stichting
mathematisch
centrum

MC

AFDELING MATHEMATISCHE BESLISKUNDE (DEPARTMENT OF OPERATIONS RESEARCH) BC 21/79 JUNI

H.C. TIJMS

SEMI-MARKOV DECISION PROCESSES WITH DENUMERABLE STATE SPACE, UNBOUNDED ONE-STEP COSTS AND THE AVERAGE COST CRITERION

2e boerhaavestraat 49 amsterdam

Printed at the Mathematical Centre, 49, 2e Boerhaavestraat, Amsterdam.

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

AMS(MOS) subject classification scheme (1970): 90C45

Semi-Markov decision processes with denumerable state space, unbounded one-step costs and the average cost criterion *)

by

H.C. Tijms

ABSTRACT. This paper surveys optimality results for average cost denumerable state semi-Markov decision processes with compact metric action sets and unbounded one-step costs under the assumption that the underlying Markov decision chains associated with the stationary policies are unichained. Also an extensive discussion of simultaneous recurrence conditions on a compact metric set of denumerable stochastic matrices is given.

KEY WORDS & PHRASES: *semi-Markov decision processes, denumerable state space, average costs, recurrence conditions, optimality results.*

---

*) Based on lectures given in the Postdoctoral Seminar on "Markov Decision Theory" held at the University of Utrecht, February-June, 1979.

1. *Introduction*

We are concerned with a dynamic system which at decision epochs beginning with epoch 0 is observed to be in one of the states of a *denumerable* state space I and subsequently is controlled by choosing an action. For any state $i \varepsilon I$, the set $A(i)$ denotes the set of pure actions available in state i. If at any decision epoch the system is in state i and action $a \varepsilon A(i)$ is taken, then, regardless of the history of the system, the following happens:

(i) an immediate cost $c(i,a)$ is incurred

(ii) the time until the next decision epoch and the state at the next decision epoch are random with joint probability distribution function $Q(.,.|i,a)$.

For any $i \varepsilon I$ and $a \varepsilon I$, let

$$p_{ij}(a) = Q(\infty,j|i,a) \text{ for } j \varepsilon I \text{ and } \tau(i,a) = \sum_{j \varepsilon I} \int_0^\infty tQ(dt,j|i,a).$$

i.e. $p_{ij}(a)$ denotes the probability that the next state will be j and $\tau(i,a)$ denotes the unconditional mean time until the next decision epoch when action a is taken in state i. Observe that $\Sigma_{j \varepsilon I} p_{ij}(a) = 1$ for alle i,a. We make the following assumption.

ASSUMPTION 1.

(a) *For any* $i \varepsilon I$, *the set* $A(i)$ *is a compact metric set.*

(b) *For any* $i \varepsilon I$, *both* $c(i,a)$, $p_{ij}(a)$ *for any* $j \varepsilon I$ *and* $\tau(i,a)$ *are continuous on* $A(i)$.

(c) *There is a number* $\varepsilon > 0$ *such that* $\tau(i,a) \geq \varepsilon$ *for all* $i \varepsilon I$ *and* $a \varepsilon A(i)$.

We now introduce some familiar notions. For n = 0,1,.., denote by X_n and a_n the state and the action at the n^{th} decision epoch (the 0^{th} decision epoch is at epoch 0). A policy π for controlling the system is any measurable rule which for each n specifies which action to choose at the n^{th} decision epoch given the current state X_n and the sequence $(X_0,a_0,\ldots,X_{n-1},a_{n-1})$ of past states and actions where the actions chosen may be randomised. A policy π is called *memoryless* when the actions chosen are independent of the history of the system except for the present state.

Define R as the class of all stochastic matrices $P = (p_{ij})$, $i,j \in I$ such that for any $i \in I$ the elements of the i^{th} row of P can be represented by

$$(1.1) \qquad p_{ij} = \int_{A(i)} p_{ij}(a)\, \pi_i(da) \text{ for all } j \in I$$

for some probability distribution $\pi_i\{.\}$ on A(i). Then any memoryless policy π can be represented by some sequence $(P_1, P_2, ..)$ in R such that the i^{th} row of P_n gives the probability distribution of the state at the n^{th} decision epoch when the current state at the (n-1)st decision epoch is i and policy π is used. Define $F = X_{i \in I} A(i)$. Observe that, under assumption 1(a), F is a compact metric set in the product topology. For any $f \in F$, let P(f) be the stochastic matrix whose (i,j)th element is $p_{ij}(f(i))$, $i,j \in I$ and for n = 1,2,... denote by the stochastic matrix $P^n(f) = (p^n_{ij}(f))$ the n-fold matrix product of P(f) with itself. A memoryless policy $\pi = (P_1, P_2, ...)$ is called randomized stationary when $P_n = P \in R$ for all $n \geq 1$ and is called stationary when in addition P = P(f) for some $f \in F$. A stationary policy which prescribes to take the single action $f(i) \in A(i)$ whenever the system is in state i will be denoted by $f^{(\infty)}$. Observe that under any randomized stationary policy with associated stochastic matrix $P \in R$ the process $\{X_n, n \geq 0\}$ is a Markov chain with one-step transition matrix P.

For n = 0,1,..., denote by τ_n the time between the n^{th} and (n+1)st decision epoch. A policy π^* is said to be *(strongly) average cost optimal* when $\limsup_{n\to\infty} \phi_n(i,\pi^*)$ is less than or equal to $\limsup_{n\to\infty} \phi_n(i,\pi)$ ($\liminf_{n\to\infty} \phi_n(i,\pi)$) for any $i \in I$ and policy π where $\phi_n(i,\pi)$ is defined by

$$(1.2) \qquad \phi_n(i,\pi) = \frac{E_\pi\{\Sigma_{k=0}^n c(X_k, a_k) \mid X_0 = i\}}{E_\pi\{\Sigma_{k=0}^n \tau_k \mid X_0 = i\}}, \; n = 0,1,...$$

with E_π is the expectation under policy π. We here assume that this quantity is well-defined for any $i \in I$ and policy π as is the case under the assumptions to be given in section 3.

For the case where A(i) is finite for all i and the quantities c(i,a) and $\tau(i,a)$ are uniformly bounded counterexamples can be given showing that an average cost optimal policy may not exist or no randomized stationary policy may exist whose average cost is within ε of the minimal average cost, cf. Ross (1970,1971). A counterexample in Fisher & Ross (1968) shows that an average cost optimal policy may exist but any average cost optimal policy is non-stationary. In this counterexample it is remarkable that for any randomized stationary policy the associated stochastic matrix $P \in R$ is irreducible and positive recurrent*). This indicates that strong recurrency conditions will be required to establish optimality results for the average cost criterion.

In general we can only say that for fixed initial state we may restrict ourselves to the class of memoryless policies. More precisely, by a slight modification of the proof of Theorem 2 in Derman & Strauch (1966), we have the well-known result that for any fixed $i_0 \in I$ and policy π_0 a memoryless policy π_M can be found such that for any $k \in I$, Borel set $A \subseteq A(k)$ and $n \geq 0$

$$(1.3) \qquad Pr_{\pi_m}\{X_n=k,\ a_n \in A \mid X_0=i_0\} = Pr_{\pi_0}\{X_n=k,\ a_n \in A \mid X_0=i_0\}.$$

We further state as general result that if a finite solution $\{g, v(i) \mid i \in I\}$ exists to the average cost optimality equation

$$(1.4) \qquad v(i) = \min_{a \in A(i)} \{c(i,a) - g\tau(i,a) + \sum_{j \in I} p_{ij}(a)v(j)\} \text{ for all } i \in I$$

such that

$$(1.5) \qquad \lim_{n \to \infty} \frac{1}{n} E_\pi\{|v(X_n)| \mid X_0=i\} = 0 \text{ for any } i \in I \text{ and policy } \pi$$

then any stationary policy $f^{(\infty)}$ such that the pure action $f(i)$ minimizes the right side of (1.4) for all $i \in I$ is strongly average cost optimal.

We shall focus our attention on the existence of a finite solution to the average cost optimality equation and the existence of a strongly average cost optimal policy. In this paper we shall present for the unichained case a rather complete theory for the denumerable state semi-Markov model with unbounded one-step expected costs and the average costs as optimality criterion. For the unichained case this theory encompasses the finite state space model and the special cases of the denumerable state space model studied so far in the literature.

---

*) cf. also Fisher (1968) for the deep result that $\inf_{f \in F} \pi_j(f) > 0$ for all $j \in I$ provided that $A(i)$ is finite for all i and for any $f \in F$ the stochastic matrix $P(f)$ is irreducible and positive recurrent where $\{\pi_j(f), j \in I\}$ is the unique stationary probability distribution of $P(f)$. In case transient states are allowed and $\mu_{i0}(f) < \infty$ for all $i \in I$ and $f \in F$ for some common recurrent state 0 where $\mu_{i0}(f)$ denotes the mean number of transitions until the first return starting from i and using $f^{(\infty)}$, it is an open question whether $\inf_{f \in F} \pi_0(f) > 0$.

This literature was initiated by the papers of Derman (1963), Taylor (1965) and Ross (1968) under the very restrictive assumption of uniformly bounded functions $c(i,a)$ and $\tau(i,a)$ and under the assumption of a common positive recurrent state for the stationary policies. This work was considerably extended in Hordijk (1974,1976) and Federguen & Tijms (1978) and in a recent paper of Federgruen, Hordijk and Tijms (1978b) a rather complete theory for the unichained case has been established. Extensions of this theory for the unichained case to the important case of "communicating" Markov decision chains (cf. Bather (1973) and Hordijk (1974)) will probably involve the other assumptions given in section 3 below but will require quite different proof techniques as possibly linear programming or fixed point methods.

To establish this results for the average cost criterion, a thorough analysis of recurrence properties of the collection of underlying stochastic matrices $P(f)$, $f \in F$ will be essential. This analysis will be presented in detail in section 2. In the final section 3 we outline how to derive the optimality results for the average cost criterion.

2. Recurrence conditions for a collection of denumerable stochastic matrices.

We consider a set $\mathcal{P} = (P(f), f \in F)$ of stochastic matrices $P(f) = (p_{ij}(f))$, $i,j \in I$ having a denumerable state space I where the parameter set F is a compact metric space. Note that, for any $f \in F$, $p_{ij}(f) \geq 0$ and $\Sigma_{j \in I} p_{ij}(f) = 1$. Throughout this section we make the following assumption.

Assumption (a) *For any* $i,j \in I$, *the function* $p_{ij}(f)$ *is continuous on the compact metric set* F.

(b) *For any* $f \in F$, *the stochastic matrix* $P(f)$ *has no two disjoint closed sets.*

For any $f \in F$, denote by the stochastic matrix $P^n(f) = (p^n_{ij}(f), i,j \in I$ the n-fold matrix product of $P(f)$ with itself for $n = 1,2,\ldots$. Note that for any $i,j \in I$ and $n \geq 1$ the function $p^n_{ij}(f)$ is continuous on F. For any $i_0 \in I$, $A \subset I$ and $f \in F$, define the taboo probability

$$(2.1) \quad t^n_{i_0 A}(f) = \sum_{i_1,\ldots,i_n \in I \setminus A} p_{i_0 i_1}(f) \ldots p_{i_{n-1} i_n}(f), \quad n=1,2,\ldots$$

i.e. $t^n_{iA}(f)$ is the probability that under the stochastic matrix $P(f)$ the first return to the set A takes more than n transitions starting from state i.

For any $i \in I$, $A \subset I$ and $f \in F$, define the (possibly infinite) mean recurrence time

$$(2.2) \qquad \mu_{iA}(f) = 1 + \Sigma_{n=1}^{\infty} t_{iA}^{n}(f).$$

We write $t_{iA}^{n}(f) = t_{ij}(f)$ and $\mu_{iA}(f) = \mu_{ij}(f)$ for $A = \{j\}$. Consider now the following simultaneous recurrence conditions on the set $P = (P(f), f \in F)$.

C1. *There is a finite set* K *and a finite number* B *such that*

$$\mu_{iK}(f) \leq B \textit{ for all } i \in I \textit{ and } f \in F.$$

C2. *There is a finite set* K, *an integer* $\nu \geq 1$ *and a number* $\rho > 0$ *such that*

$$\sum_{j \in K} p_{ij}^{\nu}(f) \geq \rho \textit{ for all } i \in I \textit{ and } f \in F.$$

C3. *There is an integer* $\nu \geq 1$ *and a number* $\rho > 0$ *such that for all* $f \in F$

$$(2.3) \qquad \inf_{i_1, i_2 \in I} \{ \sum_{j \in I} \min[p_{i_1 j}^{\nu}(f), p_{i_2 j}^{\nu}(f)]\} \geq \rho$$

C4. *There is an integer* $\nu \geq 1$ *and a number* $\rho > 0$ *such that for any* $f \in F$ *a probability distribution* $\{\pi_j(f), j \in I\}$ *(say) exists for which*

$$\left| \sum_{j \in A} p_{ij}^{n}(f) - \sum_{j \in A} \pi_j(f) \right| \leq (1-\rho)^{\lfloor n/\nu \rfloor} \textit{ for all } i \in I,\ A \subset I \textit{ and } n \geq 1.$$

where $\lfloor x \rfloor$ *denotes the largest integer less than or equal to* x.

C5. *For any* $f \in F$ *there is a probability distribution* $\{\pi_j(f), j \in I\}$ *such that*

$$p_{ij}^{n}(f) \to \pi_j(f) \textit{ uniformly in } (i,f) \in I \times F \textit{ as } n \to \infty \textit{ for any } j \in I.$$

C6. *There is a finite number* B *such that for any* $f \in F$ *a state* s_f *exists for which*

$$\mu_{is_f}(f) \leq B \textit{ for all } i \in I.$$

C7. *There is a finite set* K *and a finite number* B *such that for any* $f \in F$ *a state* $s_f \in K$ *exists for which*

$$\mu_{is_f}(f) \leq B \textit{ for all } i \in I.$$

C8. *There is an integer* $\nu \geq 1$ *and a number* $\rho > 0$ *such that for any* $f \in F$ *a state* s_f *exists for which*

$$p_{is_f}^{\nu}(f) \geq \rho \textit{ for all } i \in I.$$

C9. *There is a finite set* K*, an integer* $\nu \geq 1$ *and a number* $\rho > 0$ *such that for any* $f \in F$ *a state* $s_f \in K$ *exists for which*

$$p^{\nu}_{is_f}(f) \geq \rho \text{ for all } i \in I.$$

C10. *There exists an integer* $\nu \geq 1$*, a number* $\rho > 0$ *such that for all* $i_1, i_2 \in I$ *and* $f \in F$ *a state* j *exists with*

$$\min\{\Sigma^{\nu}_{k=1} p^{k}_{i_1 j}(f),\ \Sigma^{\nu}_{k=1} p^{k}_{i_2 j}(f)\} \geq \rho.$$

The simultaneous recurrence conditions C1 and C2 were first considered in Hordijk (1974). The condition C2 was called the simultaneous Doeblin condition since under C2 for any $f \in F$ the stochastic matrix P(f) satisfies the so-called Doeblin condition from Markov chain theory. The simultaneous recurrence conditions C3 and C4 were first studied in Federgruen & Tijms (1978). In analogy to Markov chain terminology the conditions C3 and C4 are called a simultaneous scrambling and quasi-compactness condition respectively. We not that the right side of (2.3) is called the ergodic coefficient of the stochastic matrix P(f) and the probability distribution $\{\pi_j(f), j \in I\}$ in C4 is the unique stationary probability distribution of P(f). Also note that under C3, C4, C5 and C8 for any $f \in F$ the stochastic matrix P(f) is aperiodic. Finally, the condition C10 is a generalization of C4 and allows for periodicity.

We shall now present a number of relationships between the above simultaneous recurrence conditions. The discussion is based on Hordijk (1974), Federgruen & Tijms (1978) and Federgruen, Hordijk & Tijms (1978a).

THEOREM 2.1.

(i) *Suppose that for any* $f \in F$ *the stochastic matrix* P(f) *has a unique stationary probability distribution* $\{\pi_j(f), j \in I\}$ *(say). Then for any* $j \in I$ *the function* $\pi_j(f)$ *is continuous on* F *if and only if the collection* $[\{\pi_j(f), j \in I\}]$ *is a tight collection of probability distribution, i.e. for any* $\varepsilon > 0$ *there is a set* $K(\varepsilon)$ *such that*

$$(2.4) \qquad \sum_{j \in K(\varepsilon)} \pi_j(f) \geq 1 - \varepsilon \text{ for all } f \in F.$$

(ii) *The conditions* C1 *and* C2 *are equivalent. Further, condition* C2 *implies the tightness condition* (2.4)*)

*) The condition C2 is much stronger than the tightness condition (2.4). It would be interesting to know other sufficient conditions for (2.4). It is our conjecture that (2.4) is satisfied when a finite set K exists such that sup μ (f) < ∞

THEOREM 2.2.

(i) *If the stochastic matrix* P(f) *is aperiodic for any* $f\in F$, *then condition* C2 *implies condition* C3.

(ii) *The condition* C3 *implies condition* C4.

THEOREM 2.3.

(i) *The condition* C5 *implies the conditions* C2 *and* C9.

(ii) *The conditions* C3, C4, C5, C8 *and* C9 *are equivalent.*

THEOREM 2.4.

(i) *Each of the conditions* C2, C6 *and* C10 *implies condition* C7.

Summarizing the above theorems we have

THEOREM 2.5.

(i) *The conditions* C1, C2, C6, C7 *and* C10 *are equivalent.**)

(ii) *If the stochastic matrix* P(f) *is aperiodic for any* $f\in F$, *then the conditions* C1-C10 *are equivalent.*

We now give the proofs of these Theorems.

Proof of Theorem 2.1.

(i) Suppose first that for each $\varepsilon>0$ we can find a finite set $K(\varepsilon)$ such that (2.4) holds. Now, let $\{f_n, n\geq 1\}$ be any sequence in F such that $f_n \to f^*$ as $n\to\infty$. Choose $h\in I$. We shall now verify that

$$\lim_{n\to\infty} \pi_h(f_n) = \pi_h(f^*).$$

To do this, let α_h be any limit point of $\{\pi_h(f_n), n\geq 1\}$. By the well-known diagonalization method, we can choose a subsequence $\{n_k, k\geq 1\}$ of integers for which

$$\lim_{k\to\infty} \pi_j(f_{n_k}) = \pi_j \text{ (say) exists for all } j\in I \text{ such that } \pi_h=\alpha_h.$$

*) It appears from the proof of Theorem 2.5(i) that it is no restriction to require in C10 that the states $j = j(i_1,i_2,f)$ belong to a finite set.

For any $f \in F$ we have that $\{\pi_j(f)\}$ is the unique probability distribution satisfying

$$\pi_j(f) = \sum_{i \in I} p_{ij}(f)\, \pi_i(f) \quad \text{for all } j \in I.$$

Taking $f=f_{n_k}$, letting $k\to\infty$ and using both the continuity of $p_{ij}(f)$ and Proposition 18 on p. 232 in Royden (1968), we find

$$\pi_j = \sum_{i \in I} p_{ij}(f^*)\pi_i \quad \text{for all } j \in I.$$

Further, using (2.4) we have

$$\sum_{j \in I} \pi_j = 1$$

Together the above relations and the fact that $P(f^*)$ has a unique stationary probability distribution imply $\pi_j = \pi_j(f^*)$ for all j. In particular $\alpha_h = \pi_h(f^*)$ as was to be proved.

Suppose next that $\pi_j(f)$ is continuous on F for each $j \in I$. Let now $\{K_n, n\geq 1\}$ be any sequence of finite subsets of I such that

$$K_{n+1} \supseteq K_n \text{ for all } n\geq 1 \text{ and } \lim_{n\to\infty} K_n = I.$$

Define for n=1,2,...

$$a_n(f) = \sum_{j \in K_n} \pi_j(f), \quad f \in F.$$

Then $a_n(f)$ is continuous on F for all $n\geq 1$. Further, we have for any $f \in F$ that

$$a_{n+1}(f) \geq a_n(f) \text{ for all } n\geq 1 \text{ and } \lim_{n\to\infty} a_n(f) = 1.$$

Now, by Theorem 7.13 in Rudin (1964), it follows that $a_n(f)$ converges to 1 uniformly in $f \in F$ as $n\to\infty$. Hence for each $\varepsilon>0$ we can find a finite n such that $a_n(f) \geq 1-\varepsilon$ for any f F which verifies (2.4).

(ii) We shall prove that condition C2 with triple (K,ν,ρ) implies that for some finite constant B

$$\mu_{iK}(f) \leq B \quad \text{for all } i \in I \text{ and } f \in F$$

and that this result implies in its turn that for any $\varepsilon>0$ a finite set $K(\varepsilon)$ and an integer $\nu(\varepsilon)$ exist such that

$$\sum_{j\in K(\varepsilon)} p_{ij}^{n}(f) \geq 1-\varepsilon \text{ for all } i\in I \text{ and } n\geq\nu(\varepsilon).$$

The latter inequality has a consequence that condition C2 implies the tightness relation (2.4), since it is well-known from Markov chain theory (cf. Doob (1953)) that a stochastic matrix $P(f)$ which is unichained and satisfies the Doeblin condition has a unique stationary probability distribution $\{\pi_j(f), j\in I\}$ such that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} p_{ij}^{k}(f) = \pi_j(f) \text{ for all } i,j\in I.$$

Suppose now that C2 with triple (K,ν,ρ) holds. Then, $t_{iK}^{\nu}(f) \leq 1-\rho$ for all $i\in I$ and $f\in F$ and so, for all $m\geq 1$,

$$t_{iK}^{m}(f) \leq (1-\rho)^{[m/\nu]} \text{for all } i\in I \text{ and } f\in F,$$

using the fact that $t_{iK}^{m}(f)$ is non-increasing in $m\geq 1$. Now, by (2.2), we get $\mu_{iK}(f) \leq B$ for all $i\in I$ and $f\in F$ for some constant B. From this result on, we proceed as follows. Fix $0<\varepsilon<1$ and choose $0<\gamma<1$ such that $(1-\gamma)^2 \geq 1-\varepsilon$. Then we can find an integer $N\geq 1$ such that

(2.5) $\quad t_{iK}^{N}(f) \leq \gamma$ for all $i\in I$ and $f\in F$.

To prove this, suppose to the contrary that for any $m\geq 1$ a state $i\in I$ and a $f\in F$ can be found such that $t_{iK}^{m}(f)>\gamma$. Since this probability is non-increasing in m, we have by (2.2) $\mu_{iK}(f)>1+m\gamma$ which gives a contraction by choosing m large enough. We next show that there is a finite set A such that

(2.6) $\quad \sum_{j\in A} p_{ij}^{m}(f) \geq 1-\gamma$ for all $i\in K$, $1\leq m\leq N$ and $f\in F$.

To do this, fix $i\in K$ and $1\leq m\leq N$. In the same way as in the last part of the proof of (i), we find that for any $\gamma>0$ a finite set $A(\gamma)$ exists with $\Sigma_{j\in A(\gamma)} p_{ij}^{k}(f) \geq 1-\gamma$ for all $f\in F$. By this result and the finiteness of K, we get (2.6). Now, using (2.6)-(2.7) and defining $t_{iK}^{n}(k,f)$ as the probability that the first return to the set K occurs in state k at the nth transition, we find

$$\sum_{j\in A} p_{ij}^{N+1}(f) \geq \sum_{n=1}^{N} \sum_{k\in K} t_{iK}^{n}(k,f) \sum_{j\in A} p_{kj}^{N+1-n}(f) \geq (1-\gamma)(1-t_{iK}^{N}(f)) \geq (1-\gamma)^2 \geq 1-\varepsilon$$

which completes the proof.

Proof of Theorem 2.2.

(i) Suppose first that C2 with triple (K,ν,ρ) holds and that every $P(f)$ is aperiodic. We shall then verify condition C3. Since for any $f\in F$ the stochastic matrix $P(f)$ satisfies the Doeblin condition, has no two disjoint closed sets and is aperiodic, we have from Markov chain theory that every $P(f)$ has a unique stationary probability distribution $\{\pi_j(f), j\in I\}$ (say) such that

(2.7) $\lim_{n\to\infty} p_{ij}^{n}(f) = \pi_j(f)$ for all $i,j\in I$.

Since C2 implies $\Sigma_{j\in K} p_{ij}^{n}(f) \geq \rho$ for all $i\in I$, $f\in F$ and $n\geq\nu$, we have

(2.8) $\sum_{j\in K} \pi_j(f) \geq \rho$ for all $f\in F$.

Define now

(2.9) $F_k = \{f\in F \mid \pi_k(f) \geq \frac{\rho}{|K|}\}$ for $k\in K$.

where $|K|$ denotes the number of states in K. Then, by (2.8)

$$F = \bigcup_{k\in K} F_k.$$

Using part (ii) of th. 2.1 and the fact that F is a compact metric space, it follows that for any $k\in K$ the set F_k is closed and hence compact. For any $i\in I$ and $k\in K$, define

(2.10) $n(i,k,f) = \min\{n\geq 1 \mid p_{ik}^{n}(f) > \frac{\rho}{2|K|}\}$ for $f\in F_k$.

By (2.7), $n(i,k,f)$ exists and is finite. Using the fact that $P^n(f)$ is continuous on F for each $n\geq 1$, it is immediately verified that for each $i\in I$ and $k\in K$ the set $\{f\in F_k \mid n(i,k,f)\geq\alpha\}$ is closed for any real α, i.e. for each $i\in I$ and $k\in K$ the function $n(i,k,f)$ is upper semi-continuous on the compact set F_k.

Now, by Proposition 10 on p.161 in Royden (1968) we have that for each $i\in I$ and $k\in K$ the function $n(i,k,f)$ assumes a finite maximum on F_k. Hence, using the finiteness of K, we can find an integer $\mu\geq 1$ such that

(2.11) $n(i,k,f) \leq \mu$ for all $i\in K$, $k\in K$ and $f\in F_k$.

Next define for any $k\in K$

(2.12) $m(k,f) = \min\{n\geq 1 \mid p_{kk}^{m}(f) > \frac{\rho}{2|K|}$ for all $n\leq m\leq n+\mu\}$ for $f\in F_k$.

We now verify that for each $k\in K$ the set $S_\alpha = \{f\in F_k \mid m(k,f)\geq\alpha\}$ is closed for any real α. Fix $k\in K$ and an integer $\alpha>1$. Suppose that $f_n\in S_\alpha$ for $n\geq 1$ and that $f_n\to f^*$ as $n\to\infty$. Then we can find a subsequence $\{n_h, h\geq 1\}$ of integers, and integers r and t with $1\leq r\leq\alpha-1$ and $r\leq t\leq r+\mu$ such that $p_{kk}^{t}(f_{n_h}) \leq \rho/(2|K|)$ for all $h\geq 1$. Hence, by the fact that $p_{kk}^{t}(f)$ is continuous on F, we find $p_{kk}^{t}(f^*) \leq \rho/(2|K|)$ and so $f^* \in S_\alpha$. We have now proved that for any $k\in K$, the function $m(k,f)$ is upper semi-continuous on the compact set F_k. Hence there exists an integer $N\geq 1$ such that

$$m(k,f)<N \quad \text{for all } k\in K \text{ and } f\in F_k.$$

For any $k\in K$ and $f\in F_k$, we have by (2.10)-(2.12)

$$p_{ik}^{\mu+m(k,f)}(f)\geq p_{ik}^{n(i,k,f)}(f)p_{kk}^{m(k,f)+\mu-n(i,k,f)}(f) > \frac{\rho^2}{4|K|^2} \text{ for all } i\in K.$$

Hence for any $k\in K$ and $f\in F_k$,

$$p_{ik}^{\nu+\mu+m(k,f)}(f) \geq \sum_{j\in K} p_{ij}^{\nu}(f)p_{jk}^{\mu+m(k,f)}(f) > \frac{\rho^3}{4|K|^2} \text{ for all } i\in I.$$

Using this result, we now find for any $k\in K$ and $f\in F_k$,

$$\sum_{j\in I} \min[p_{i_1j}^{\nu+\mu+N}(f), p_{i_2j}^{\nu+\mu+N}(f)] \geq$$

$$\sum_{j\in I} \min[p_{i_1k}^{\nu+\mu+m(k,f)}(f)p_{kj}^{N-m(k,f)}(f), p_{i_2k}^{\nu+\mu+m(k,f)}(f)p_{kj}^{N-m(k,f)}(f)] \geq$$

$$\geq \frac{\rho^3}{4|K|^2} \sum_{j\in I} p_{kj}^{N-m(k,f)}(f) = \frac{\rho^3}{4|K|^2} \text{ for all } i_1,i_2\in I,$$

which verifies C3.

Proof of Theorem 2.2.

(ii) The proof of this theorem proceeds along the same lines as that of theorem 1 in Anthonisse & Tijms (1977).

Assume C3 holds with the pair (ν,ρ). Fix $f\in F$ and $A\subseteq I$. For $n=1,2,..$ define

$$M_n = \sup_{i\in I} \Sigma_{j\in A} p_{ij}^{n}(f) \text{ and } m_n = \inf_{i\in I} \Sigma_{j\in A} p_{ij}^{n}(f).$$

By induction, it follows from $p_{ij}^{n+1}(f) = \Sigma_{k\in I} p_{ik}(f)\, p_{kj}^{n}(f)$ that

(2.13) $\quad M_{n+1} \le M_n$ and $m_{n+1} \ge m_n$ for all $n \ge 1$.

For any number a, let a^+ and a^- be defined by $a^+ = \max(a,0)$ and $a^- = \min(a,0)$. Then, $a = a^+ + a^-$ and for any sequence $\{a_j, j\in I\}$ of numbers such that $\Sigma_{j\in I}|a_j| < \infty$ and $\Sigma_{j\in I} a_j = 0$ we have $\Sigma_j a_j^+ = -\Sigma_j a_j^-$. Further, we note that $(a-b)^+ = a - \min(a,b)$ for any pair of numbers a,b. Fix now $i\in I$ and $n > \nu$. Then,

$$\sum_{j\in A} p_{ij}^n(f) - \sum_{j\in A} p_{rj}^n(f) = \sum_{k\in I} \{p_{ik}^{\nu}(f) - p_{rk}^{\nu}(f)\} \sum_{j\in A} p_{kj}^{n-\nu}(f) =$$

$$= \sum_{k\in I} \{p_{ik}^{\nu}(f) - p_{rk}^{\nu}(f)\}^+ \sum_{j\in A} p_{kj}^{n-\nu}(f) + \sum_{k\in I} \{p_{ik}^{\nu}(f) - p_{rk}^{\nu}(f)\}^- \sum_{j\in A} p_{kj}^{n-\nu}(f)$$

$$\le \{M_{n-\nu} - m_{n-\nu}\} \sum_{k\in I} \{p_{ik}^{\nu}(f) - p_{rk}^{\nu}(f)\}^+ =$$

$$= \{M_{n-\nu} - m_{n-\nu}\}\{1 - \sum_{k\in I} \min[p_{ik}^{\nu}(f), p_{rk}^{\nu}(f)]\} \le$$

$$\le (1-\rho)(M_{n-\nu} - m_{n-\nu}).$$

Since i and r were arbitrarily chosen, it follows that

$$M_n - m_n \le (1-\rho)\, \{M_{n-\nu} - m_{n-\nu}\} \text{ for all } n \ge \nu.$$

Hence, since $M_n - m_n$ is non-increasing in $n \ge 1$,

(2.14) $\quad M_n - m_n \le (1-\rho)^{\lfloor n/\nu \rfloor}$ for all $n \ge 1$.

Together, (2.11) and (2.12) imply that for some finite non-negative number $\pi(A)$

$$\lim_{n\to\infty} M_n = \lim_{n\to\infty} m_n = \pi(A).$$

Further for any $n \ge 1$,

(2.15) $\quad m_n \le \pi(A) \le M_n$ and $m_n \le \sum_{j\in A} p_{ij}^n(f) \le M_n$ for all $i\in I$.

It now follows from (2.12) and (2.13) that

$$\Big|\sum_{j\in A} p_{ij}^n(f) - \pi(A)\Big| \le (1-\rho)^{\lfloor n/\nu \rfloor} \text{ for all } n \ge 1.$$

Since this relation holds for any $A \subseteq I$, it follows that $\pi(.)$ is a probability measure on the class of all subsets of I which completes the proof.

Proof of Theorem 2.3.

(i) Suppose that condition C5 holds. Since for any $i,j \in I$ and $n \geq 1$ the function $p_{ij}^n(f)$ is continuous on F, it follows from C5 that for any $j \in I$ the function $\pi_j(f)$ is continuous in $f \in F$. Now, by Theorem 2.1(i) we can find a finite set K and a number $\delta > 0$ such that

$$(2.16) \qquad \sum_{j \in K} \pi_j(f) \geq \delta \quad \text{for all } f \in F.$$

By C5 and the finiteness of K, we can find an integer $\nu \geq 1$ such that $p_{ij}^{\nu}(f) \geq \pi_j(f) - \delta/(2|K|)$ for all $i \in I$, $f \in F$ and $j \in K$ where $|K|$ denotes the number of states in K. Together this inequality and (2.14) imply condition C2. Further, we get from (2.14) that for any $f \in F$ there is a state s_f such that $\pi_{s_f}(f) \geq \delta/|K|$ and so $p_{is_f}^{\nu}(f) \geq \delta/(2|K|)$ for all $i \in I$ and $f \in F$. This inequality verifies condition C9 which completes the proof of part (i).

(ii) Since C9 implies C8 and in its turn C8 implies C3 and since C4 implies C5, this part follows by using theorem 2.2.(ii) and theorem 2.3(i).

Proof of Theorem 2.4.

To prove the theorem we shall use a classical perturbation of the stochastic matrices P(f), $f \in F$. Fix any number τ with $0<\tau<1$ and let $\bar{\mathcal{P}} = (\bar{P}(f), f \in F)$ be the set of stochastic matrices $\bar{P}(f) = (\bar{p}_{ij}(f))$, $i,j \in I$ such that for any $f \in F$ and $i,j \in I$:

$$\bar{p}_{ij} = \begin{cases} \tau p_{ij}(f) & \text{for } j \neq i \\ 1 - \tau + \tau p_{ii}(f) & \text{for } j = i \end{cases}$$

Note that, by $\bar{p}_{ii}(f) \geq 1 - \tau > 0$ for all $i \in I$ and $f \in F$, the stochastic matrix $\bar{P}(f)$ is aperiodic for all $f \in F$. Also note that for any $i,j \in I$ the function $\bar{p}_{ij}(f)$ is continuous in $f \in F$ and for any $f \in F$, the stochastic matrix $\bar{P}(f)$ has no two disjoint closed sets. Define for the stochastic matrices $\bar{P}(f)$ the taboo probabilities $\bar{t}_{iA}^n(f)$ and the mean recurrence times $\bar{\mu}_{iA}(f)$ as in (2.2) and (2.3).

By induction on n, it is straightforward to verify that for any $f \in F$

$$(2.17) \quad \bar{t}^{n}_{ij}(f) = \sum_{k=0}^{n} \binom{n}{k}(1-\tau)^{n-k}\tau^{k} t^{k}_{ij}(f) \text{ for all } n = 0,1,\ldots \text{ and } i,j \in I \text{ with } i \neq j,$$

where $\bar{t}^{0}_{ij}(f) = t^{0}_{ij}(f) = 1$. From the relations (2.3) and (2.17) we get

$$(2.18) \quad \bar{\mu}_{ij}(f) = \frac{\mu_{ij}(f)}{\tau} \quad \text{for all } i,j \in I \text{ with } i \neq j \text{ and } f \in F.$$

We note that this relation is intuitively clear by a direct probabilistic interpretation.

Suppose that the condition C2 holds with triple (K,ν,ρ). Then, by $\bar{p}_{ij}(f) \geq \tau p_{ij}(f)$ for all $i,j \in I$ and $f \in F$, we have

$$\sum_{j \in K} \bar{p}^{\nu}_{ij}(f) \geq \tau^{\nu} \sum_{j \in K} p_{ij}(f) \geq \tau^{\nu}\rho \quad \text{for all } i \in I \text{ and } f \in F.$$

Hence, the condition C2 applies to the set $\bar{P} = (\bar{P}(f), f \in F)$. Moreover we have that any stochastic matrix $\bar{P}(f)$, $f \in F$ is aperiodic. Now, by the combination of theorem 2.2 and part (ii) of theorem 2.3, it follows that condition C9 applies to the set $\bar{P}$. Now, by invoking (2.18), it follows that the condition C7 holds for the set $P = (P(f), f \in F)$ as was to be proved.

Suppose that condition C6 holds. Then, by invoking again (2.18), we have that condition C6 applies to the set $\bar{P}$. Hence there is a finite number B such that for any $f \in F$ there exists a state s_f such that

$$(2.19) \quad \bar{\mu}_{is_f}(f) = 1 + \sum_{n=1}^{\infty} \bar{t}^{n}_{is_f}(f) \leq B \text{ for all } i \in I.$$

Fix now $0<\gamma<1$. Since for any $f \in F$ and $i \in I$ the taboo probability $\bar{t}^{n}_{is_f}(f)$ is non-increasing in n, it follows that there is an integer $N \geq 1$ such that (cf. (2.5))

$$(2.20) \quad \bar{t}^{N}_{is_f}(f) \leq \gamma \quad \text{for all } i \in I \text{ and } f \in F.$$

Together the inquality (2.20) and the fact that $\bar{p}_{kk}(f) \geq 1-\tau$ for all $k \in I$ and $f \in F$ imply

$$\bar{p}^{N}_{is_f}(f) \geq (1-\tau)^{N-1}(1-\gamma) \quad \text{for all } i \in I \text{ and } f \in F.$$

This shows that condition C8 applies to the set $\bar{P}$. Next by part (ii) of theorem 2.3 condition C9 applies to the set $\bar{P}$. Since C9 implies C7, it follows that condition C7 applies to the set $\bar{P}$. Now by invoking again (2.18) we have that condition C7 holds for the stochastic matrices P(f), $f \in F$ as was to be verified.

Finally suppose that condition C10 holds. Then, by $\bar{p}_{ij}^{k}(f) \geq \tau^k p_{ij}^{k}(f)$ for all $i,j \in I$, $f \in F$ and $n \geq 1$, we have that C10 applies to the set $\bar{P}$ with ρ replaced by $\rho\tau^{\nu}$. Since C10 applies to $\bar{P}$, we can find for any $i_1, i_2 \in I$ and $f \in F$ a state j and integers $\ell_1, \ell_2 \leq \nu$ such that

$$\bar{p}_{i_1 j}^{\ell_1}(f) \geq \frac{\rho\tau^{\nu}}{\nu} \quad \text{and} \quad \bar{p}_{i_2 j}^{\ell_2}(f) \geq \frac{\rho\tau^{\nu}}{\nu}.$$

Then, by $\bar{p}_{ii}(f) \geq 1-\tau > 0$ for all $i \in I$ and $f \in F$, it follows that C3 applies to $\bar{P}$. Now, by part (ii) of Theorem 2.3 we have that C9 and consequently C7 apply to $\bar{P}$. Then, by (2.16), it follows that C7 applies to $\bar{P}$ which completes the proof.

<u>Proof of Theorem 2.5.</u>

(i) Since C7 trivially implies C1 and C6, we have by Theorem 2.1(ii) and Theorem 2.4 that C1, C2, C6 and C7 are equivalent. Since we have shown in Theorem 2.4 that C10 implies C7, it remains to verify that C7 implies C10. Now, by C7 and (2.2), it follows that for any $0<\gamma<1$ we can find an integer $\nu \geq 1$ such that for any $f \in F$ we have $t_{is_f}^{N} \leq \gamma$ for all $i \in I$. This implies C10.

(ii) This part is an immediate consequence of Theorem 2.2(i), Theorem 2.3(ii) and Theorem 2.5(i).

We conclude this section with a result due to Isaacson and Luecke (1978) which roughly states the under the scrambling condition C3 the "subdominant eigenvalues" of the stochastic matrices P(f), $f \in F$are uniformly below 1. Therefore let X be the Banach space of all bounded complex-valued functions on I with $\|x\| = \sup_{i \in I} |x(i)|$. For any matrix $A = (a_{ij})$, $i,j \in I$ such that A corresponds to a linear operator from X to X, define $\sigma(A)$ and $\sigma_p(A)$ as the spectrum and point spectrum of A respectively, i.e.

$$\sigma(A) = \mathbb{C} \setminus \{\lambda \in \mathbb{C} \mid A-\lambda I \text{ is 1-1 and on to}\}$$

and

$$\sigma_p(A) = \{\lambda \in \mathbb{C} \mid A-\lambda I \text{ is not 1-1}\} = \{\lambda \mid Ax = \lambda x \text{ for some } x \neq 0\}.$$

Let $r(A)$ be the spectral radius of A, i.e. $r(A) = \sup\{|\lambda| \mid \lambda\in\sigma(A)\}$. In case A is a stochastic matrix, define

$$\beta(A) = \sup\{|\lambda| \mid \lambda\in\sigma(A),\ \lambda\neq 1\}.$$

It is well-known that in case I is finite, then $\sigma(A) = \sigma_p(A)$. Moreover, in case A is a finite stochastic matrix which has no two disjoint closed sets and is aperiodic then the eigenvalue $\lambda=1$ has multiplicity one and any other eigenvalue of A is less than 1 in absolute value so that $\beta(A)$ represents the subdominant eigenvalue of A.

For any stochastic matrix $P = (p_{ij})$, $i,j\in I$, define

$$\delta(P) = 1 - \inf_{i_1,i_2\in I} \sum_{j\in I} \min(p_{i_1 j}, p_{i_2 j}),$$

i.e. the delta coefficient of P is one minus the ergodic coefficient of P. If the stochastic matrix P has a (unique) stationary probability distribution $\{\pi_j, j\in I\}$, then (see the proof of Theorem 3.5 in Isaacson & Luecke (1978)),

$$r(P-P^*) = \inf_{n\geq 1} [\delta(P^n)]^{1/n} = \lim_{n\to\infty} [\delta(P^n)]^{1/n},$$

where P^* is the stochastic matrix with identical rows given by $\{\pi_j, j\in I\}$. Moreover, by the theorems 3.5 and 3.8 in Isaacson & Luecke (1978) we have the next theorem.

THEOREM 2.6.
Suppose that C3 *holds with triple* (ν,ρ). *Then*

$$\beta(P(f)) = r(P(f)-P^*(f)) \leq (1-\rho)^{1/\nu} \textit{ for all } f\in F.$$

It is an open question whether the converse of the theorem holds, that is whether C3 holds when any $P(f)$, $f\in F$ is strongly ergodic and $\beta(P(f)-P^*(f))$ is is uniformly below 1.

3. Optimality results for the average cost criterion

We first need some notation. For any set $A \subset I$, define the random variable

$$N(A) = \inf \{n \geq 1 | X_n \in A\},$$

i.e. $N(A)$ denotes the number of transitions until the first return to the set A where $N(A) = \infty$ if $X_n \notin A$ for all $n \geq 1$. Also for any $A \subset I$ and $f \in F$, define the taboo probability

$$(3.1) \qquad {}_A p_{ij}^n(f) = \Pr_{f^{(\infty)}}\{X_n = j,\ X_k \notin A \text{ for } 1 \leq k \leq n-1 | X_0 = i\},$$

$$i,j \in I \text{ and } n = 1,2,\ldots\ .$$

Observe that

$$(3.2) \qquad E_{f^{(\infty)}}\{N(A)|X_0=i\} = 1 + \sum_{n=1}^{\infty} \sum_{j \notin A} {}_A p_{ij}^n(f).$$

In addition to assumption 1 of section 1 we now introduce the following assumption.

ASSUMPTION 2.

(a) *There is a finite set* K *such that for any* $i \in I$ *the quantities* $u^*(i)$ *and* $y^*(i)$ *are finite where*

$$(3.3) \qquad \sup_{f \in F} E_{f^{(\infty)}}\{\sum_{k=0}^{N(K)-1} \tau_k | X_0 = i\} = u^*(i) \text{ for all } i \in I$$

and

$$(3.4) \qquad \sup_{f \in F} E_{f^{(\infty)}}\{\sum_{k=0}^{N(K)-1} |c(X_k,a_k)|\,|X_0 = i\} = y^*(i) \text{ for all } i \in I.$$

(b) *For any* $f \in F$, *the stochastic matrix* $P(f)$ *has no two disjoint closed sets.*

(c) *For any* $i \in I$, *both* $\Sigma_{j \in I} p_{ij}(a) u^*(j)$ *and* $\Sigma_{j \in I} p_{ij}(a) y^*(j)$ *are continuous on* $A(i)$.

In words, assumption 2(a) requires the existence of a finite set K such that the supremum over all stationary policies of both the expected time ant the total expected absolute cost incurred until the first return to the set K are finite for any starting state. To satisfy assumption 2(a) in applications it may be necessary to exclude in certain states those actions which are far from being "optimal", e.g. in an M/M/c queueing system with a controllable number of operating servers we may consider only policies under which all c servers will be operating when the queue size exceeds some large value. In case the quantities $c(i,a)$ and $\tau(i,a)$ are uniformly bounded and the set $(P(f), f\in F)$ of stochastic matrices satisfies one of the conditions C1-C10 of section 2 then assumptions 2(a) and 2(c) hold with bounded functions u^* and y^*. Using results from positive dynamic programming, it is readily verified that assumption 2(a) is equivalent to the Liapunov condition requiring the existence of a finite set K and a finite non-negative function $x(i)$, $i\in I$ such that (cf. Federgruen, Hordijk & Tijms (1978b))

(3.5) $$|c(i,a)| + \tau(i,a) + \sum_{j\notin K} p_{ij}(a)x(j) \le x(i) \text{ for all } i\in I \text{ and } a\in A(i).$$

By assumption 1(c) and 2(a), we have

(3.6) $$E_{f^{(\infty)}}\{N(K)|X_0=i\} \le \frac{u^*(i)}{\varepsilon} \text{ for all } i\in I \text{ and } f\in F.$$

Define

(3.7) $$q_{ij}(f) = \sum_{n=1}^{\infty} {}_K p_{ij}^n(f),\ i\in I,\ j\in K,$$

i.e. $q_{ij}(f)$ is the probability that at the first return to the set K the transition occurs into state j starting from state i and using policy $f^{(\infty)}$. Observe that, by (3.6),

(3.8) $$\sum_{j\in K} q_{ij}(f) = 1 \text{ for all } i\in I.$$

For any $f\in F$, define for $i\in I$ and $j\in K$ the (possibly infinite) quantity

(3.9) $\nu_{ij}(f)$ = expected number of returns to the set K until the first transition into state j occurs starting from state i and using policy $f^{(\infty)}$.

We now have the following Theorem, see Federgruen, Hordijk & Tijms (1978b).

THEOREM 3.1. *Suppose that the assumptions* 1-2(a),(b) *hold. Then*

(a) *For any* $f \in F$, *the finite stochastic matrix* $(q_{ij}(f))$, $i,j \in K$ *has no two disjoint closed sets.*

(b) *For any* $i \in I$ *and* $j \in K$, *the probability* $q_{ij}(f)$ *is continuous on* F.

(c) *There is a finite number* B *such that for any* $f \in F$ *a state* $s_f \in K$ *exists for which* $\nu_{is_f}(f) \leq B$ *for all* $i \in I$ *and*

$$(3.10)\quad E_{f^{(\infty)}}\{\sum_{k=0}^{N(\{s_f\})-1} \tau(X_k,a_k) \mid X_0=i\} \leq u^*(i)+B \text{ for all } i \in I$$

and

$$(3.11)\quad E_{f^{(\infty)}}\{\sum_{k=0}^{N(\{s_f\})-1} |c(X_k,a_k)| \mid X_0=i\} \leq y^*(i)+B \text{ for all } i \in I.$$

This Theorem is essential in establishing the optimality result for the average cost criterion. To do this, we shall use a standard technique developed in Taylor (1965) and Ross (1968). We first analyse a discounted cost model and therefore define for any $i \in I$, policy π and discount factor $\beta > 0$

$$(3.12)\quad V_\beta(i,\pi) = E_\pi\{\sum_{n=0}^{\infty} e^{-\beta\Sigma_{k=0}^{n-1}\tau(X_k,a_k)} c(X_n,a_n) \mid X_0=i\}.$$

Furhter, let

$$(3.13)\quad V_\beta(i) = \inf_\pi V_\beta(i,\pi) \text{ for } i \in I.$$

It is not difficult to verify that for any $\beta>0$ the quantity $V_\beta(i,\pi)$ is well-defined for any $i \in I$ and policy π and that, for some constant c_β,

$$(3.14)\quad |V_\beta(i)| \leq c_\beta y^*(i) \text{ for all } i \in I.$$

Moreover, cf. Federgruen, Hordijk & Tijms (1978b)

$$(3.15)\quad V_\beta(i) = \min_{a \in A(i)} \{c(i,a) + e^{-\beta\tau(i,a)} \sum_{j \in I} p_{ij}(a) V_\beta(j)\} \text{ for all } i \in I,$$

where for any stationary policy $f_\beta^{(\infty)}$ such that the action $f_\beta(i)$ minimizes the right size of (3.15) for all $i \in I$ holds

$$(3.16)\quad V_\beta(i,f_\beta^{(\infty)}) = V_\beta(i) \text{ for all } i \in I.$$

Using Theorem 3.1 and (3.16) it can be shown that for any fixed state $s\in I$ there exists finite constants $\beta^*>0$ and c such that for all $0<\beta<\beta^*$

$$(3.17)\quad |\beta V_\beta(s)| \le c \text{ and } |V_\beta(i)-V_\beta(s)| \le c(u^*(i)+y^*(i)) \text{ for all } i\in I.$$

For any $\beta>0$, let $f_\beta\in F$ be such that $f_\beta(i)$ minimizes the right size of (3.15) for all $i\in I$. Now, using (3.17), assumption 1(a) and the diagonalization method, we can find a sequence $\{\beta_k\}$ with $\beta_k\to 0$ as $k\to\infty$, a constant g, a function v(i), $i\in I$ and an action $a(i)\in A(i)$ for any i I such that

$$(3.18)\quad \lim_{k\to\infty} \beta_k V_{\beta_k}(s) = g,\ \lim_{k\to\infty} V_{\beta_k}(i)-V_{\beta_k}(s) = v(i) \text{ and } \lim_{k\to\infty} f_{\beta_k}(i) = a(i) \text{ for all } i\in I.$$

Then we obtain by using (3.15) the following Theorem (see Federgruen, Hordijk & Tijms (1978b))

THEOREM 3.2. *Suppose that the assumptions* 1-2 *hold. Then there exists a constant* g *and a function* v(i), $i\in I$ *such that*

$$(3.19)\quad \sup_{i\in I} \frac{|v(i)|}{u^*(i)+y^*(i)} < \infty$$

satisfying the average cost optimality equation

$$(3.20)\quad v(i) = \min_{a\in A(i)} \{\tau(i,a)-g\tau(i,a) + \sum_{j\in I} p_{ij}(a)v(j)\} \text{ for all } i\in I.$$

The assumptions 1-2 are in general not enough to guarantee that an average cost optimal stationary policy exists as follows from the fact that the assumptions 1-2 are satisfied in the counterexample of Fisher & Ross (1968) in which example any average cost optimal policy is non-stationary. In case the condition (1.5) holds, then any stationary $f^{(\infty)}$ such that the action f(i) minimizes the right side of (3.20) for all $i\in I$ is (strongly) average cost optimal and has g as its average cost for any starting state i. It has been shown in Federgruen, Hordijk & Tijms (1978b) that a sufficient condition for (1.5) is given by

$$(3.21)\quad \lim_{n\to\infty} \hat{P}^n(f)(u^*+y^*) = 0 \text{ for all } f\in F$$

with $\hat{P}^n(f)$ is the n-fold matrix product of the substochastic matrix $\hat{P}(f)=(\hat{p}_{ij}(f))$, $i,j\in I$ with itself where $\hat{p}_{ij}(f) = p_{ij}(f)$ for $i\in I$ and $j\notin K$ and $\hat{p}_{ij}(f)=0$ for $i\in I$ and $j\in K$.

Under the assumptions 1-2 and condition (3.21) we have that for any solution $\{g, v(i), i \in I\}$ to (3.21) such that (3.19) holds, the constant g is *uniquely determined* as the minimal average expected cost per unit time. Under the additional assumption that for any strongly average cost optimal stationary policy the total expected costs incurred until the first return to the finite set K is finite for any starting state when the one step costs in state i are given by $u^*(i)+y^*(i)$, the function $v(i)$, $i \in I$ is *uniquely determined up to an additive constant* as can be shown by using the proof of Lemma 3 in Hordijk, Schweitzer & Tijms (1975).

References

1. Bather, J. (1973), Optimal decision procedures for finite Markov chains, Parts I, II and III, *Adv. Appl. Prob.* 5, 328-339, 521-540, 541-553.
2. Derman, C. and R. Strauch, (1966), A note on memoryless rules for controlling sequential control processes, *Ann. Math. Statist.* 37, 276-278.
3. Derman, C. (1966), Denumerable state Markovian decision processes-average cost criterion, *Ann. Math. Statist.* 37, 1545-1554.
4. Doob, J.L. (1953), Stochastic Processes, Wiley, New York.
5. Federgruen, A. and H.C. Tijms (1978), The optimality equation in average cost denumerable state semi-Markov decision problems. recurrency conditions and algorithms, *J. Appl. Prob.* 15, 356-373.
6. Federgruen, A., Hordijk, A. and H.C. Tijms (1978a), A note on simultaneous recurrence conditions on a set of denumerable stochastic matrices, *J. Appl. Prob.* 15, 842-847.
7. Federgruen, A, Hordijk, A. and H.C. Tijms (1978b), Denumerable state semi-Markov decision processes with unbounded cost criterion, Report BW 92/78, Mathematisch Centrum, Amsterdam.
8. Fisher, L. (1968), On recurrent denumerable decision processes, *Ann. Math. Statist.* 39, 424-434.
9. Fisher, L. and S.M. Ross (1968), An example in denumerable decision processes, *Ann. Math. Statist.* 39, 674-675.
10. Hordijk, A. (1974), *Dynamic Programming and Markov Potential Theory*, Mathematical Centre Tract no. 51, Mathematisch Centrum, Amsterdam.
11. Hordijk, A., Schweitzer, P.J. and H.C. Tijms (1975), The asymptotic behaviour of the minimal total expected cost for the denumerable state Markov decision model, *J. Appl. Prob.* 12. 298-305.
12. Hordijk, A. (1976), Regenerative Markov decision models, in: Mathematical Programming Study, 6, 49-72, North-Holland, Amsterdam.
13. Isaacson, D. and Luecke, G.R. (1978), Strongly ergodic Markov chains and rates of convergence using spectral conditions, *Stoch. Proc. Appl.* 7, 113-121.
14. Ross, S.M. (1968), Non-discounted denumerable Markov decision models. *Ann. Math. Statist.* 39, 412-423.
15. Ross, S.M. (1970), *Applied Probability Models with Optimization Applications*, Holden-Day Inc., San Francisco.
16. Ross, S.M. (1971), On the nonexistence of ε-optimal randomized stationary policies in average cost Markov decision models. *Ann. Math. Statist.* 42, 1767-1768.
17. Royden, H.L. (1968), Real Analysis (2nd. ed.), MacMillan, New York
18. Rudin, W. (1964), Principles of Mathematical Analysis (2nd. ed.) McGraw-Hill, New York.
19. Taylor, H.M. (1965), Markovian sequential replacement processes, *Ann. Math. Statist.* 36, 1677-1694.