM(y)$$

is the unique bounded solution of equation (6.3.27)

Hence

$$u(x) = B(x) + \int_0^{b-x} e^{\delta y} B(x+y)\, dM(y)$$

is the unique bounded solution of (6.3.19). □

<u>THEOREM 6.3.7.</u> The probability distribution function ψ on S is given by

$$\psi[v,\infty) = B(0) + \int_0^{D_1} e^{\delta y} B(y)\, dM(y) \qquad \text{for } v > D_1$$

$$\psi(\{D_1\}) = 1 - \lim_{v \downarrow D_1} \psi[v,\infty),$$

where δ is defined by (6.3.20) (with $\kappa = \nu_1\sigma^{-1}$) and (6.3.21) ; M(.) is defined by (6.3.22) and (6.3.23) while

$$B(s) := \nu_1\sigma^{-1} \int_s^{D_1} \bar{F}(v-y)\, dy - \nu_1\sigma^{-1} \int_{y=0}^{D_1 - s} \int_{z=s+y}^{D_1} \bar{F}(v-z)\, dz\, dH(y),$$

where $\bar{F}(x) := 1 - F(x)$, $x \geq 0$.

<u>PROOF</u>. Define for all $0 \leq s < D_1$ and all $v > D_1$

$$p(s,v) := P^{(s)}_{(V,R)}\{x\colon E_1(x) \in [v,\infty)\}.$$

Then $\psi[v,\infty) = p(0,v)$, for all $v > D_1$.
For fixed v the function $p(.,v)$ is continuous on $[0,D_1)$. Next, by using standard arguments, we have for all $0 \leq s < D_1$

$$p(s-\Delta s,v) = \frac{\nu_1 \Delta s}{\sigma} \{ \int_{v-s}^{\infty} dF(y) + \int_0^{D_1-s} p(s+y,v)\, dF(y)\} +$$
$$+ (1 - \frac{\nu_1\Delta s}{\sigma})\, p(s,v) + o(\Delta s),$$

from which we get for all $0 \leq s < D_1$

$$\frac{\partial p(s,v)}{\partial s} = -\frac{\nu_1}{\sigma}\bar{F}(v-s) + \frac{\nu_1}{\sigma}\{p(s,v) - \int_0^{D_1-s} p(s+y,v)\, dF(y)\}.$$

Moreover, $p(s,v)$ satisfies the boundary condition

$$\lim_{s \uparrow D_1} p(s,v) = 0 \qquad \text{for all } v > D_1.$$

Application of the previous lemma completes the proof. □

REMARK 6.3.8. An explicit expression for $\psi(.)$ can be obtained for the special case, where

$$F(t) = 1 - e^{-\eta t}, \qquad t \geq 0.$$

Elementary, but lengthy calculations yield

$$H(x) = \nu_1(\sigma\eta)^{-1}(1 - e^{-\eta x}) \qquad \text{for } x \geq 0$$

$$\delta = \nu_1\sigma^{-1} - \eta$$

$$G(x) = 1 - e^{-\nu_1 x\sigma^{-1}} \qquad \text{for } x \geq 0$$

$$\frac{dM(x)}{dx} = \tfrac{1}{2}\nu_1\sigma^{-1}(1 - e^{-2\nu_1 x\sigma^{-1}}) \qquad \text{for } x \geq 0$$

$$\psi[v,\infty) = \nu_1(\sigma\eta + \nu_1)^{-1}e^{-\eta v}(e^{\eta D_1} - e^{-\nu_1 D_1\sigma^{-1}}) \qquad \text{for } v > D_1$$

$$\psi(\{D_1\}) = (\sigma\eta + \nu_1)^{-1}(\sigma\eta + \nu_1 e^{-(\eta+\nu_1\sigma^{-1})D_1}).$$

THEOREM 6.3.9. For $0 \leq s < D_1$

$$k_1(s) = B(s) + \int_0^{D_1-s} e^{\delta y} B(s+y)\, dM(y) \tag{6.3.28}$$

$$t_1(s) = D(s) + \int_0^{D_1-s} e^{\delta y} D(s+y)\, dM(y), \tag{6.3.29}$$

where δ and $M(.)$ are defined by (6.3.20) upto (6.3.23), with $\kappa = \nu_1\sigma^{-1}$ and

$$B(s) := \sigma^{-1}\int_s^{D_1} b(y)\,dy - \sigma^{-1}\int_{y=0}^{D_1-s}\int_{x=s+y}^{D_1} b(x)\,dx\, dH(y)$$

$$D(s) := \sigma^{-1}(D_1-s) - \sigma^{-1}\int_{y=0}^{D_1-s}(D_1-s-y)\, dH(y).$$

PROOF. First we note that $k_1(.)$ and $t_1(.)$ are continuous on $[0,D_1)$. Then, for any $0 \le s < D_1$

$$k_1(s-\Delta s) = \frac{b(s)\Delta s}{\sigma} + \frac{\nu_1 \Delta s}{\sigma} \int_0^{D_1-s} k_1(s+y)\, dF(y) + $$
$$+ (1 - \frac{\nu_1 \Delta s}{\sigma})\, k_1(s) + o(\Delta s),$$

from which we find, for all $0 \le s < D_1$

$$k_1'(s) = -\frac{b(s)}{\sigma} + \frac{\nu_1}{\sigma}\{k_1(s) - \int_0^{D_1-s} k_1(s+y)\, dF(y)\}.$$

Moreover

$$\lim_{s\uparrow D_1} k_1(s) = 0.$$

Now (6.3.28) is obtained by application of lemma 6.3.6 , while (6.3.29) follows from (6.3.28) by putting $b(s) = 1$ for all $s \in [0,D_1)$. □

REMARK 6.3.10. Formulas for $k_2(s)$ and $t_2(s)$ for $s \in [D_1,D_2)$ are obtained from those for $k_1(s)$ and $t_1(s)$ by replacing ν_1 by ν_2 and D_1 by D_2.

EXAMPLE 6.3.11. For $F(x) = 1 - e^{-\eta x}$, $x \ge 0$ and $b(x) = b.x$, $x \ge 0$ for some positive constant b, we find after lengthy calculations for all $s \in [0,D_1)$

$$k_1(s) = \tfrac{1}{2} b\eta\, (\sigma\eta + \nu_1)^{-1} (D_1^2 - s^2) +$$
$$+ b\nu_1 (\sigma\eta + \nu_1)^{-2} (s - D_1 \exp\{-(D_1 - s)(\eta + \nu_1\sigma^{-1})\}) +$$
$$+ b\nu_1\sigma(\sigma\eta + \nu_1)^{-3}(1 - \exp\{-(D_1 - s)(\eta + \nu_1\sigma^{-1})\})$$

and

$$t_1(s) = \eta(\sigma\eta + \nu_1)^{-1}(D_1 - s) +$$
$$+ \nu_1(\sigma\eta + \nu_1)^{-2}(1 - \exp\{-(D_1 - s)(\eta + \nu_1\sigma^{-1})\}).$$

CHAPTER 7

AN INVENTORY MODEL

7.1. INTRODUCTION AND ASSUMPTIONS.

Our third application concerns an inventory model which has got ample attention in the literature. However, as far as the author knows, the final results obtained in this section, are new. The model that we consider here is as follows.

Customers arrive at a warehouse according to a Poisson process with constant arrival rate ν. The demands of the customers are independent random variables having common probability distribution function F, with $F(0) = 0$. Unfilled demands are backlogged. At every moment the decision maker can place an order of any size. We assume that an order is delivered without lead time.

The cost structure consists of two parts: a holding- or penalty cost rate b(s) is incurred if the inventory level is s (negative values of the state variable indicate the amount to be backlogged) and a lump cost m(a) is incurred when an order of size a is placed.

For this model we define the following CTMDP:

$$
\begin{aligned}
&S: = \mathbb{R} \\
&A_1: = \{\nu\} \\
&A_2: = [0,\infty) \\
&q(s,\nu): = \nu, \qquad s \in S \\
(7.1.1) \qquad &\Pi(s,\nu,[s,s-t]): = F(t), \qquad s \in S \text{ and } t \geq 0 \\
&p(s,a,\{s+a\}): = 1, \qquad (s,a) \in S \times A_2 \\
&c_1(s,\nu): = b(s), \qquad s \in S \\
&c_2(s,a): = m(a), \qquad (s,a) \in S \times A_2 \\
&f(s,t): = s, \qquad (s,t) \in S \times [0,\infty).
\end{aligned}
$$

The state of the system denotes the inventory level, where negative values indicate the amount to be backlogged. The only possible control is the choice of the fixed arrival rate ν. An impulsive control represents the size of an order. Further we assume the drift-function to be constant in its time variable.
We make the following assumptions on the model parameters.

ASSUMPTION 7.1.1.

(i) $F(.)$ has a density function;

(ii) $\mu: = \int_0^\infty t dF(t) < \infty$ and $\mu^{(2)}: = \int_0^\infty t^2 dF(t) < \infty$

(iii) $b(s) = \begin{cases} bs, & s \geq 0 \\ -ps, & s < 0 \end{cases}$

for finite, non-negative constants b and p

(iv) $m(a) = M\delta(a) + ma, \quad a \geq 0$

for finite, non-negative constants m and M while $\delta(a) = 1$ if $a > 0$ and $\delta(a) = 0$ if $a = 0$.

With this assumption model (7.1.1) is a well defined CTMDP for which assumption 2.2.2. holds.
The k th approximating DTMDP is defined by $(S, A_1, A_2, p_1^{(k)}, p_2^{(k)}, c_1^{(k)}, c_2^{(k)}, f, k)$ with (cf. remark 2.4.2.(ii))

$$p_1^{(k)}(s,\nu,[s,s-t]) = 1-k^{-1}\nu+k^{-1}\nu F(t); \qquad s \in S,\ t \geq 0$$
$$p_2^{(k)}(s,a,[s+a,s+a-t]) = 1-k^{-1}\nu+k^{-1}\nu F(t); \qquad (s,a) \in S \times A_2,\ t \geq 0.$$
$$c_1^{(k)}(s,\nu) = k^{-1}b(s); \qquad s \in S$$
$$c_2^{(k)}(s,a) = m(a)+k^{-1}b(s+a); \qquad (s,a) \in S \times A_2.$$

REMARK 7.1.2. (i) Since $A_1=\{\nu\}$ it follows for every policy (V,R) with $R=(R_1,R_2)$ that $R_1(x,t)$ is a probability measure degenerated in $\{\nu\}$ for all (x,t). Hence in this chapter a policy is denoted by (V,R), where V is a closed subset of S and R a transition probability from $J[0,\infty)\times[0,\infty)$ to A_2.
(ii) Let R be a transition probability from $J[0,\infty)\times[0,\infty)$ to A_2 such that $R(.,t)$ is r_t-measurable and let V be a closed subset of S. Then (V,R) is a policy iff $R(x,t)\{a: a+\pi_t x \in V\} = 0$ for all (x,t) with $\pi_t x \in V$.

7.2. THE α-DISCOUNTED COST CASE.

PROPOSITION 7.2.1. Let P_0 be an initial distribution on S. A policy (V,R) is regular if

(7.2.1) δV is countable

(7.2.2) $$P_{(V,R)}\{x: R(x_k,t_k) \overset{w}{\not\to} R(x,t) \text{ for some } t, \text{ while } d(x_k,x) \to 0,\ t_k \to t\} = 0.$$

PROOF. An immediate consequence of assumption 7.1.1.(i) and theorem 2.4.3. □

PROPOSITION 7.2.2. Let P_0 be an initial distribution on S, with

(7.2.3) $$\int_S s^2 \, dP_0(s) < \infty.$$

A policy (V,R) is strong regular if (7.2.1) and (7.2.2) hold and if there exists a state $s_0 \in S$ such that for all $(x,t) \in J[0,\infty)\times[0,\infty)$ with $\pi_t x \in V$

(7.2.4) $$R(x,t)([0,s_0-\pi_t x]) = 1$$

(this implies that $V \cap [s_0,\infty) = \emptyset$).

PROOF. Let (V,R) be a policy for which (7.2.1) and (7.2.2) hold. Define for the k th approximating DTMDP policies $(V^{(k)},R^{(k)})$ by $V^{(k)}=V$ and $R^{(k)}=R$, $k \geq 1$. We will show that (V,R) and $(V^{(k)},R^{(k)})$ satisfy the conditions of theorem 4.2.1.

First of all we note that condition (2.4.3) follows from (7.2.1) and assumption 7.1.1.(i), that condition (2.4.4) follows from (7.2.2) while condition (2.4.5) is trivially satisfied.

Next we note that condition (4.2.1) holds for all $\delta > 0$ and all $k \geq 1$. Put

$$A_1(j) := A_1, \qquad j \in S$$

$$A_2(j) := \begin{cases} [0,s_0-j], & j < s_0 \\ \emptyset, & j \geq s_0 \end{cases}$$

Then $R^{(k)}(x,t)(A_2^c(j)) = 0$ for all $(x,t) \in J_k[0,\infty) \times L_k$ for which $\pi_t x = j \in V^{(k)}$. Choose $\alpha > 0$ and $\beta > 1$ (depending on α) such that (4.2.11) holds for some $\delta > 0$.

Define

$$h_1(s) = \begin{cases} (b+p)s_0 & \text{for } |s| < s_0 \\ (b+p)|s| & \text{for } |s| \geq s_0 \end{cases}$$

$$h_2(s) = \begin{cases} (M+m)s_0 & \text{for } s \geq 0 \\ (M+m)(s_0-s) & \text{for } s < 0, \end{cases}$$

where we assume, without loss of generality, that $s_0 > 1$ and $(1 + 2\mu s_0^{-1} + \mu^{(2)} s_0^{-2}) < \beta$.

Now one easily verifies that condition (4.2.2) is satisfied. Conditions (4.2.3), (4.2.4) and (4.2.5) are trivially fulfilled for $\ell(.) \equiv 1$. To verify (4.2.6) and (4.2.7) we observe that

$$\sup_{a\epsilon A_1(s)} \int_S h_1^2(s_1)\ d\Pi(s,a,s_1) = \int_0^\infty h_1^2(s-t)\ dF(t) \leq$$

$$\leq \begin{cases} (b+p)^2\ s^2\ (1+\mu^{(2)} s^{-2}) & \text{for } s > s_0 \\ (b+p)^2\ s_0^2\ (1+2\mu s_0^{-1}+\mu^{(2)} s_0^{-2}) & \text{for } -s_0 \leq s \leq s_0; \\ (b+p)^2\ s^2\ (1+2\mu s^{-1}+\mu^{(2)} s^{-2}) & \text{for } s < -s_0 \end{cases}$$

$$\sup_{a\epsilon A_1(s)} \int_S h_2^2(s_1)\ d\Pi(s,a,s_1) = \int_0^\infty h_2^2(s-t)\ dF(t) \leq$$

$$\leq \begin{cases} (M+m)^2 s_0^2 (1+2\mu s_0^{-1}+\mu^{(2)} s_0^{-2}) & \text{for } s \geq 0 \\ (M+m)^2 (s_0-s)^2 (1+2\mu(s_0-s)^{-1}+\mu^{(2)}(s_0-s)^{-2}) & \text{for } s < 0 \end{cases};$$

$$\sup_{a\epsilon A_2(s)} \int_S h_1^2(s_1)\ dp(s,a,s_1) = \sup_{a\epsilon A_2(s)} h_1^2(s+a) \leq h_1^2(s)$$

and

$$\sup_{a\epsilon A_2(s)} \int_S h_2^2(s_1)\ dp(s,a,s_1) = \sup_{a\epsilon A_2(s)} h_2^2(s+a) \leq h_2^2(s).$$

Finally condition (4.2.8) follows from (7.2.3) while (4.2.9) and (4.2.10) follow from the bounded convergence theorem and assumption 7.1.1.(ii). □

DEFINITION 7.2.3. Let (V,R) be a memoryless deterministic policy for the CTMDP. Hence there exists a function $r(.,.)$ on $S \times [0,\infty)$ with values in A_2 such that $r(s,t) = R(x,t)$ for all $(x,t) \in J[0,\infty) \times [0,\infty)$ with $\pi_t x = s$. The policy (V,R) is of *(s,S) type* if there exist states $s^* \le S^*$, such that for all $s \in S$ and $t \ge 0$

$$V = (-\infty, s^*]$$

$$r(s,t) = [S^*-s]^+. \quad ^{*)}$$

REMARK 7.2.4. As in the previous chapter we treat the k th approximating DTMDP for finite as well as infinite horizon as a classical discrete time Markov decision process $(S,A,p^{(k)},c^{(k)})$ with

$$A: = [0,\infty)$$
$$p^{(k)}(s,a,[s+a,s+a-t]): = 1-k^{-1}\nu+k^{-1}\nu F(t); \ (s,a) \in S \times A, \ t \ge 0$$
$$c^{(k)}(s,a): = M\delta(a)+ma+k^{-1}b(s+a).$$

A policy for this discrete time Markov decision process is denoted by R, a transition probability from $J_k[0,\infty) \times L_k$ to A, such that $R(.,t)$ is r_t-measurable.

DEFINITION 7.2.5. Let R be a deterministic memoryless policy for the k th approximating DTMDP. Hence there exists a function $r(.,.)$ on $S \times [0,\infty)$ with values in A such that $r(s,t) = R(x,t)$ for all $(x,t) \in J_k[0,\infty) \times L_k$ with $\pi_t x = s$.
The policy R is of *(s,S) type* if there exist for all $t \ge 0$ states $s_t^* \le S_t^*$ such that

$$r(s,t) = \begin{cases} S_t^*-s & \text{if } s \le s_t^* \\ 0 & \text{otherwise.} \end{cases}$$

*) $[x]^+: = \max(x,0)$.

Conditions under which the optimal policy for the discrete time, finite- and infinite horizon model is of (s,S) type have been derived by several authors. Here we mention SCARF (1960), ZABEL (1962), IGLEHART (1963), VEINOTT (1966), PORTEUS (1971) and SCHÄL (1976).

THEOREM 7.2.6. Consider the k th approximating DTMDP with some initial distribution P_0 on S. Put $\alpha > 0$ and $\alpha_k := \exp(-\alpha k^{-1})$. For all $n \geq 1$ there exists a memoryless deterministic policy $R_n^{(k)}$ such that

(i) $R_n^{(k)}$ is α_k-discounted, n-horizon optimal in the class of all policies

(ii) $R_n^{(k)}$ is of (s,S) type.

PROOF. The proof follows the lines of SCARF (1960) and ZABEL (1962). For $n \geq 1$ and $s \in S$ we denote

$$f_{n,\alpha}^{(k)}(s) := \inf_{R \in \mathcal{R}} \int_{J_k[0,\infty)} c_{\alpha_k,n,R}(x) \; dP_R(x \mid \pi_0 x = s),$$

where $\mathcal{R}$ denotes the class of all policies.
By induction on n follows for all $n \geq 1$

$$f_{n,\alpha}^{(k)}(s) = \inf_{a \geq 0} \{M\delta(a) + ma + k^{-1}b(s+a) + \alpha_k (1-\frac{\nu}{k}) f_{n-1,\alpha}^{(k)}(s+a) +$$

$$+ \alpha_k \frac{\nu}{k} \int_0^\infty f_{n-1,\alpha}^{(k)}(s+a-t) \; dF(t)\},$$

where we define

$$f_{0,\alpha}^{(k)}(s): = \begin{cases} 0 & \text{for } s \geq 0 \\ -\tilde{m}s & \text{for } s < 0 \end{cases}$$

for some finite constant $\tilde{m} > m$.
Put for $n \geq 1$

$$(7.2.5) \qquad w_{n,\alpha}^{(k)}(s): = ms + k^{-1}b(s) + \alpha_k(1 - \frac{\nu}{k})\, f_{n-1,\alpha}^{(k)}(s) + + \alpha_k \frac{\nu}{k} \int_0^\infty f_{n-1,\alpha}^{(k)}(s-t)\, dF(t).$$

Hence

$$(7.2.6) \qquad f_{n,\alpha}^{(k)}(s) = \inf_{a\geq 0} \{M\delta(a) + w_{n,\alpha}^{(k)}(s+a) - ms\}.$$

Define

$$(7.2.7) \qquad r_{n,\alpha}^{(k)}(s): = \inf \{\tilde{a}\geq 0:\ M\delta(\tilde{a})+w_{n,\alpha}^{(k)}(s+\tilde{a}) = \inf_{a\geq 0} \{M\delta(a)+w_{n,\alpha}^{(k)}(s+a)\}\}.$$

From part (i) of lemma 7.2.9. below it follows that $r_{n,\alpha}^{(k)}(s) < \infty$ for all $s \in S$. Next we define for all $(x,\ell k^{-1}) \in J_k[0,\infty) \times L_k$

$$R_n^{(k)}(x,\ell k^{-1}): = r_{n-\ell+1,\alpha}^{(k)}(\pi_{\ell k^{-1}}x),\quad 1 \leq \ell \leq n$$

$$R_n^{(k)}(x,\ell k^{-1}): = R_n^{(k)}(x,nk^{-1}),\quad \ell \geq n.$$

From part (iii) of lemma 7.2.9. below it follows that $R_n^{(k)}(.,.)$ is a well defined policy for the k th approximating DTMDP. Obviously, $R_n^{(k)}(.,.)$ is memoryless, deterministic and α_k-discounted optimal in the class of all policies. Finally, we conclude from lemma 7.2.9. below that $R_n^{(k)}(.,.)$ is of (s,S) type. □

We need the concept of K-convexity, introduced by SCARF (1960) (see SCHÄL (1976) for a generalization of this concept).

DEFINITION 7.2.7. Let $f(.)$ be a real-valued function on $\mathbb{R}$ and let K be a real constant. The function $f(.)$ is called *K-convex* if

$$\frac{f(y) - f(x)}{y - x} \le \frac{f(z) - f(x) + K}{z - x}$$

for all $x < y < z$.

First we state without proof some well-known properties of K-convex functions.

LEMMA 7.2.8.

(i) If f is K-convex then f is M-convex for all $M \ge K$.

(ii) If f and g are K-convex and M-convex, respectively, and $\gamma_1, \gamma_2 > 0$, then $\gamma_1 f + \gamma_2 g$ is $(\gamma_1 K + \gamma_2 M)$-convex.

(iii) If f is K-convex and $F(.)$ is a probability distribution function on $[0,\infty)$ with ${}_0\int^\infty |f(x-z)|\, dF(z) < \infty$ for all $x \in \mathbb{R}$, then ${}_0\int^\infty f(x-z)\, dF(z)$ is also K-convex.

(iv) Assume that f is K-convex and continuous with $\lim_{|x|\to\infty} f(x) = \infty$. Let

$$s_1 := \inf \{\tilde{s} : f(\tilde{s}) = \inf_s f(s)\}$$

and

$$s_2 := \inf \{s \le s_1 : K + f(s_1) = f(s)\}.$$

Then

$$\begin{aligned} f(s) &> f(s_2) && \text{for all } s < s_2 \\ f(y) &\ge f(x) - K && \text{for all } y \ge x \ge s_1 \\ f(s) &\le f(s_2) && \text{for all } s_2 \le s \le s_1. \end{aligned}$$

LEMMA 7.2.9. Let $w^{(k)}_{n,\alpha}(.)$ and $r^{(k)}_{n,\alpha}(.)$ be as given by (7.2.5) and (7.2.7) in the proof of the previous theorem. Then for all k large enough and all α small enough

(i) $w^{(k)}_{n,\alpha}(.)$ is M-convex and continuous

(ii) $\lim_{s\to\infty} w^{(k)}_{n,\alpha}(s) = \lim_{s\to-\infty} w^{(k)}_{n,\alpha}(s) = \infty$

(iii) there exist states $s^{(k)}_{\ell,\alpha} \le S^{(k)}_{\ell,\alpha}$, $1 \le \ell \le n$, such that

$$r^{(k)}_{\ell,\alpha}(s) = \begin{cases} S^{(k)}_{\ell,\alpha} - s & \text{if } s < s^{(k)}_{\ell,\alpha} \\ 0 & \text{otherwise.} \end{cases}$$

PROOF. The proof proceeds by induction on n. For n = 1 we have

$$w_{1,\alpha}^{(k)}(s) = \begin{cases} (m+bk^{-1}-\tilde{m}\alpha_k\nu k^{-1})s + \alpha_k\nu k^{-1}\tilde{m}\int_s^\infty t\,dF(t), & s \geq 0 \\ (m-pk^{-1}-\tilde{m}\alpha_k)s + \alpha_k\nu k^{-1}\tilde{m}\mu & , \quad s < 0. \end{cases}$$

Choose k_0 such that for all $k \geq k_0$

$$m + bk^{-1} - \tilde{m}\alpha_k\nu k^{-1} > 0$$

$$m - pk^{-1} - \tilde{m}\alpha_k < 0$$

$$\alpha_k\nu k^{-1}\tilde{m}\mu < M.$$

One easily checks that for all $k \geq k_0$ the function $w_{1,\alpha}^{(k)}(.)$ is M-convex and continuous and that $\lim_{s\to\infty} w_{1,\alpha}^{(k)}(s) = \lim_{s\to-\infty} w_{1,\alpha}^{(k)}(s) = \infty$. This implies that $w_{1,\alpha}^{(k)}(.)$ attains its infimum. Put

$$S_{1,\alpha}^{(k)} := \inf\{\tilde{s}: w_{1,\alpha}^{(k)}(\tilde{s}) = \inf_s w_{1,\alpha}^{(k)}(s)\}$$

and

$$s_{1,\alpha}^{(k)} := \inf\{s \leq S_{1,\alpha}^{(k)}: M + w_{1,\alpha}^{(k)}(S_{1,\alpha}^{(k)}) = w_{1,\alpha}^{(k)}(s)\}.$$

Since $w_{1,\alpha}^{(k)}(.)$ is M-convex and continuous we find from part (iv) of lemma 7.2.8.

$$w_{1,\alpha}^{(k)}(s) > w_{1,\alpha}^{(k)}(s_{1,\alpha}^{(k)}) \qquad \text{for all } s < s_{1,\alpha}^{(k)}$$

$$w_{1,\alpha}^{(k)}(y) \geq w_{1,\alpha}^{(k)}(x) - M \qquad \text{for all } y \geq x \geq S_{1,\alpha}^{(k)}$$

$$w_{1,\alpha}^{(k)}(s) \leq w_{1,\alpha}^{(k)}(s_{1,\alpha}^{(k)}) \qquad \text{for all } s_{1,\alpha}^{(k)} \leq s \leq S_{1,\alpha}^{(k)}.$$

This implies

$$r_{1,\alpha}^{(k)}(s) = \begin{cases} S_{1,\alpha}^{(k)} - s & \text{for } s < s_{1,\alpha}^{(k)} \\ 0 & \text{for } s \geq s_{1,\alpha}^{(k)}. \end{cases}$$

This completes the proof for n = 1.

Suppose the lemma is true for n. Hence there exist $s_{n,\alpha}^{(k)} \le S_{n,\alpha}^{(k)}$ such that

$$S_{n,\alpha}^{(k)} := \inf \{\tilde{s}: w_{n,\alpha}^{(k)}(\tilde{s}) = \inf_{s} w_{n,\alpha}^{(k)}(s)\} \tag{7.2.8}$$

$$s_{n,\alpha}^{(k)} := \inf \{s \le S_{n,\alpha}^{(k)}: M + w_{n,\alpha}^{(k)}(S_{n,\alpha}^{(k)}) = w_{n,\alpha}^{(k)}(s)\} \tag{7.2.9}$$

$$f_{n,\alpha}^{(k)}(s) = \begin{cases} M + w_{n,\alpha}^{(k)}(S_{n,\alpha}^{(k)}) - ms & \text{for } s < s_{n,\alpha}^{(k)} \\ w_{n,\alpha}^{(k)}(s) - ms & \text{for } s \ge s_{n,\alpha}^{(k)} \end{cases} \tag{7.2.10}$$

Reasoning along the same lines as e.g. on page 174 in ROSS (1970), it follows that $f_{n,\alpha}^{(k)}(.)$ is M-convex and continuous. This implies with parts (i), (ii) and (iii) of lemma 7.2.8. that $w_{n+1,\alpha}^{(k)}(.)$ is M-convex and continuous, which proves (i) for n+1.
Moreover, we have for $s \le s_{n,\alpha}^{(k)}$

$$w_{n+1,\alpha}^{(k)}(s) = ms(1-\alpha_k) + k^{-1}b(s) + \alpha_k \{M+w_{n,\alpha}^{(k)}(S_{n,\alpha}^{(k)}) + \nu m\mu k^{-1}\}.$$

Choose $\alpha_0 < pm^{-1}$. Then $m(1-\alpha_k)-pk^{-1} < 0$ for all $0 < \alpha \le \alpha_0$, which implies that

$$\lim_{s\to-\infty} w_{n+1,\alpha}^{(k)}(s) = \infty.$$

Next we note that for $y \ge x$

$$\begin{aligned} mx + f_{n,\alpha}^{(k)}(x) &= \inf_{a\ge0} \{M\delta(a) + w_{n,\alpha}^{(k)}(x+a)\} \le \\ &\le \inf_{a\ge0} \{M\delta(a) + w_{n,\alpha}^{(k)}(y+a)\} + M = \\ &= my + f_{n,\alpha}^{(k)}(y) + M. \end{aligned} \tag{7.2.11}$$

Hence, by (7.2.5) we have for $y \ge x$

$$w_{n+1,\alpha}^{(k)}(y) - w_{n+1,\alpha}^{(k)}(x) \ge m(1-\alpha_k)(y-x) + k^{-1}b(y) - k^{-1}b(x) - \alpha_k M \tag{7.2.12}$$

which in turn gives

$$\lim_{s\to\infty} w^{(k)}_{n+1,\alpha}(s) = \infty.$$

This proves (ii) for n+1. Finally the proof of (iii) for n+1 proceeds similar as for n=1. □

LEMMA 7.2.10. Let $S^{(k)}_{n,\alpha}$ and $s^{(k)}_{n,\alpha}$ be defined by (7.2.8) and (7.2.9) in the proof of the previous lemma. Then

$$0 \le S^{(k)}_{n,\alpha} \le kMb^{-1}$$

and

$$-Mk(p-m\alpha)^{-1} \le s^{(k)}_{n,\alpha} \le 0.$$

PROOF. From relation (7.2.12) follows for all $y \ge x \ge 0$

$$w^{(k)}_{n,\alpha}(y) - w^{(k)}_{n,\alpha}(x) \ge -\alpha_k M.$$

This implies with the definition of $s^{(k)}_{n,\alpha}$ that

$$s^{(k)}_{n,\alpha} \le 0.$$

On the other hand we conclude from (7.2.12) that for $y \ge x \ge 0$

$$w^{(k)}_{n,\alpha}(y) - w^{(k)}_{n,\alpha}(x) \ge k^{-1}b(y-x) - M.$$

Since $w^{(k)}_{n,\alpha}$ attains its infimum in $S^{(k)}_{n,\alpha}$ we conclude that

$$S^{(k)}_{n,\alpha} \le kMb^{-1}.$$

Next we show by induction on n that $w^{(k)}_{n,\alpha}(.)$ is non-increasing on $(-\infty,0]$. This is obvious for n=1. Suppose it is true for n. Then (7.2.10) implies that for $x \le y \le 0$

$$f^{(k)}_{n,\alpha}(x) - f^{(k)}_{n,\alpha}(y) \ge -m(x-y). \tag{7.2.13}$$

Hence it follows by (7.2.5) that for all $x \le y \le 0$

$$(7.2.14) \qquad w^{(k)}_{n+1,\alpha}(x) - w^{(k)}_{n+1,\alpha}(y) \geq m(1-\alpha_k)(x-y) - k^{-1}p(x-y) \geq 0.$$

From the fact that $w^{(k)}_{n,\alpha}(.)$ is non-increasing on $(-\infty,0]$ we conclude that

$$S^{(k)}_{n,\alpha} \geq 0.$$

Finally (7.2.14) yields for $x \leq 0$

$$w^{(k)}_{n,\alpha}(x) \geq m(1-\alpha_k)x - k^{-1}px + w^{(k)}_{n,\alpha}(0) \geq$$

$$\geq m\alpha k^{-1}x - k^{-1}px + w^{(k)}_{n,\alpha}(S^{(k)}_{n,\alpha}).$$

Hence, by the definition of $s^{(k)}_{n,\alpha}$

$$s^{(k)}_{n,\alpha} \geq -Mk(p-m\alpha)^{-1}. \qquad \square$$

<u>THEOREM 7.2.11</u>. Consider the k th approximating DTMDP with some initial distribution P_0 on S. Choose $\alpha > 0$ and put $\alpha_k := \exp(-\alpha k^{-1})$. There exists a stationary deterministic policy $R^{(k)}_*$ such that

(i) $R^{(k)}_*$ is α_k-discounted optimal in the class of all policies.

(ii) $R^{(k)}_*$ is of (s,S) type.

<u>PROOF</u>. Let $S^{(k)}_{n,\alpha}$ and $s^{(k)}_{n,\alpha}$ be defined by (7.2.8) and (7.2.9). From the previous lemma follows the existence of real numbers $S^{(k)}_{\alpha}$ and $s^{(k)}_{\alpha}$ and a sequence $(n_j)_{j=1}^{\infty}$ of natural numbers such that

$$(7.2.15) \qquad \lim_{j\to\infty} S^{(k)}_{n_j,\alpha} =: S^{(k)}_{\alpha}$$

and

$$(7.2.16) \qquad \lim_{j\to\infty} s^{(k)}_{n_j,\alpha} =: s^{(k)}_{\alpha}.$$

Put

$$R^{(k)}_*(x,t) = \begin{cases} S^{(k)}_{\alpha} - \pi_t x & \text{if } \pi_t x < s^{(k)}_{\alpha} \\ 0 & \text{otherwise.} \end{cases}$$

Then $R_*^{(k)}(.,.)$ is a well-defined stationary deterministic policy, which is obviously of (s,S) type.
To prove that $R_*^{(k)}(.,.)$ is α_k-discounted optimal we define

$$(7.2.17) \qquad f_\alpha^{(k)}(s):=\inf_{R\in\mathcal{R}} \int_{J_k[0,\infty)} c_{\alpha_k,R}(x)\, dP_R(x|\pi_0 x = s),\; s \in S.$$

Since $(f_{n,\alpha}^{(k)}(.))$ is a non-decreasing sequence of non-negative functions on S

$$v(s):=\lim_{n\to\infty} f_{n,\alpha}^{(k)}(s) \geq 0$$

exists for all $s \in S$.
From Dini's theorem (see page 162 of ROYDEN (1968)) it follows that

$$(7.2.18) \qquad f_{n,\alpha}^{(k)}(.) \overset{c}{\to} v(.).$$

By considering the (s,S) type policy with $s = S = 0$ we easily find that

$$(7.2.19) \qquad f_{n,\alpha}^{(k)}(s) \leq f_\alpha^{(k)}(s) \leq M - ms + \nu\alpha^{-1}(M+m\mu) \qquad \text{for } s < 0.$$

Combining (7.2.18) and (7.2.19) with (7.2.5) we find with the bounded convergence theorem that

$$(7.2.20) \qquad w_{n,\alpha}^{(k)}(.) \overset{c}{\to} w(.),$$

where

$$w(s):= ms + k^{-1}b(s) + \alpha_k(1-\frac{\nu}{k})\, v(s) + \alpha_k \frac{\nu}{k}\int_0^\infty v(s-t)\, dF(t).$$

From (7.2.10), (7.2.18) and (7.2.20) follows

$$(7.2.21) \qquad v(s) = \begin{cases} M + w(S_\alpha^{(k)}) - ms & \text{for } s < s_\alpha^{(k)} \\ w(s) - ms & \text{for } s \geq s_\alpha^{(k)}. \end{cases}$$

Iteration of (7.2.21) yields for $s \in S$ and $n \geq 1$

$$v(s) \geq \int_{J_k[0,\infty)} c_{\alpha_k,n,R_*^{(k)}}(x) \; dP_{R_*^{(k)}}(x|\pi_0 x = s).$$

By letting $n \to \infty$ we find

$$(7.2.22) \qquad v(s) \geq \int_{J_k[0,\infty)} c_{\alpha_k,R_*^{(k)}}(x) \; dP_{R_*^{(k)}}(x|\pi_0 x = s) \geq f_\alpha^{(k)}(s).$$

Together (7.2.22) and $v(s) \leq f_\alpha^{(k)}(s)$ yield the α_k-discounted optimality of $R_*^{(k)}(.,.)$ in $\mathcal{R}$. □

LEMMA 7.2.12. Let $S_\alpha^{(k)}$ and $s_\alpha^{(k)}$ be defined by (7.2.15) and (7.2.16). Then for α small enough

$$0 \leq S_\alpha^{(k)} \leq M(\alpha+\nu)b^{-1}$$

and

$$-M(\alpha+\nu)(p-\alpha m)^{-1} \leq s_\alpha^{(k)} \leq 0.$$

PROOF. Let $f_\alpha^{(k)}(.)$ be defined by (7.2.17). From the previous theorem it follows that for all $s \in S$

$$f_\alpha^{(k)}(s) = \inf_{R\in\mathcal{R}} \int_{J_k[0,\infty)} c_{\alpha_k,R}(x) \; dP_R(x|\pi_0 x = s),$$

where $\mathcal{R}$ is the collection of all semi-Markov strategies, i.e. $R \in \mathcal{R}$ if for fixed x the function $R(x,.)$ is constant on $(T_n(x),T_{n+1}(x)]$ for all $n \geq 1$. Stated otherwise, the function $f_\alpha^{(k)}(.)$ is the optimal value function of the semi-Markov decision process, where the sojourn time $T(s,a)$ in state $s \in S$, when action a is chosen is a random variable with the following distribution function

$$\mathbb{P}\{T(s,a) = nk^{-1}\} = (1-\nu k^{-1})^{n-1}\nu k^{-1}, \qquad a = 0,\ s \in S$$
$$\mathbb{P}\{T(s,a) = k^{-1}\} = 1 \qquad\qquad , \qquad a > 0,\ s \in S.$$

Reasoning along the same lines as e.g. in chapter 7 of ROSS (1970) or in LIPPMAN (1973) it follows that $f_\alpha^{(k)}(.)$ satisfies the functional equation

$$f_\alpha^{(k)}(s) = \inf_{a\geq 0} \{M\delta(a) + ma + k^{-1} b(s+a) E(\sum_{n=0}^{kT(s+a,0)-1} \alpha_k^n) +$$

$$+ E(\alpha_k^{kT(s+a,0)}) \int_0^\infty f_\alpha^{(k)}(s+a-t)\, dF(t)\} =$$

$$= \inf_{a\geq 0} \{M\delta(a) + ma + \frac{b(s+a)}{k(1-\alpha_k)+\alpha_k \nu} +$$

$$+ \frac{\alpha_k \nu}{k(1-\alpha_k)+\alpha_k \nu} \int_0^\infty f_\alpha^{(k)}(s+a-t)\, dF(t)\}.$$

Put

$$\tilde{w}_\alpha^{(k)}(s) := ms + \frac{b(s)}{k(1-\alpha_k)+\alpha_k \nu} + \frac{\alpha_k \nu}{k(1-\alpha_k)+\alpha_k \nu} \int_0^\infty f_\alpha^{(k)}(s-t)\, dF(t).$$

Then

$$\text{(7.2.23)} \qquad f_\alpha^{(k)}(s) = \inf_{a\geq 0} \{M\delta(a) + \tilde{w}_\alpha^{(k)}(s+a) - ms\}.$$

From the previous theorem follows also that the policy $R_*^{(k)}(.,.)$ defined by

$$R_*^{(k)}(x,t) = \begin{cases} S_\alpha^{(k)} - \pi_t x & \text{if } \pi_t x < s_\alpha^{(k)} \\ 0 & \text{otherwise} \end{cases}$$

is α_k-discounted optimal. Moreover $R_*^{(k)}(x,t)$ minimizes the right hand side of (7.2.23) if $\pi_t x = s$.
Hence

$$f_\alpha^{(k)}(s) = \begin{cases} M + \tilde{w}_\alpha^{(k)}(S_\alpha^{(k)}) - ms & \text{for } s < s_\alpha^{(k)} \\ \tilde{w}_\alpha^{(k)}(s) - ms & \text{for } s \geq s_\alpha^{(k)} \end{cases}$$

which in turn implies that $\tilde{w}_\alpha^{(k)}(.)$ attains its infimum at $S_\alpha^{(k)}$, while

$$\text{(7.2.24)} \qquad \tilde{w}_\alpha^{(k)}(s_\alpha^{(k)}) \leq M + \tilde{w}_\alpha^{(k)}(S_\alpha^{(k)}).$$

For $y \geq x$ we have by (7.2.11)

$$mx + f_\alpha^{(k)}(x) \leq my + M + f_\alpha^{(k)}(y).$$

Hence

$$\tilde{w}_\alpha^{(k)}(y) - \tilde{w}_\alpha^{(k)}(x) \geq \frac{b(y)-b(x)}{k(1-\alpha_k)+\alpha_k\nu} - M \geq$$

$$\geq (\alpha+\nu)^{-1} (b(y)-b(x)) - M.$$

Since $\tilde{w}_\alpha^{(k)}(.)$ attains its infimum at $S_\alpha^{(k)}$ it follows that

$$0 \leq S_\alpha^{(k)} \leq M(\alpha+\nu)b^{-1}.$$

Moreover, (7.2.13) implies that for $x \leq y \leq 0$

$$f_\alpha^{(k)}(x) - f_\alpha^{(k)}(y) \geq -m(x-y).$$

Hence we have for $x \leq y \leq 0$

$$\tilde{w}_\alpha^{(k)}(x) - \tilde{w}_\alpha^{(k)}(y) \geq m(x-y) + \frac{b(x)-b(y)}{k(1-\alpha_k)+\alpha_k\nu} - \frac{m(x-y)\alpha_k\nu}{k(1-\alpha_k)+\alpha_k\nu} \geq$$

$$\geq \frac{p(y-x)}{k(1-\alpha_k)+\alpha_k\nu} + \frac{m(x-y)k(1-\alpha_k)}{k(1-\alpha_k)+\alpha_k\nu}.$$

From this inequality follows for α small enough and for $s \leq 0$

$$\tilde{w}_\alpha^{(k)}(s) \geq \tilde{w}_\alpha^{(k)}(0) - (\alpha+\nu)^{-1}(p-m\alpha)s \geq$$

$$\geq \tilde{w}_\alpha^{(k)}(S_\alpha^{(k)}) - (\alpha+\nu)^{-1}(p-m\alpha)s.$$

This implies with (7.2.24).

$$s_\alpha^{(k)} \geq -M(\nu+\alpha)(p-m\alpha)^{-1}. \qquad \Box$$

THEOREM 7.2.13. Consider the CTMDP with some initial distribution P_0 on S, such that (7.2.3) holds. For all $\alpha > 0$ there exists a stationary deterministic policy (V_α, R_α) such that

(i) (V_α, R_α) is α-discounted optimal in the class of strong regular policies.

(ii) (V_α, R_α) is of (s,S) type.

PROOF. Let $S_\alpha^{(k)}$ and $s_\alpha^{(k)}$ be defined by (7.2.15) and (7.2.16). From the previous lemma follows the existence of real numbers S_α and s_α and a sequence $(k_j)_{j=1}^\infty$ of natural numbers such that

$$(7.2.25)\qquad \lim_{j\to\infty} S_\alpha^{(k_j)} =: S_\alpha^*$$

$$(7.2.26)\qquad \lim_{j\to\infty} s_\alpha^{(k_j)} =: s_\alpha^*.$$

Put

$$(7.2.27)\qquad V_\alpha := (-\infty, s_\alpha^*]$$

$$(7.2.28)\qquad R_\alpha(x,t) := [S_\alpha^* - \pi_t x]^+.$$

Then (V_α, R_α) is a well-defined policy for the CTMDP, which is obviously stationary, deterministic and of (s,S) type. According to theorem 4.2.9. the policy (V_α, R_α) is α-discounted optimal for the CTMDP if (V_α, R_α) and $(R_*^{(k_j)})_{j=1}^\infty$ satisfy condition (4.1.8). Theorem 4.2.1. gives sufficient conditions for (4.1.8). Finally theorem 2.5.4 and the proof of proposition 7.2.2. show that these sufficient conditions are satisfied for (V_α, R_α) and $(R_*^{(k_j)})_{j=1}^\infty$. □

7.3. THE AVERAGE COST CASE.

THEOREM 7.3.1. Consider the CTMDP with some initial distribution P_0 on S such that (7.2.3) holds. Then there exists a stationary deterministic policy (V,R) such that

(i) (V,R) is average optimal in the class of strong regular policies.

(ii) (V,R) is of (s,S) type.

PROOF. Let S_α^* and s_α^* for $\alpha > 0$ be defined by (7.2.25) and (7.2.26). By lemma 7.2.12. there exists for any sequence $(\beta_k)_{k=1}^\infty$ with $\lim_{k\to\infty}\beta_k = 0$ a subsequence $(\beta_{k_j})_{j=1}^\infty$ such that

$$S^* := \lim_{j\to\infty} S_{\beta_{k_j}}^*$$

and

$$s^* := \lim_{j\to\infty} s_{\beta_{k_j}}^*$$

exist.

Moreover it follows from lemma 7.2.12 that $0 \le S^* \le M\nu b^{-1}$ and $-M\nu p^{-1} \le s^* \le 0$.

Put

$$V: = (-\infty, s^*]$$

and

$$R(x,t): = [S^* - \pi_t x]^+.$$

Obviously (V,R) is a well-defined, deterministic stationary policy of (s,S) type. To show that (V,R) is average optimal in the class of strong regular policies it is sufficient to show that conditions (4.4.6) and (4.4.7) of theorem 4.4.6. hold.

Consider the policy (V_α, R_α) defined by (7.2.27) and (7.2.28). Using standard renewal arguments, it can be shown that (see e.g. VEINOTT and WAGNER (1965))

$$c(V_\alpha, R_\alpha) = \nu g(s_\alpha^*, S_\alpha^*)(1+M_F(S_\alpha^* - s_\alpha^*))^{-1}, \tag{7.3.1}$$

where $M_F(.)$ is the renewal function of the probability distribution F and

$$\begin{aligned} g(u,v): = {} & b(v)\nu^{-1} + M\{1-F(v-u)\} + m\int_{v-u}^{\infty} t dF(t) + \\ & + \int_0^{v-u} \{b(v-y)\nu^{-1} + \{1-F(v-u-y)\}\{M+my\} + \\ & + m\int_{v-u-y}^{\infty} t dF(t)\} \, dM_F(y) \qquad \text{for } u \le v. \end{aligned}$$

Together (7.3.1) and assumption 7.1.1.(i) imply (4.4.6).

Finally we put for $x \in J[0,\infty)$

$$\tau_{n,\alpha}(x): = \text{epoch of } n^{\text{th}} \text{ entrance of } x \text{ into state } S_\alpha^*.$$

Then the policy (V_α, R_α) induces on $J[0,\infty)$ a regenerative stochastic process X_α, with sequence of regeneration epochs $(\tau_{n,\alpha}(X_\alpha))_{n=0}^{\infty}$. One easily verifies that for this regenerative stochastic process the conditions (4.4.13), (4.4.14) and (4.4.15) hold. Hence it follows from theorem 4.4.8. that (4.4.7) holds. □

CHAPTER 8

RELATED LITERATURE

8.1. INTRODUCTION.

In this final chapter we give a brief overview of those parts of the literature, that are closely related in method or result to the work presented in this monograph.
The reader should not expect a complete picture of all the literature on Markov decision processes with continuous time parameter, nor an overview of all those papers, in which approximation methods based on weak convergence are used. The literature that is considered, is separated into three sections. In section 8.2 we mention publications in which Markov decision processes with continuous time parameter are treated from a general point of view. In section 8.3 literature on approximation methods for continuous time Markov decision processes is gathered. Finally, we give in section 8.4 some references with respect to the specific models presented in the chapters 5, 6 and 7.

8.2. MARKOV DECISION PROCESSES WITH CONTINUOUS TIME PARAMETER FROM A GENERAL POINT OF VIEW.

During the first ten years after the initiating work of BELLMAN (1957) and HOWARD (1960) on Markov decision processes, the attention was mainly directed to the theory of discrete time- and semi-Markov decision processes. For references we refer to ROSS (1970) and HORDIJK (1974). The theory of semi-Markov decision processes is closely connected with the theory of Markov decision drift processes with continuous time parameter. Indeed, any CTMDP with drift function constant in its time variable becomes a semi-Markov decision process if the class of policies is restricted to the stationary ones. Hence, the theory of semi-Markov decision processes is applicable, once we know for a certain CTMDP that a stationary optimal policy exists.

MILLER (1968) considered continuous time Markov decision processes with finite state space S and finite set of controls A_1, while in his analysis the set of impulsive controls is empty. The cost functional that is considered is the 0-discounted, T-horizon cost functional. The class of policies is restricted to the pure, memoryless ones. By a constructive proof Miller shows, that within this class there exists a piecewise constant 0-discounted, T-horizon optimal policy i.e. for all $i \in S$ there exists a piecewise constant function $r(i,.)$ on $[0,T]$ with values in A_1, such that the policy R, defined by $R(x,t):= r(\pi_t x,t)$ is 0-discounted, T-horizon optimal.

In a second paper MILLER (1968[a]) treats the α-discounted and average cost functionals for continuous time Markov decision processes with finite state and action space. The existence of a stationary, α-discounted optimal and a stationary average optimal policy in the class of piecewise constant policies is established. These results are also obtainable from our theorems 4.2.1. and 4.4.6., using the corresponding well-known results for discrete time processes.

KAKUMANU (1971) generalizes the results of Miller to the countable state space. He also assumes an empty set of impulsive controls. In the case in which the set of controls is finite, the existence of a stationary, α-discounted optimal policy, within the class of pure memoryless policies is established. Moreover, he proves for a countable set of controls, that within the class of pure, memoryless policies there exists a stationary, α-discounted, ε-optimal policy. These results can also be obtained from our theorem 4.2.1. with some additional argumentation.

KAKUMANU (1975) considered in a subsequent paper the average cost functional for the case of a countable state space, finite set of controls and empty set of impulsive controls. The existence of a stationary, average optimal policy in the class of memoryless policies is proved, under the condition that the probability measure induced by any pure, memoryless policy generates a positive recurrent Markov process, with only one recurrent class. However, his conditions seem to be insufficient, since his results contradict a counter example of FISHER and ROSS (1968). Our conditions of theorem 4.4.6. which ensure the average optimality of a policy for the continuous time model are stronger than those of Kakumanu. DOSHI (1974 and 1976) deals with continuous time Markov decision (drift) processes, with the restrictions that the set of impulsive controls is empty and the drift function is constant in

its time variable. He derives conditions, which guarantee the existence of a stationary, α-discounted optimal and a stationary, average optimal policy, within the class of pure, memoryless policies. This kind of result (for the general state space) cannot be obtained from our theory without any additional information about the optimal policies for the approximating discrete time processes. The numerous conditions of Doshi, however, are not easy to verify.
PLISKA (1975) considers continuous time Markov decision (drift) processes under the same restrictions as those of DOSHI (1976). He shows, that for the 0-discounted, T-horizon cost functional, as well as the 0-discounted cost functional the optimal value function is the unique solution of a certain differential equation.
Finally, YUSHKEVICH (1977) treats continuous time Markov decision processes with countable state space and (logically) a drift function that is constant in its time variable. He takes into account the class of pure policies. His main result is the derivation of sufficient conditions, which guarantee the existence of a pure, memoryless, 0-discounted, ε-optimal policy in the class of pure policies. We obtained a similar result in chapter 3.
As YUSHKEVICH (1977) is the only author who admits history remembering policies, only he is concerned with the question of the existence of a probability measure on the space of all sample paths, induced by a general (non-Markovian) policy. MANDL (1973) had earlier partially answered this question for the case of a finite state space. He defined the transition probabilities of the embedded process, describing the jump times and jump states. For a finite state space and empty set of impulsive controls the transition probabilities $Q^{(n)}(.)$ defined on page 38 of this monograph (definition 2.2.11.) simplify to the transition probabilities given by Mandl.
In dealing with the problem of the existence of stationary *average* optimal policies for discrete or continuous time Markov decision (drift) processes we can roughly distinguish two approaches. The first approach takes the optimality equation for the average cost criterion as a starting point. The difficult part of that analysis is to find sufficient conditions under which this optimality equation has a bounded solution. Some authors provide sets of sufficient conditions in terms of the solution of the optimality equation for the discounted cost criterion (TAYLOR (1965), ROSS (1968,

1968a, 1970a), TIJMS (1975), HORDIJK (1976) and DOSHI (1976)). Also recurrence conditions on the underlying stochastic processes are proposed as sufficient for the optimality equation to have a bounded solution (DERMAN (1966), DERMAN and VEINOTT (1967), HORDIJK (1974), FEDERGRUEN, HORDIJK and TIJMS (1979) and FEDERGRUEN, SCHWEITZER and TIJMS (1980)).

In the second approach the value function of the average cost problem is directly related to the value function of the discounted problem. In SCHÄL (1977) this is established by interchanging limit and infimum while BLACKWELL (1962), MILLER (1968) and HORDIJK (1971) make use of a well known Abelian theorem. It is worthwhile to notice that in this Abelian approach no use is made of the optimality equation. In section 4.5. of this monograph we used the latter method to find sufficient conditions for the average optimality of a limit point of a sequence of discounted optimal policies.

8.3. APPROXIMATION METHODS.

Approximation methods for continuous time Markov decision processes have been applied during the last decade with considerable success. In view of the fast growing literature in which these approximation methods are used, they seem to become a main tool in the analysis of Markov decision processes. In particular the theory on weak convergence of probability measures plays a key role.
For applications before 1974 in queueing theory and other areas of applied probability the reader is referred to the survey papers by IGLEHART (1973 and 1974). In this section we briefly discuss some publications in which continuous time Markov decision processes are treated with employment of approximation methods with respect to the time parameter. Beyond the scope of this monograph are the important methods which approximate or discretize the input parameters or the state space.
MITCHEL (1973) considers an M/G/1 queueing system with controllable service rate under the average cost functional. He proves, under certain assumptions on the cost functions, the existence of a stationary, ε-optimal, monotone policy i.e. the service rate used increases with the amount of work in the system. First he proves, via the α-discounted, finite horizon analysis the existence of a monotone, stationary, average

optimal policy for the k th approximating decision process. Next, he shows that given k and a policy R for the CTMDP, there exists a policy R_k for the k th approximating DTMDP, such that the expected average costs under both policies differ by a function of k, which goes to zero as $k \to \infty$. Finally he proves by a method depending on the structure of the problem, that the optimal value function for the sequence of approximating decision processes converges to the optimal value function of the CTMDP. This justifies the conclusion, that there exists a monotone, stationary, ε-optimal policy for the CTMDP.

WINSTON (1976) considered continuous time Markov decision processes with finite state space, finite set of controls and empty set of impulsive controls. By analysis of the discrete time and continuous time optimality equations he proves for these processes our theorem 4.2.9. under the assumption that the optimal policies for the approximating decision processes are stationary. Winston applies this result to a maintenance system, consisting of a finite number of machines and a single server repair facility, that can be operated at a finite number of rates. Conditions are derived which ensure, that the optimal repair rate for the CTMDP is a non-increasing function of the number of machines in good condition.

The results of Winston are generalized in chapter 10 of WHITT (1975). He proves, besides a lot of other interesting convergence results on input parameters, a theorem similar to theorem 2.4.3. in this monograph. The processes considered by Whitt have countable state space and empty set of impulsive controls, while the policies are restricted to the class of pure memoryless policies.

LIPPMAN (1976) considers continuous time Markov decision processes with countable state space, finite (state-depending) sets of controls and empty set of impulsive controls. He proposes a procedure by which a CTMDP is approximated by a sequence of semi-Markov processes.

Finally, we mention a paper by KAKUMANU (1977), in which he, in stead of using an approximation procedure, constructs for every CTMDP a single DTMDP, such that the total expected α-discounted costs for the CTMDP and the total expected β-discounted costs for the DTMDP are proportional (β depends only on α and the model parameters). The proof of this result is based on an analysis of the optimality equation.

8.4. PROCESSES WITH SPECIAL STRUCTURE.

In this section we mention some papers, in which, although with quite different methods, results are obtained related to our results on the specific models of chapters 5, 6 and 7.
With respect to the M/M/1 queueing model with controllable arrival and service rate (chapter 5) we mention a paper by SERFOZO (1981) in which he considers a model which differs in three aspects from our model. The possible arrival- and service rates in Serfozo's model can be chosen from a finite set, in our model from a finite interval. In Serfozo's model the rates can only be changed when a customer arrives or a service is completed, while we considered the continuous time control case. Finally, Serfozo proves the existence of a monotone α-discounted optimal and a monotone average optimal policy, within the class of stationary policies, while we obtained the same results in the class of strong regular policies.
For an overview of the literature on controlled queueing systems upto 1974 we refer to PRABHU and STIDHAM (1974).
Regarding the maintenance replacement model (chapter 6) we mention papers by TAYLOR (1975) and ZUCKERMAN (1977). TAYLOR (1975) treats the continuous time replacement model, where the only possible control of the system is replacement of an old device by a new one. He proves the existence of an average optimal control limit type policy.
ZUCKERMAN (1977) extends the results of Taylor for more general cost functions. Moreover, he gets a similar result for the α-discounted cost functional. The results of Zuckerman are more general than ours with respect to the assumptions on the cost functions. However, we assumed that besides by replacement the system can be controlled by maintenance at several levels. Moreover, in our model a continuous decay occurs besides the shockswise decay. Zuckerman considers the model without maintenance, while the damage accumulates only by shocks.
For an overview of the literature on maintenance replacement models before 1975 we refer to a survey paper by PIERSKALLA and VOELKER (1976).
The conditions for optimality of (s,S) type policies in discrete time inventory models (chapter 7) have been discussed in detail in the literature. As far as the author knows, a mathematical satisfactory proof of the corresponding result for the continuous time inventory model has never been given before.

REFERENCES

BELLMAN, R. (1957). *Dynamic programming*. Princeton University Press, Princeton.

BILLINGSLEY, P. (1968). *Convergence of probability measures*. Wiley, New York.

BLACKWELL, D. (1962). *Discrete dynamic programming*. Ann. Math. Statist. 33, 719-726.

CINLAR, E. and PINSKY, M. (1971). *A stochastic integral in storage theory*. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 17, 227-240.

CINLAR, E. (1975). *Introduction to stochastic processes*. Prentice Hall, New Jersey.

COHEN, J.W. (1976). *On regenerative processes in queueing theory*. Lecture notes in Economics and Mathematical Systems, 121. Springer Verlag, Berlin.

CRABILL, T. (1972). *Optimal control of a service facility with variable exponential service time and constant arrival rate*. Management Sci. 18, 560-566.

DERMAN, C. and STRAUCH, R.E. (1966). *A note on memoryless rules for controlling sequential control processes*. Ann. Math. Statist. 37, 276-287.

DERMAN, C. (1966). *Denumerable state Markovian decision processes; average cost criterion*. Ann. Math. Statist. 37, 1545-1553.

DERMAN, C. and VEINOTT, A.F. jr. (1967). *A solution to a countable system of equations arising in Markovian decision processes*. Ann. Math. Statist. 38, 582-584.

DOSHI. B.T. (1974). *Continuous time control of Markov processes on an arbitrary state space*. Tech. Summ. Report, 1468. Mathematical Research Centre, Madison, Wisconsin.

DOSHI, B.T. (1976). *Continuous time control of Markov processes on an arbitrary state space: average return criterion*. Stochastic Processes Appl. 4, 55-77.

DUDLEY, R.M. (1968). *Distances of probability measures and random variables*. Ann. Math. Statist. 39, 1563-1572.

FEDERGRUEN, A., HORDIJK, A. and TIJMS, H.C. (1979). *Denumerable state semi-Markov decision processes with unbounded costs, average cost criterion*. Stochastic Processes Appl. 9, 223-235.

FEDERGRUEN, A., SCHWEITZER, P.J. and TIJMS, H.C. (1980). *Denumerable undiscounted semi-Markov decision processes with unbounded regions*.

FELLER, W. (1966). *An introduction to probability theory and its applications II*. Wiley, New York.

FISHER, L. and ROSS, S.M. (1968). *An example in denumerable decision processes*. Ann. Math. Statist. 39, 674-675.

HORDIJK, A. (1971). *A sufficient condition for the existence of an optimal policy with respect to the average cost criterion in Markovian decision processes*. Transactions of the sixth Prague conference on information theory. Academia, Prague.

HORDIJK, A. (1974). *Dynamic programming and Markov potential theory*. Mathematical Centre Tract, 51. Mathematical Centre, Amsterdam.

HORDIJK, A. (1976). *Regenerative Markov decision models*. Mathematical Programming Study 6, 49-72. North-Holland, Amsterdam.

HOWARD, R.A. (1960). *Dynamic programming and Markov processes*. The Massachusetts Technology Press, Cambridge, Massachusetts.

IGLEHART, D.L. (1963). *Optimality of (s,S) policies in the infinite horizon dynamic inventory problem*. Management Sci. 9, 259-267.

IGLEHART, D.L. (1973). *Weak convergence in queueing theory*. Advances in Appl. Probability 5, 570-594.

IGLEHART, D.L. (1974). *Weak convergence in applied probability*. Stochastic Processes Appl. 2, 211-241.

JEWELL, W. (1963). *Markov renewal programming I and II*. Operations Res. 2, 938-971.

KAKUMANU, P. (1971). *Continuously discounted Markov decision model with countable state and action space*. Ann. Math. Statist. 42, 919-926.

KAKUMANU, P. (1975). *Continuous time Markovian decision processes; average return criterion*. J. Math. Anal. Appl. 52, 173-188.

KAKUMANU, P. (1977). *Relation between continuous and discrete time Markovian decision problems*. Naval Res. Logist. Quart. 24, 431-439.

KURATOWSKI, K. (1966). *Topology I*. Academic Press, New York.

KUSHNER, H.J. (1977). *Probability methods for approximations in stochastic control and for elliptic equations*. Academic Press, New York.

LANGEN, H.J. (1981). *Convergence of dynamic programming models*. Math. Oper. Res. 6, 493-512.

LINDVALL, T. (1973). *Weak convergence of probability measures and random functions in the function space* $D[0,\infty)$. J. Appl. Probability 10, 109-121.

LIPPMAN, S.A. (1973). *Semi-Markov decision processes with unbounded rewards*. Management Sci. 19, 717-731.

LIPPMAN, S.A. (1975). *Applying a new device in the optimization of exponential queueing systems*. Operations Res. 23, 687-710.

LIPPMAN, S.A. (1976). *Countable state continuous time dynamic programming with structure*. Operations Res. 24, 477-490.

LOW, D.W. (1974). *Optimal dynamic pricing policies for an M/M/s queue*. Operations Res. 22, 545-561.

MANDL, P. (1973). *On the adaptive control of finite state Markov processes*. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 27, 263-276.

MILLER, B.L. (1968). *Finite state continuous time Markov decision processes with a finite planning horizon*. SIAM J. Control 6, 266-280.

MILLER, B.L. (1968[a]). *Finite state continuous time Markov decision processes with an infinite planning horizon*. J. Math. Anal. Appl. 22, 552-569.

MITCHELL, B. (1973). *Optimal service rate selection in an* $M/G/\hat{1}$ *queue*. SIAM J. Appl. Math. 24, 19-35.

NEVEU, J. (1965). *Mathematical foundations of the calculus of probability*. Holden-Day, San Francisco.

PARTHASARATHY, K.R. (1967). *Probability measures on metric spaces*. Academic Press, New York.

PIERSKALLA, W.P. and VOELKER, J.A. (1976). *A survey of maintenance models: the control and surveillance of deteriorating systems*. Naval Res. Logist. Quart. 23, 353-388.

PLISKA, S.R. (1975). *Controlled jump processes*. Stochastic Processes Appl. 3, 259-282.

PORTEUS, E.L. (1971). *On the optimality of generalized (s,S) policies*. Management Sci. 17, 411-426.

PRABHU, N.U. and STIDHAM, S., Jr. (1974). *Optimal control of queueing systems*. In: Mathematical methods in queueing theory, 263-294. Lecture notes in Economics and Mathematical Systems, 98. Springer Verlag, Berlin.

PROHOROV, J.V. (1956). *Convergence of random processes and limit theorems in probability theory*. Theor. Probability Appl. 1, 157-214.

ROBIN, M. (1978). *Contrôle impulsionnel des processus de Markov*. Thesis, Paris IX.

ROSS, S.M. (1968). *Non-discounted denumerable Markovian decision models*. Ann. Math. Statist. 39, 412-423.

ROSS, S.M. (1968a). *Arbitrary state Markovian decision processes*. Ann. Math. Statist. 39, 2118-2122.

ROSS, S.M. (1970). *Applied probability models with optimization applications*. Holden-Day, San Francisco.

ROSS, S.M. (1970a). *Average cost semi-Markov decision processes*. J. Appl. Probability 7, 649-656.

ROYDEN, H.L. (1968). *Real analysis*. Macmillan, New York.

SABETI, H. (1973). *Optimal selection of service rates in queueing with different costs*. J. Operations Res. Soc. Japan 16, 15-35.

SCARF, H. (1960). *The optimality of (S,s) policies in the dynamic inventory problem*. Mathematical methods in the Social sciences. Stanford University Press, Stanford.

SCHÄL, M. (1976). *On the optimality of (s,S) policies in dynamic inventory models with finite horizon*. SIAM J. Appl. Math. 30, 528-537.

SCHÄL, M. (1977). *On negative dynamic programming with irreducible Markov chains and the average cost criterion*. Bonner Mathematische Schriften, 98, 93-97.

SERFOZO, R.F. (1981). *Optimal control of random walks, birth and death processes and queues*. Advances in Appl. Probability 13, pp. 61-83.

SKOROHOD, A.V. (1956). *Limit theorems for stochastic processes*. Theor. Probability Appl. 1, 261-290.

STONE, C. (1963). *Weak convergence of stochastic processes defined on semi-infinite time intervals*. Proc. Amer. Math. Soc. 14, 694-696.

STONE, C.J. (1972). *An upper bound for the renewal function*. Ann. Math. Statist. 43, 2050-2052.

STOYAN, D. (1977). *Qualitative Eigenschaften und Abschätzungen stochastischer Modelle*. Akademie-Verlag, Berlin.

TAYLOR, H.M. (1965). *Markovian sequential replacement processes*. Ann. Math. Statist. 36, 1677-1694.

TAYLOR, H.M. (1975). *Optimal replacement under additive damage and other failure models*. Naval Res. Logist. Quart. 22, 1-18.

TITCHMARSH, E.C. (1939). *The theory of functions*. Oxford University Press, Oxford.

TIJMS, H.C. (1975). *On dynamic programming with arbitrary state space, compact action space and the average return as criterion*. Research Report BW 55/75. Mathematical Centre. Amsterdam.

TIJMS, H.C. and VAN DER DUYN SCHOUTEN, F.A. (1978). *Inventory control with two switch-over levels for a class of M/G/1 queueing systems with variable arrival and service rate*. Stochastic Processes Appl. 6, 213-222.

VEINOTT, A.F.,Jr. and WAGNER, H.M. (1965). *Computing optimal (s,S) inventory policies*. Management Sci. 11, 525-552.

VEINOTT, A.F.,Jr. (1966). *On the optimality of (s,S) inventory policies: new conditions and a new proof*. SIAM J. Appl. Math. 14, 1067-1083.

WHITT, W. (1970). *Weak convergence of probability measures on the function space* $C[0,\infty)$. Ann. Math. Statist. 41, 939-944.

WHITT, W. (1975). *Continuity of Markov processes and dynamic programs*. School of Organization and Management, Yale University, New Haven.

WHITT, W. (1980). *Some useful functions for functional limit theorems*. Math. Oper. Res. 5, 67-85.

WICHURA, M.J. (1970). *On the construction of almost uniformly convergent random variables with given weakly convergent image laws*. Ann. Math. Statist. 41, 284-291.

WIDDER, D.V. (1946). *The Laplace transform*. Princeton University Press, Princeton.

WINSTON, W. (1976). *Optimal control of discrete and continuous time maintenance systems with variable service rates*. Indiana University, Indiana.

YUSHKEVICH, A.A. (1977). *Controlled Markov models with countable state space and continuous time*. Theor. Probability Appl. 22, 215-234.

YUSHKEVICH, A.A. and FAINBERG, E.A. (1979). *Onhomogeneous Markov models with continuous time and finite or countable state space.* Theor. Probability Appl. 24, 156-161.

ZABEL, E. (1962). *A note on the optimality of (S,s) policies in inventory theory*. Management Sci. 9, 123-125.

ZUCKERMAN, D. (1977). *Replacement models under additive damage*. Naval Res. Logist. Quart. 24, 549-558.

INDEX OF SYMBOLS AND NOTATIONS

(Symbols with only local significance are not included in this list).

INDEX OF TERMS

SUMMARY

This monograph deals with the optimal control of Markov decision drift processes with continuous time parameter (CTMDP). Such a decision process can be described as follows. A stochastic process $\{X_t, t \geq 0\}$ on a given metric space is continuously observed and at each $t \in [0,\infty)$ an action is chosen based on the history of the process upto time t. The process $\{X_t, t \geq 0\}$ is assumed to be a jump process with a deterministic drift between two successive jumps. There are two kinds of actions available, which differ in their impact on the evolution of the process. The (generator) controls are actions which affect the process only in its infinitesimal parameters (e.g. the arrival rate in an arrival process), while the impulsive controls are actions which have an impulsive influence on the process; i.e. an impulsive control causes an immediate change in the state of the system (e.g. replacement in a maintenance model). At time t a cost is incurred at a rate depending on the actual state of the system and the control that is chosen. Moreover, when an impulsive control is chosen, a lump cost is incurred, which depends on the actual impulsive control and the state of the system. A policy is a rule for choosing actions. That is, given the history of the process and the present state at time t, the policy prescribes the control and impulsive control to be chosen.
Our investigation of a CTMDP proceeds along the following lines. First the class of decision processes with discrete time parameter (DTMDP) is introduced. A DTMDP is a decision process that is not continuously observed and controlled, but only on equidistant time points. For any CTMDP a sequence of approximating DTMDP's is constructed, for which the distance between two successive decision epochs decreases to zero. It is shown that for any CTMDP-policy, with certain regularity conditions, there exist policies for the sequence of approximating DTMDP's, such that the induced stochastic processes converge weakly on the space of all possible sample paths.
This type of result can be applied to transfer properties of a DTMDP to a CTMDP. For example, the powerful method of mathematical induction by which the structure of optimal policies in discrete time problems can be determined is not applicable for continuous time problems. However, by the above stated convergence result, the method of mathematical induction becomes again the basis for the analysis of a CTMDP.
We collected the necessary material on weak convergence of probability

measures on metric spaces in chapter 1. Moreover, we introduce and analyse in this chapter a sequence space, which can be used as the space of the possible sample paths for the processes to be considered in this monograph.

In chapter 2 we formally introduce a CTMDP and DTMDP, the class of possible policies is defined and it is shown that any policy induces a unique probability measure on the space of possible sample paths. Theorem 2.4.3. constitutes the main result of this chapter. This theorem provides conditions for a CTMDP-policy, which ensure the existence of policies for the sequence of approximating DTMDP's, such that weak convergence of the induced probability measures occurs.

Using the results of chapter 2 we prove in chapter 3 a theorem for a CTMDP, which justifies the frequently made assumption that history remembering policies do not yield any improvement in the expected (discounted or average) costs.

For a number of applications the results of chapter 2 seem to be insufficient. From the weak convergence of the induced probability measures we cannot conclude the convergence of the total expected discounted costs, when the cost functions are unbounded. Hence, we give in chapter 4 (theorem 4.2.1.), for the model with unbounded cost functions, conditions which ensure not only the weak convergence of the induced probability measures, but also the convergence of the total expected discounted costs. From this theorem sufficient conditions for the discounted optimality of a CTMDP-policy are deduced. Using the results for the discounted cost criterion together with a well-known Abelian theorem, we derive also conditions, which guarantee the average optimality of a CTMDP-policy (theorem 4.4.6.).

The convergence theorems of the chapters 2 and 4 can be successfully applied to transpose results to a CTMDP, which are alreay known for a corresponding DTMDP (or which can be obtained by standard methods). Especially, structural properties of optimal policies, which are usually obtained for a DTMDP by the method of mathematical induction, can be transposed in this way. As illustration we consider in chapter 5, 6 and 7 three quite different continuous time Markov decision processes on which the proposed technique is applied. Successively, we treat an M/M/1 waiting line model, a maintenance replacement model and an inventory model. Structural results for the optimal policies are obtained for the α-discounted as well as the average cost criterion.

TITLES IN THE SERIES MATHEMATICAL CENTRE TRACTS

(An asterisk before the MCT number indicates that the tract is under preparation).

A leaflet containing an order form and abstracts of all publications mentioned below is available at the Mathematisch Centrum, Kruislaan 413, 1098 SJ Amsterdam, The Netherlands. Orders should be sent to the same address.

MCT 1 T. VAN DER WALT, *Fixed and almost fixed points*, 1963. ISBN 90 6196 002 9.

MCT 2 A.R. BLOEMENA, *Sampling from a graph*, 1964. ISBN 90 6196 003 7.

MCT 3 G. DE LEVE, *Generalized Markovian decision processes, part I: Model and method*, 1964. ISBN 90 6196 004 5.

MCT 4 G. DE LEVE, *Generalized Markovian decision processes, part II: Probabilistic background*, 1964. ISBN 90 6196 005 3.

MCT 5 G. DE LEVE, H.C. TIJMS & P.J. WEEDA, *Generalized Markovian decision processes, Applications*, 1970. ISBN 90 6196 051 7.

MCT 6 M.A. MAURICE, *Compact ordered spaces*, 1964. ISBN 90 6196 006 1.

MCT 7 W.R. VAN ZWET, *Convex transformations of random variables*, 1964. ISBN 90 6196 007 X.

MCT 8 J.A. ZONNEVELD, *Automatic numerical integration*, 1964. ISBN 90 6196 008 8.

MCT 9 P.C. BAAYEN, *Universal morphisms*, 1964. ISBN 90 6196 009 6.

MCT 10 E.M. DE JAGER, *Applications of distrubutions in mathematical physics*, 1964. ISBN 90 6196 010 X.

MCT 11 A.B. PAALMAN-DE MIRANDA, *Topological semigroups*, 1964. ISBN 90 6196 011 8.

MCT 12 J.A.Th.M. VAN BERCKEL, H. BRANDT CORSTIUS, R.J. MOKKEN & A. VAN WIJNGAARDEN, *Formal properties of newspaper Dutch*, 1965. ISBN 90 6196 013 4.

MCT 13 H.A. LAUWERIER, *Asymptotic expansions*, 1966, out of print; replaced by MCT 54.

MCT 14 H.A. LAUWERIER, *Calculus of variations in mathematical physics*, 1966. ISBN 90 6196 020 7.

MCT 15 R. DOORNBOS, *Slippage tests*, 1966. ISBN 90 6196 021 5.

MCT 16 J.W. DE BAKKER, *Formal definition of programming languages with an application to the definition of ALGOL 60*, 1967. ISBN 90 6196 022 3.

MCT 17 R.P. VAN DE RIET, *Formula manipulation in ALGOL 60, part 1*, 1968. ISBN 90 6196 025 8.

MCT 18 R.P. VAN DE RIET, *Formula manipulation in ALGOL 60, part 2*, 1968. ISBN 90 6196 038 X.

MCT 19 J. VAN DER SLOT, *Some properties related to compactness*, 1968. ISBN 90 6196 026 6.

MCT 20 P.J. VAN DER HOUWEN, *Finite difference methods for solving partial differential equations*, 1968. ISBN 90 6196 027 4.

MCT 21 E. WATTEL, *The compactness operator in set theory and topology*, 1968. ISBN 90 6196 028 2.

MCT 22 T.J. DEKKER, *ALGOL 60 procedures in numerical algebra, part 1*, 1968. ISBN 90 6196 029 0.

MCT 23 T.J. DEKKER & W. HOFFMANN, *ALGOL 60 procedures in numerical algebra, part 2*, 1968. ISBN 90 6196 030 4.

MCT 24 J.W. DE BAKKER, *Recursive procedures*, 1971. ISBN 90 6196 060 6.

MCT 25 E.R. PAËRL, *Representations of the Lorentz group and projective geometry*, 1969. ISBN 90 6196 039 8.

MCT 26 EUROPEAN MEETING 1968, *Selected statistical papers, part I*, 1968. ISBN 90 6196 031 2.

MCT 27 EUROPEAN MEETING 1968, *Selected statistical papers, part II*, 1969. ISBN 90 6196 040 1.

MCT 28 J. OOSTERHOFF, *Combination of one-sided statistical tests*, 1969. ISBN 90 6196 041 X.

MCT 29 J. VERHOEFF, *Error detecting decimal codes*, 1969. ISBN 90 6196 042 8.

MCT 30 H. BRANDT CORSTIUS, *Exercises in computational linguistics*, 1970. ISBN 90 6196 052 5.

MCT 31 W. MOLENAAR, *Approximations to the Poisson, binomial and hypergeometric distribution functions*, 1970. ISBN 90 6196 053 3.

MCT 32 L. DE HAAN, *On regular variation and its application to the weak convergence of sample extremes*, 1970. ISBN 90 6196 054 1.

MCT 33 F.W. STEUTEL, *Preservation of infinite divisibility under mixing and related topics*, 1970. ISBN 90 6196 061 4.

MCT 34 I. JUHÁSZ, A. VERBEEK & N.S. KROONENBERG, *Cardinal functions in topology*, 1971. ISBN 90 6196 062 2.

MCT 35 M.H. VAN EMDEN, *An analysis of complexity*, 1971. ISBN 90 6196 063 0.

MCT 36 J. GRASMAN, *On the birth of boundary layers*, 1971. ISBN 90 6196 064 9.

MCT 37 J.W. DE BAKKER, G.A. BLAAUW, A.J.W. DUIJVESTIJN, E.W. DIJKSTRA, P.J. VAN DER HOUWEN, G.A.M. KAMSTEEG-KEMPER, F.E.J. KRUSEMAN ARETZ, W.L. VAN DER POEL, J.P. SCHAAP-KRUSEMAN, M.V. WILKES & G. ZOUTENDIJK, *MC-25 Informatica Symposium* 1971. ISBN 90 6196 065 7.

MCT 38 W.A. VERLOREN VAN THEMAAT, *Automatic analysis of Dutch compound words*, 1971. ISBN 90 6196 073 8.

MCT 39 H. BAVINCK, *Jacobi series and approximation*, 1972. ISBN 90 6196 074 6.

MCT 40 H.C. TIJMS, *Analysis of (s,S) inventory models*, 1972. ISBN 90 6196 075 4.

MCT 41 A. VERBEEK, *Superextensions of topological spaces*, 1972. ISBN 90 6196 076 2.

MCT 42 W. VERVAAT, *Success epochs in Bernoulli trials (with applications in number theory)*, 1972. ISBN 90 6196 077 0.

MCT 43 F.H. RUYMGAART, *Asymptotic theory of rank tests for independence*, 1973. ISBN 90 6196 081 9.

MCT 44 H. BART, *Meromorphic operator valued functions*, 1973. ISBN 90 6196 082 7.

MCT 45 A.A. BALKEMA, *Monotone transformations and limit laws* 1973. ISBN 90 6196 083 5.

MCT 46 R.P. VAN DE RIET, *ABC ALGOL, A portable language for formula manipulation systems, part 1: The language*, 1973. ISBN 90 6196 084 3.

MCT 47 R.P. VAN DE RIET, *ABC ALGOL, A portable language for formula manipulation systems, part 2: The compiler*, 1973. ISBN 90 6196 085 1.

MCT 48 F.E.J. KRUSEMAN ARETZ, P.J.W. TEN HAGEN & H.L. OUDSHOORN, *An ALGOL 60 compiler in ALGOL 60, Text of the MC-compiler for the EL-X8*, 1973. ISBN 90 6196 086 X.

MCT 49 H. KOK, *Connected orderable spaces*, 1974. ISBN 90 6196 088 6.

MCT 50 A. VAN WIJNGAARDEN, B.J. MAILLOUX, J.E.L. PECK, C.H.A. KOSTER, M. SINTZOFF, C.H. LINDSEY, L.G.L.T. MEERTENS & R.G. FISKER (eds), *Revised report on the algorithmic language ALGOL 68*, 1976. ISBN 90 6196 089 4.

MCT 51 A. HORDIJK, *Dynamic programming and Markov potential theory*, 1974. ISBN 90 6196 095 9.

MCT 52 P.C. BAAYEN (ed.), *Topological structures*, 1974. ISBN 90 6196 096 7.

MCT 53 M.J. FABER, *Metrizability in generalized ordered spaces*, 1974. ISBN 90 6196 097 5.

MCT 54 H.A. LAUWERIER, *Asymptotic analysis, part 1*, 1974. ISBN 90 6196 098 3.

MCT 55 M. HALL JR. & J.H. VAN LINT (eds), *Combinatorics, part 1: Theory of designs, finite geometry and coding theory*, 1974. ISBN 90 6196 099 1.

MCT 56 M. HALL JR. & J.H. VAN LINT (eds), *Combinatorics, part 2: Graph theory, foundations, partitions and combinatorial geometry*, 1974. ISBN 90 6196 100 9.

MCT 57 M. HALL JR. & J.H. VAN LINT (eds), *Combinatorics, part 3: Combinatorial group theory*, 1974. ISBN 90 6196 101 7.

MCT 58 W. ALBERS, *Asymptotic expansions and the deficiency concept in statistics*, 1975. ISBN 90 6196 102 5.

MCT 59 J.L. MIJNHEER, *Sample path properties of stable processes*, 1975. ISBN 90 6196 107 6.

MCT 60 F. GÖBEL, *Queueing models involving buffers*, 1975. ISBN 90 6196 108 4.

*MCT 61 P. VAN EMDE BOAS, *Abstract resource-bound classes, part 1*, ISBN 90 6196 109 2.

*MCT 62 P. VAN EMDE BOAS, *Abstract resource-bound classes, part 2*, ISBN 90 6196 110 6.

MCT 63 J.W. DE BAKKER (ed.), *Foundations of computer science*, 1975. ISBN 90 6196 111 4.

MCT 64 W.J. DE SCHIPPER, *Symmetric closed categories*, 1975. ISBN 90 6196 112 2.

MCT 65 J. DE VRIES, *Topological transformation groups 1 A categorical approach*, 1975. ISBN 90 6196 113 0.

MCT 66 H.G.J. PIJLS, *Locally convex algebras in spectral theory and eigenfunction expansions*, 1976. ISBN 90 6196 114 9.

*MCT 67 H.A. LAUWERIER, *Asymptotic analysis, part 2*, ISBN 90 6196 119 X.

MCT 68 P.P.N. DE GROEN, *Singularly perturbed differential operators of second order*, 1976. ISBN 90 6196 120 3.

MCT 69 J.K. LENSTRA, *Sequencing by enumerative methods*, 1977. ISBN 90 6196 125 4.

MCT 70 W.P. DE ROEVER JR., *Recursive program schemes: Semantics and proof theory*, 1976. ISBN 90 6196 127 0.

MCT 71 J.A.E.E. VAN NUNEN, *Contracting Markov decision processes*, 1976. ISBN 90 6196 129 7.

MCT 72 J.K.M. JANSEN, *Simple periodic and nonperiodic Lamé functions and their applications in the theory of conical waveguides*, 1977. ISBN 90 6196 130 0.

MCT 73 D.M.R. LEIVANT, *Absoluteness of intuitionistic logic*, 1979. ISBN 90 6196 122 X.

MCT 74 H.J.J. TE RIELE, *A theoretical and computational study of generalized aliquot sequences*, 1976. ISBN 90 6196 131 9.

MCT 75 A.E. BROUWER, *Treelike spaces and related connected topological spaces*, 1977. ISBN 90 6196 132 7.

MCT 76 M. REM, *Associations and the closure statement*, 1976. ISBN 90 6196 135 1.

MCT 77 W.C.M. KALLENBERG, *Asymptotic optimality of likelihood ratio tests in exponential families*, 1977. ISBN 90 6196 134 3.

MCT 78 E. DE JONGE & A.C.M. VAN ROOIJ, *Introduction to Riesz spaces*, 1977. ISBN 90 6196 133 5.

MCT 79 M.C.A. VAN ZUIJLEN, *Empirical distributions and rank statistics*, 1977. ISBN 90 6196 145 9.

MCT 80 P.W. HEMKER, *A numerical study of stiff two-point boundary problems*, 1977. ISBN 90 6196 146 7.

MCT 81 K.R. APT & J.W. DE BAKKER (eds), *Foundations of computer science II*, part 1, 1976. ISBN 90 6196 140 8.

MCT 82 K.R. APT & J.W. DE BAKKER (eds), *Foundations of computer science II*, part 2, 1976. ISBN 90 6196 141 6.

MCT 83 L.S. BENTHEM JUTTING, *Checking Landau's "Grundlagen" in the AUTOMATH system*, 1979. ISBN 90 6196 147 5.

MCT 84 H.L.L. BUSARD, *The translation of the elements of Euclid from the Arabic into Latin by Hermann of Carinthia (?) books vii-xii*, 1977. ISBN 90 6196 148 3.

MCT 85 J. VAN MILL, *Supercompactness and Wallman spaces*, 1977. ISBN 90 6196 151 3.

MCT 86 S.G. VAN DER MEULEN & M. VELDHORST, *Torrix I, A programming system for operations on vectors and matrices over arbitrary fields and of variable size*. 1978. ISBN 90 6196 152 1.

*MCT 87 S.G. VAN DER MEULEN & M. VELDHORST, *Torrix II*, ISBN 90 6196 153 X.

MCT 88 A. SCHRIJVER, *Matroids and linking systems*, 1977. ISBN 90 6196 154 8.

MCT 89 J.W. DE ROEVER, *Complex Fourier transformation and analytic functionals with unbounded carriers*, 1978. ISBN 90 6196 155 6.

MCT 90 L.P.J. GROENEWEGEN, *Characterization of optimal strategies in dynamic games*, 1981 . ISBN 90 6196 156 4.

MCT 91 J.M. GEYSEL, *Transcendence in fields of positive characteristic*, 1979. ISBN 90 6196 157 2.

MCT 92 P.J. WEEDA, *Finite generalized Markov programming*, 1979. ISBN 90 6196 158 0.

MCT 93 H.C. TIJMS & J. WESSELS (eds), *Markov decision theory*, 1977. ISBN 90 6196 160 2.

MCT 94 A. BIJLSMA, *Simultaneous approximations in transcendental number theory*, 1978. ISBN 90 6196 162 9.

MCT 95 K.M. VAN HEE, *Bayesian control of Markov chains*, 1978. ISBN 90 6196 163 7.

MCT 96 P.M.B. VITÁNYI, *Lindenmayer systems: Structure, languages, and growth functions*, 1980. ISBN 90 6196 164 5.

*MCT 97 A. FEDERGRUEN, *Markovian control problems; functional equations and algorithms*, . ISBN 90 6196 165 3.

MCT 98 R. GEEL, *Singular perturbations of hyperbolic type*, 1978. ISBN 90 6196 166 1.

MCT 99 J.K. LENSTRA, A.H.G. RINNOOY KAN & P. VAN EMDE BOAS, *Interfaces between computer science and operations research*, 1978. ISBN 90 6196 170 X.

MCT 100 P.C. BAAYEN, D. VAN DULST & J. OOSTERHOFF (eds), *Proceedings bicentennial congress of the Wiskundig Genootschap, part 1*, 1979. ISBN 90 6196 168 8.

MCT 101 P.C. BAAYEN, D. VAN DULST & J. OOSTERHOFF (eds), *Proceedings bicentennial congress of the Wiskundig Genootschap, part 2*, 1979. ISBN 90 6196 169 6.

MCT 102 D. VAN DULST, *Reflexive and superreflexive Banach spaces*, 1978. ISBN 90 6196 171 8.

MCT 103 K. VAN HARN, *Classifying infinitely divisible distributions by functional equations*, 1978. ISBN 90 6196 172 6.

MCT 104 J.M. VAN WOUWE, *Go-spaces and generalizations of metrizability*, 1979. ISBN 90 6196 173 4.

MCT 105 R. HELMERS, *Edgeworth expansions for linear combinations of order statistics*, 1982. ISBN 90 6196 174 2.

MCT 106 A. SCHRIJVER (ed.), *Packing and covering in combinatorics*, 1979. ISBN 90 6196 180 7.

MCT 107 C. DEN HEIJER, *The numerical solution of nonlinear operator equations by imbedding methods*, 1979. ISBN 90 6196 175 0.

MCT 108 J.W. DE BAKKER & J. VAN LEEUWEN (eds), *Foundations of computer science III*, part *1*, 1979. ISBN 90 6196 176 9.

MCT 109 J.W. DE BAKKER & J. VAN LEEUWEN (eds), *Foundations of computer science III*, part *2*, 1979. ISBN 90 6196 177 7.

MCT 110 J.C. VAN VLIET, *ALGOL 68 transput*, part *I*: *Historical review and discussion of the implementation model*, 1979. ISBN 90 6196 178 5.

MCT 111 J.C. VAN VLIET, *ALGOL 68 transput*, part *II*: *An implementation model*, 1979. ISBN 90 6196 179 3.

MCT 112 H.C.P. BERBEE, *Random walks with stationary increments and renewal theory*, 1979. ISBN 90 6196 182 3.

MCT 113 T.A.B. SNIJDERS, *Asymptotic optimality theory for testing problems with restricted alternatives*, 1979. ISBN 90 6196 183 1.

MCT 114 A.J.E.M. JANSSEN, *Application of the Wigner distribution to harmonic analysis of generalized stochastic processes*, 1979. ISBN 90 6196 184 X.

MCT 115 P.C. BAAYEN & J. VAN MILL (eds), *Topological Structures II*, part *1*, 1979. ISBN 90 6196 185 5.

MCT 116 P.C. BAAYEN & J. VAN MILL (eds), *Topological Structures II*, part *2*, 1979. ISBN 90 6196 186 6.

MCT 117 P.J.M. KALLENBERG, *Branching processes with continuous state space*, 1979. ISBN 90 6196 188 2.

MCT 118 P. GROENEBOOM, *Large deviations and asymptotic efficiencies*, 1980. ISBN 90 6196 190 4.

MCT 119 F.J. PETERS, *Sparse matrices and substructures, with a novel implementation of finite element algorithms*, 1980. ISBN 90 6196 192 0.

MCT 120 W.P.M. DE RUYTER, *On the asymptotic analysis of large-scale ocean circulation*, 1980. ISBN 90 6196 192 9.

MCT 121 W.H. HAEMERS, *Eigenvalue techniques in design and graph theory*, 1980. ISBN 90 6196 194 7.

MCT 122 J.C.P. BUS, *Numerical solution of systems of nonlinear equations*, 1980. ISBN 90 6196 195 5.

MCT 123 I. YUHÁSZ, *Cardinal functions in topology - ten years later*, 1980. ISBN 90 6196 196 3.

MCT 124 R.D. GILL, *Censoring and stochastic integrals*, 1980. ISBN 90 6196 197 1.

MCT 125 R. EISING, *2-D systems, an algebraic approach*, 1980. ISBN 90 6196 198 X.

MCT 126 G. VAN DER HOEK, *Reduction methods in nonlinear programming*, 1980. ISBN 90 6196 199 8.

MCT 127 J.W. KLOP, *Combinatory reduction systems*, 1980. ISBN 90 6196 200 5.

MCT 128 A.J.J. TALMAN, *Variable dimension fixed point algorithms and triangulations*, 1980. ISBN 90 6196 201 3.

MCT 129 G. VAN DER LAAN, *Simplicial fixed point algorithms*, 1980. ISBN 90 6196 202 1.

MCT 130 P.J.W. TEN HAGEN et al., *ILP Intermediate language for pictures*, 1980. ISBN 90 6196 204 8.

MCT 131 R.J.R. BACK, *Correctness preserving program refinements: Proof theory and applications*, 1980. ISBN 90 6196 207 2.

MCT 132 H.M. MULDER, *The interval function of a graph*, 1980. ISBN 90 6196 208 0.

MCT 133 C.A.J. KLAASSEN, *Statistical performance of location estimators*, 1981. ISBN 90 6196 209 9.

MCT 134 J.C. VAN VLIET & H. WUPPER (eds), *Proceedings international conference on ALGOL 68*, 1981. ISBN 90 6196 210 2.

MCT 135 J.A.G. GROENENDIJK, T.M.V. JANSSEN & M.J.B. STOKHOF (eds), *Formal methods in the study of language*, part *I*, 1981. ISBN 90 6196 211 0.

MCT 136 J.A.G. GROENENDIJK, T.M.V. JANSSEN & M.J.B. STOKHOF (eds), *Formal methods in the study of language*, part *II*, 1981. ISBN 90 6196 213 7.

MCT 137 J. TELGEN, *Redundancy and linear programs*, 1981. ISBN 90 6196 215 3.

MCT 138 H.A. LAUWERIER, *Mathematical models of epidemics*, 1981. ISBN 90 6196 216 1.

MCT 139 J. VAN DER WAL, *Stochastic dynamic programming, successive approximations and nearly optimal strategies for Markov decision processes and Markov games*, 1980. ISBN 90 6196 218 8.

MCT 140 J.H. VAN GELDROP, *A mathematical theory of pure exchange economies without the no-critical-point hypothesis*, 1981.
ISBN 90 6196 219 6.

MCT 141 G.E. WELTERS, *Abel-Jacobi isogenies for certain types of Fano three-folds*, 1981.
ISBN 90 6196 227 7.

MCT 142 H.R. BENNETT & D.J. LUTZER (eds), *Topology and order structures*, part 1, 1981.
ISBN 90 6196 228 5.

MCT 143 H. J.M. SCHUMACHER, *Dynamic feedback in finite- and infinite dimensional linear systems*, 1981.
ISBN 90 6196 229 3.

MCT 144 P. EIJGENRAAM, *The solution of initial value problems using interval arithmetic. Formulation and analysis of an algorithm*, 1981.
ISBN 90 6196 230 7.

MCT 145 A.J. BRENTJES, *Multi-dimensional continued fraction algorithms*, 1981. ISBN 90 6196 231 5.

MCT 146 C. VAN DER MEE, *Semigroup and factorization methods in transport theory*, 1982. ISBN 90 6196 233 1.

MCT 147 H.H. TIGELAAR, *Identification and informative sample size*, 1982.
ISBN 90 6196 235 8.

MCT 148 L.C.M. KALLENBERG, *Linear programming and finite Markovian control problems*, 1983. ISBN 90 6196 236 6.

MCT 149 C.B. HUIJSMANS, M.A. KAASHOEK, W.A.J. LUXEMBURG & W.K. VIETSCH, (eds), *From A to Z, proceeding of a symposium in honour of A.C. Zaanen*, 1982. ISBN 90 6196 241 2.

MCT 150 M. VELDHORST, *An analysis of sparse matrix storage schemes*, 1982.
ISBN 90 6196 242 0.

MCT 151 R.J.M.M. DOES, *Higher order asymptotics for simple linear Rank statistics*, 1982. ISBN 90 6196 243 9.

MCT 152 G.F. VAN DER HOEVEN, *Projections of Lawless sequences*, 1982.
ISBN 90 6196 244 7.

MCT 153 J.P.C. BLANC, *Application of the theory of boundary value problems in the analysis of a queueing model with paired services*, 1982.
ISBN 90 6196 247 1.

MCT 154 H.W. LENSTRA, JR. & R. TIJDEMAN (eds), *Computational methods in number theory, part I*, 1982.
ISBN 90 6196 248 X.

MCT 155 H.W. LENSTRA, JR. & R. TIJDEMAN (eds), *Computational methods in number theory, part II*, 1982.
ISBN 90 6196 249 8,

MCT 156 P.M.G. APERS, *Query processing and data allocation in distributed database systems*, 1983.
ISBN 90 6196 251 X.

MCT 157 H.A.W.M. KNEPPERS, *The covariant classification of two-dimensional smooth commutative formal groups over an algebraically closed field of positive characteristic*, 1983.
ISBN 90 6196 252 8.

MCT 158 J.W. DE BAKKER & J. VAN LEEUWEN (eds), *Foundations of computer science IV, Distributed systems*, part 1, 1983.
ISBN 90 6196 254 4.

MCT 159 J.W. DE BAKKER & J. VAN LEEUWEN (eds), *Foundations of computer science IV, Distributed systems*, part 2, 1983.
ISBN 90 6196 255 0.

MCT 160 A. REZUS, *Abstract automat*, 1983.
ISBN 90 6196 256 0.

MCT 161 G.F. HELMINCK, *Eisenstein series on the metaplectic group, An algebraic approach*, 1983.
ISBN 90 6196 257 9.

MCT 162 J.J. DIK, *Tests for preference*, 1983.
ISBN 90 6196 259 5

:) MCT 163 H. SCHIPPERS, *Multiple grid methods for equations of the second kind with applications in fluid mechanics*, 1983.
ISBN 90 6196 260 9.

MCT 164 F.A. VAN DER DUYN SCHOUTEN, *Markov decision processes with continuous time parameter*, 1983.
ISBN 90 6196 261 7.

An asterisk before the number means "to appear"