Stationary Markovian decision problems II

Document status and date:
Published: 01/01/1975

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

Download date: 16. Nov. 2023

TECHNOLOGICAL UNIVERSITY EINDHOVEN

Department of Mathematics

STATISTICS AND OPERATIONS RESEARCH GROUP

Memorandum COSOR 75-15

Stationary Markovian decision problems II

by

J. Wijngaard

Stationary Markovian decision problems II

by

J. Wijngaard

1. Introduction

A stationary Markovian decision problem (SMD) is a set of pairs $\{(P_\alpha, r_\alpha)\}$, $\alpha \in A$ where P_α is a Markov provess and r_α a nonnegative function on the state space (cost function). The elements $\alpha \in A$ are called strategies. In [5] conditions are derived for the existence of an average optimal strategy. The most important conditions were the boundedness of r_α and the quasi-compactness of P_α, which is equivalent to the Doeblin-condition. For a countable state space the Doeblin-condition for a Markov process P is equivalent to the existence of a finite set A, an integer n, and an $\epsilon > 0$, such that the probability of being in the set A after n transitions $P^{(n)}(u,A) \geq \epsilon$ for each starting state u. To show how severe this condition and hence quasi-compactness is we consider the following inventory problem:

> At the beginning of each period the inventory level is assumed to be
> ..... $-2,-1,0,1,2,\ldots$. One may order a quantity of at most R units,
> the delivery is instantaneous. During the period there is a demand for
> $0,1,2,\ldots$ units with a probability of $p_0, p_1, p_2, \ldots$. The transition
> probability under order policy α is $P_\alpha(i,j) = p_{i+\alpha(i)-j}$, $\alpha(i)$ is the
> quantity to order in state i).

If R is large enough there are policies α such that for each state i one can find a finite set A, an integer n, and an $\epsilon > 0$ such that $P_\alpha^{(n)}(i,A) \geq \epsilon$. However, if j is more than nR units below the lowest element of A, then $P_\alpha^{(n)}(j,A) = 0$. Hence there is no policy such that the corresponding Markov process satisfies the Doeblin-condition. Such decision processes can be studied by introducing embedded Markov processes. In this paper we do not assume the quasi-compactness of P_α nor the boundedness of r_α. Instead of that we state the existence of a subset A of the state space such that the embedded Markov process of P_α on A exists and is quasi-compact for all $\alpha \in A$ and the recurrence time and costs until A are bounded on A.

Embedded Markov processes are introduced in section 2. We derive some properties of Markov processes with a quasi-compact embedded Markov process. Section 3 deals with the existence of the average costs for these Markov processes with unbounded cost function. The continuity of the average costs and

the existence of an optimal α for problems $\{(P_{\alpha}, r_{\alpha})\}$, $\alpha \in A$ is worked out in section 4. In section 5 the results of section 4 are applied to the case with a countable state space and related to the results of Ross [3] and Hordijk [1].

2. Embedded Markov processes

We assume that P is a Markov process on the measurable space (V, Σ). For $A \in \Sigma$ the sub-markov process I_A is defined by the sub-transition probability

$$I_A(u, E) := 1_{A \cap E}(u), \quad u \in V, \ E \in \Sigma .$$

Instead of PI_A we shall write P_A.

The next lemma serves as an introduction to the concept of an embedded Markov process.

Lemma 1. Let $A \in \Sigma$, $B := V \ A$. Define the function Q on $V \times \Sigma$ by

$$Q(u, E) = \sum_{n=0}^{\infty} (P_B^n P_A 1_E)(u) \quad \text{for all } u \in V, \ E \in \Sigma .$$

Then Q is a sub-transition probability on $V \times \Sigma$, the operator Q on $\mathcal{B}(V, \Sigma)$ is given by

$$(1) \qquad (Qf)(u) = \sum_{n=0}^{\infty} (P_B^n P_A f)(u) \quad \text{for } u \in V, \ f \in \mathcal{B}(V, \Sigma) ,$$

and the operator Q on $M(V, \Sigma)$ by

$$(2) \qquad (\mu Q)(E) = \sum_{n=0}^{\infty} (\mu P_B^n P_A)(E) \quad \text{for } E \in \Sigma, \ \mu \in M(V, \Sigma) .$$

Furthermore, Q is a Markov process on (V, Σ) if and only if

$$\lim_{n \to \infty} (P_B^n 1_V)(u) = 0 \quad \text{for all } u \in V .$$

Proof. We have $P_A = P - P_B$. Hence

$$\sum_{n=0}^{N} P_B^n P_A 1_V = \sum_{n=0}^{N} P_B^n P 1_V - \sum_{n=0}^{N} P_B^{n+1} 1_V = \sum_{n=0}^{N} P_B^n 1_V - \sum_{n=0}^{N} P_B^{n+1} 1_V = 1_V - P_B^{N+1} 1_V ,$$

which implies that $Q(u,E) \leq Q(u,V) \leq 1$ for $u \in V$, $E \in \Sigma$. The measurability of Q as function of u and the σ-additivity as function of E are easy to verify. Hence Q is a sub-transition probability on $V \times \Sigma$ and a transition probability if and only if $\lim_{n \to \infty} (P_B^n 1_V)(u) = 0$ for all $u \in V$. The equations (1) and (2) are direct consequences of the definition of Q. $\quad\quad\quad\square$

Let $A \in \Sigma$, $B := V \backslash A$. The sub-Markov process Q on (V,Σ) with sub-transition probability

$$Q(u,E) := \sum_{n=0}^{\infty} (P_B^n P_A 1_E)(u), \quad u \in V, \; E \in \Sigma$$

is called the sub-Markov process of P *induced* by A.
It is clear that the restriction of Q to $A \times \Sigma_A$ is a sub-transition probability on $A \times \Sigma_A$. The sub-Markov process on (A, Σ_A) corresponding to this sub-transition probability is called the *embedded* sub-Markov process of P on A. By lemma 1 the sub-Markov process induced by A is a Markov process if and only if $\lim_{n \to \infty} (P_B^n 1_V)(u) = 0$ for all $u \in V$, that means, if the probability that the system will never reach the set A is zero.
The relationship between invariant sets, functions and measures of P and those of a process Q of P induced by a set $A \in \Sigma$ is shown in the next lemma.

<u>Lemma 2.</u> Let $A \in \Sigma$, $B := V\backslash A$. Assume that $\lim_{n \to \infty} (P_B^n 1_V)(u) = 0$ for all $u \in V$ and let Q be the embedded Markov process of P on A. If $\mu \in M(V,\Sigma)$ and $f \in B(V,\Sigma)$ are invariant under P, then $\mu I_A Q = \mu I_A$ and $Qf = f$. Conversely, if $Qf = f$, then $Pf = f$ and if E is an invariant set under Q then

$$\bar{E} := \{u \mid Q(u,E) = 1\}$$

is an invariant set under P.

<u>Proof.</u> The proof of the invariance of μI_A and f under Q is straightforward using

$$\mu I_A P_B^n P_A = (\mu - \mu I_B) P_B^n P_A = \mu P_B^n P_A - \mu I_B P_B^n P_A =$$

$$= \mu P I_B P_B^{n-1} P_A - \mu I_B P_B^n P_A = \mu I_B P_B^{n-1} P_A - \mu I_B P_B^n P_A \, ,$$

and

$$P_B^n P_A f = P_B^n Pf - P_B^{n+1} f = P_B^n f - P_B^{n+1} f .$$

Conversely, suppose $Qf = f$. Then

$$Pf = P_A f + P_B f = P_A f + P_B Qf = P_A f + P_B \sum_{n=0}^{\infty} P_B^n P_A f =$$

$$= \sum_{n=0}^{\infty} P_B^n P_A f = Qf = f .$$

Finally, let E be an invariant set under Q and let $\bar{E} := \{u \mid Q(u,E) = 1\}$. From $Q = P_A + P_B Q$ we conclude

$$Q1_{A\backslash E} = P1_{A\backslash E} + P_B Q1_{A\backslash E} \geq P1_{A\backslash \bar{E}} + P_{B\backslash \bar{E}} Q1_{A\backslash E} .$$

Since on $\bar{E}$ we have $Q1_{A\backslash E} = 0$, it follows that $P1_{A\backslash \bar{E}} = 0$ on $\bar{E}$ and $P_{B\backslash \bar{E}} Q1_{A\backslash E} = 0$ on $\bar{E}$. From $Q1_A = 1$ and the definition of $\bar{E}$ we infer that on $V\backslash \bar{E}$ and in particular on $B\backslash \bar{E}$ we have $Q1_{A\backslash E} > 0$. It follows that $P1_{B\backslash \bar{E}} = 0$ on $\bar{E}$. Therefore $P1_{V\backslash \bar{E}} = 0$, $P1_{\bar{E}} = 1$ on $\bar{E}$. □

3. Average costs

Let P be a Markov process on (V,Σ) and r a nonnegative measurable function on V. The average costs of (P,r), starting in u, $g(u)$, are equal to $\lim_{n\to\infty} \frac{1}{n} \sum_{\ell=0}^{n-1} (P^\ell r)(u)$, if this limit exists. (The integral $\int_V P(u,ds)f(s)$, if existing, is denoted by $(Pf)(u)$.)

In this section conditions sufficient for the existence of the average costs will be given and it will be shown that these costs can be written as the quotient of the recurrence costs and recurrence time to a set A. This will be done by considering the equations

$$x = Px$$
$$y = r - x + Py$$

in the complex valued measurable functions x,y on V. These equations are called the (P,r)-*equations*. If P is quasi-compact and r is bounded, the existence of a solution of the (P,r)-equations is a consequence of the strong ergodic theorem (see [4]). The solution is given by

$$x := \lim_{n\to\infty} \frac{1}{n} \sum_{\ell=0}^{n-1} P^\ell r = g$$

$$y := \frac{1}{d} \sum_{m=0}^{d-1} \lim_{k \to \infty} \sum_{\ell=0}^{kd-1+m} P^{\ell}(r-x) \; ,$$

where d is an integer such that $\lambda^d = 1$ for all eigenvalues λ of P on the unit circle. Now the existence of a solution of the (P,r)-equations will be proved under somewhat weaker conditions. The quasi-compactness of P is replaced by the quasi-compactness of the embedded Markov process of P on some set $A \in \Sigma$. The boundedness of r is replaced by the boundedness of the expected time and expected costs until the first recurrence to A. The function x will again turn out to be equal to the average costs.

Definition 3. Let f be a nonnegative measurable real valued function on V and let A be a measurable set. The Markov process P is said to be (A,f)-*recurrent* if

i) $P_B^m f$ exists for all $m \in \mathbb{N}$, $(B := V \backslash A)$,

ii) the sum $\sum_{m=0}^{\infty} (P_B^m f)(u)$ exists for all $u \in V$,

iii) the convergence of $\sum_{m=0}^{\infty} (P_B^m f)(u)$ is uniform on A and $\sum_{m=0}^{\infty} (P_B^m f)(u)$ is bounded on A.

In the rest of this section we assume that A is a fixed measurable set such that P is $(A,1_V)$-recurrent and (A,r)-recurrent and further that the embedded Markov process Q of P on A is quasi-compact, (Q interpreted as a Markov process on (A,Σ_A)). The $(A,1_V)$-recurrency implies that the embedded sub-Markov process of P on A is a Markov process.

Let E_j, $j = 1,\ldots,n$ be the maximal invariant sets of Q, $F := \bigcup_{j=1}^{n} E_j$, and $\Delta := A \backslash F$.

Theorem 4. The (P,r)-equations have a solution.

Proof. By the strong ergodic theorem the spectral radius of $Q_\Delta := Q I_\Delta$ is smaller that 1. Hence, each of the equations $x = Q_{E_j} 1_{E_j} + Q_\Delta x$ in $\mathcal{B}(A,\Sigma_A)$ has a unique solution

$$g_j := \sum_{\ell=0}^{\infty} Q_\Delta^\ell Q_{E_j} 1_{E_j} \; .$$

Using $(Q_\Delta f)(u) = 0$ for all $f \in B(A, \Sigma_A)$, $u \in E_j$, $j = 1, 2, \ldots, n$, we get
$g_j(u) = 1$ for $u \in E_j$ and $g_j(u) = 0$ for $u \in E_i$ if $i \neq j$. This means that
$Q_{E_j} 1_{E_j} = Q_{E_j} g_j = (Q - Q_\Delta) g_j$ and that g_j is a solution of the equation
$x - Qx = 0$ in $B(A, \Sigma_A)$.

It is possible to extend g_j to a solution g_j^* of the equation $x - Qx = 0$ in
$B(V, \Sigma)$ by defining $g_j^* := Qg_j$, where Q is used as an operator on $B(A, \Sigma_A)$ to
$B(V, \Sigma)$.

Each function g_j^*, $j = 1, \ldots, n$ is a solution of the equation $x = Px$ since

$$Pg_j^* = P_B g_j^* + P_A g_j^* = P_B Q g_j^* + P_A g_j^* = \sum_{\ell=1}^{\infty} P_B^\ell P_A g_j^* + P_A g_j^* = Q g_j^* = g_j^* .$$

The problem is to choose a linear combination x of the g_j^* such that the
equation $y = r - x + Px$ has also a solution. Let Q_j be the restriction of
Q to $E_j \times \Sigma_{E_j}$. The (A, r)-recurrency and $(A, 1_V)$-recurrency of P imply the
boundedness on A of the functions $\sum_{\ell=0}^{\infty} P_B^\ell r$ and $\sum_{\ell=0}^{\infty} P_B^\ell g_j^*$, $j = 1, \ldots, n$. For
convenience we shall write Tf instead of $\sum_{\ell=0}^{\infty} P_B^\ell f$ for $f = r$ or f is bounded.
Notice that $P_B Tf = \sum_{\ell=1}^{\infty} P_B^\ell f$.

The restrictions of Tr and Tg_j^* to E_j are elements of $B(E_j, \Sigma_{E_j})$. Therefore
both $\lim_{n \to \infty} \frac{1}{n} \sum_{\ell=0}^{n-1} Q_j^\ell Tr$ and $\lim_{n \to \infty} \frac{1}{n} \sum_{\ell=0}^{n-1} Q_j^\ell Tg_j^*$ are elements of $N(I - Q_j)$. Since
$\dim N(I - Q_j) = 1$ there is a constant c_j such that

$$\lim_{n \to \infty} \frac{1}{n} \sum_{\ell=0}^{n-1} Q_j^\ell (Tr - c_j Tg_j^*) = 0 .$$

Using this it is straightforward to prove that

$$\lim_{n \to \infty} \frac{1}{n} \sum_{\ell=0}^{n-1} Q^\ell (Tr - Tg^*) = 0 \text{ on } A, \text{ where } g^* := \sum_{j=1}^{n} c_j g_j^* .$$

We shall show that the equation $y = r - g^* + Py$ has a solution.
Let the integer d be such that $\lambda^d = 1$ for all eigenvalues λ of Q on the
unit circle. By the strong ergodic theorem the function f' on A defined by

$$f' := \frac{1}{d} \sum_{m=0}^{d-1} \lim_{k \to \infty} \sum_{\ell=0}^{kd-1+m} Q^\ell (Tr - Tg^*)$$

is a solution of the equation $y = Tr - Tg^* + Qy$ in $B(A, \Sigma_A)$. The function f'
can be extended to a function f on V by defining $f := Tr - Tg^* + Qf'$. The

function f is a solution of the equation $y = r - g^* + Py$. This can be seen as follows:

$$Pf = P_Bf + P_Af = P_B(Tr - Tg^* + Qf) + P_Af = P_BT(r - g^*) + P_BQf +$$

$$+ P_Af = P_BT(r - g^*) + \sum_{\ell=1}^{\infty} P_B^{\ell}P_Af + P_Af = P_BT(r - g^*) + Qf =$$

$$= P_BT(r - g^*) + f - T(r - g^*) = f + \sum_{\ell=1}^{\infty} P_B^{\ell}(r - g^*) -$$

$$+ \sum_{\ell=0}^{\infty} P_B^{\ell}(r - g^*) = f - r + g^* . \qquad \square$$

Now some properties of the solution g^*, f^* of the (P,r)-equations are given. The next lemma is preliminary.

Lemma 5. Let

$$Tr := \sum_{\ell=0}^{\infty} P_B^{\ell}r \quad \text{and} \quad Tl_V := \sum_{\ell=0}^{\infty} P_B^{\ell}l_V .$$

Then P^mTr and P^mTl_V exist for all $m \in \mathbb{N}$ and

$$\lim_{m \to \infty} \frac{1}{m} (P^mTr)(u) = \lim_{m \to \infty} \frac{1}{m} (P^mTl_V)(u) = 0 \quad \text{for all } u \in V .$$

Proof. Substitution of $P = P_A + P_B$ in P^{m+1} yields

$$P^{m+1} = P^mP_A + P^mP_B = P^mP_A + P^{m-1}P_AP_B + P^{m-1}P_B^2 = \ldots = \sum_{k=0}^{m} P^{m-k}P_AP_B^k + P_B^{m+1} .$$

Hence

$$P^{m+1}Tr = \sum_{k=0}^{m} P^{m-k}P_A \sum_{\ell=k}^{\infty} P_B^{\ell}r + \sum_{\ell=m+1}^{\infty} P_B^{\ell}r \leq \sum_{k=0}^{m} P^{m-k}P_ATr + Tr .$$

The existence of $P^{m+1}Tr$ is implied by the existence of Tr and the boundedness of Tr on A. The existence of $P^{m+1}Tl_V$ is proved similarly. For each $\epsilon > 0$ there is an integer N_ϵ such that

$$\sum_{\ell=N_\varepsilon}^{\infty} (P_B^\ell r)(u) < \varepsilon \qquad \text{for all } u \in A \ .$$

For $m > N_\varepsilon$ we have

$$P^{m+1}Tr = \sum_{k=0}^{N_\varepsilon} P^{m-k}P_A \sum_{\ell=k}^{\infty} P_B^\ell r + \sum_{k=N_\varepsilon+1}^{m} P^{m-k}P_A \sum_{\ell=k}^{\infty} P_B^\ell r + \sum_{\ell=m+1}^{\infty} P_B^\ell r \ .$$

Let $\| Tr \|_A := \sup_{u \in A} (Tr)(u)$. Then

$$(P^{m+1}Tr)(u) \leq (N_\varepsilon + 1)\| Tr \|_A + (m - N_\varepsilon)\varepsilon + \sum_{\ell=m+1}^{\infty} (P_B^\ell r)(u) \ .$$

Using standard arguments we can show that $\lim\limits_{m \to \infty} \frac{1}{m} (P^m Tr)(u) = 0$ for all $u \in V$. That $\lim\limits_{m \to \infty} \frac{1}{m} (P^m T1_V)(u) = 0$ can be proved similarly. $\qquad \square$

Theorem 6. Let the functions g^* and f be as constructed in the proof of theorem 4. Then $P^m f$ exists for all $m \in \mathbb{N}$, $\lim\limits_{m \to \infty} \frac{1}{m} P^m f = 0$, and $g^* = g$ (the average costs of (P,r)). Let g_1, f_1 be another solution of the (P,r)-equations, such that $\lim\limits_{m \to \infty} \frac{1}{m} P^m f_1 = 0$, then $g_1 = g$ and $f - f_1 = Q(f - f_1)$.

Proof. The functions g^* and f on V were defined in the following way:

$$g^* := \sum_{j=1}^{n} c_j g_j^*, \quad \text{where } g_j^* := Q g_j \text{ and } g_j \in B(A, \Sigma_A) \ ;$$

$$f := Tr - Tg^* + Qf' \quad \text{where } f' \in B(A, \Sigma_A) \ .$$

Hence g^* and Qf' are bounded. By lemma 5 $P^m f$ exists for all $m \in \mathbb{N}$ and

$$(1) \qquad \lim\limits_{m \to \infty} \frac{1}{m} (P^m f) = 0 \ .$$

Repeated substitution of $f(u) = r(u) - g^*(u) + (Pf)(u)$ in its right-hand side yields

$$f(u) = \sum_{\ell=0}^{m-1} (P^\ell r)(u) - \sum_{\ell=0}^{m-1} (P^\ell g^*)(u) + (P^m f)(u) \ .$$

Hence

$$\frac{f(u) - (P^m f)(u)}{m} = \frac{1}{m} \sum_{\ell=0}^{m-1} (P^\ell r)(u) - g^*(u)$$

and by (1)

$$g^*(u) = \lim_{m\to\infty} \frac{1}{m} \sum_{\ell=0}^{m-1} (P^\ell r)(u) = g(u) \quad \text{for all } u \in V .$$

Now we consider the solution (g_1, f_1) of the (P,r)-equations. As for the solution (g^*, f) we can prove

$$g_1(u) = \lim_{m\to\infty} \frac{1}{m} \sum_{\ell=0}^{m-1} (P^\ell r)(u) .$$

Further the function $f - f_1$ satisfies $f - f_1 = P(f - f_1)$, hence by lemma 2, $f - f_1 = Q(f - f_1)$. $\qquad\qquad\square$

For $u \in E_j$ it is possible to write the average costs in a somewhat different way. To show this we need the following lemma.

<u>Lemma 7.</u> For all $f \in B(V, \Sigma)$ and for all $m \in \mathbb{N}$ the following relation holds, $(B := V\backslash A)$

$$(P_B^m Q f)(u) = (P_B^m Q f_{E_j}), \quad u \in E_j, \ j = 1, 2, \ldots, n .$$

<u>Proof.</u> It is sufficient to prove the assertion for nonnegative functions, namely, each $f \in B(V, \Sigma)$ can be written as $f = f_1 - f_2 + i(f_3 - f_4)$, where the functions f_1, f_2, f_3 and f_4 are nonnegative elements of B. Now assume that f is a nonnegative function in B. Substitution of $Q f_E = \sum_{\ell=0}^{\infty} P_B^\ell P_A f_E$ in $P_B^m Q f_E$ yields

(1) $\qquad P_B^m Q f_E = \sum_{\ell=m}^{\infty} P_B^\ell P_A f_E \le \sum_{\ell=0}^{\infty} P_B^\ell P_A f_E = Q f_E \quad \text{for all } E \in \Sigma .$

Furthermore

(2) $\qquad (Qf)(u) = (Q f_{E_j})(u) \quad \text{for } u \in E_j, \ j = 1, \ldots, n .$

Let $\bar{E}_j := V\backslash E_j$, then $f = f_{E_j} + f_{\bar{E}_j}$. By (1) and (2)

$$(P_B^m Qf_{\overline{E_j}})(u) \leq (Qf_{\overline{E_j}})(u) = 0 \quad \text{for } u \in E_j .$$

Hence

$$(P_B^m Qf)(u) = (P_B^m Qf_{E_j})(u) \quad \text{for } u \in E_j, \ j = 1,\ldots,n . \qquad \square$$

Theorem 8. For $u \in E_j$, $j = 1,\ldots,n$

$$\lim_{m \to \infty} \frac{1}{m} \sum_{\ell=0}^{m-1} (P^\ell r)(u) = \frac{\displaystyle\lim_{m \to \infty} \frac{1}{m} \sum_{\ell=0}^{m-1} (Q^\ell Tr)(u)}{\displaystyle\lim_{m \to \infty} \frac{1}{m} \sum_{\ell=0}^{m-1} (Q^\ell Tl_V)(u)} .$$

Proof. Let g^* be as constructed in the proof of theorem 4. Then $g^*(u) = c_j$ for $u \in E_j$, where

$$c_j = \frac{\displaystyle\lim_{m \to \infty} \frac{1}{m} \sum_{\ell=0}^{m-1} (Q_{E_j}^\ell Tr)(u)}{\displaystyle\lim_{m \to \infty} \frac{1}{m} \sum_{\ell=0}^{m-1} (Q_{E_j}^\ell Tg_j^*)(u)} .$$

But for $u \in E_j$

$$(Q_{E_j}^\ell Tr)(u) = (Q^\ell Tr)(u)$$

and

$$(Q_{E_j}^\ell Tg_j^*)(u) = (Q^\ell Tg_j^*)(u) .$$

Further

$$Tg_j^* = TQg_j^* \quad \text{and} \quad Tl_V = TQl_V .$$

By lemma 7

$$(TQg_j^*)(u) = (TQl_{E_j})(u) = (TQl_V)(u) \quad \text{for } u \in E_j .$$

Hence $(Tg_j^*)(u) = (Tl_V)(u)$ for $u \in E_j$. This completes the proof. $\qquad \square$

A more general result of this type is given by de Leve [2], part II, lemma 1.57.

4. Stationary Markovian decision problems

In this section we consider a stationary Markovian decision problem $\{(P_\alpha, r_\alpha)\}$, $\alpha \in A$ on (V, Σ) (for a definition see [5]). In [5] it was assumed that P_α is quasi-compact and r_α bounded. Now we assume the existence of a measurable set A such that

i)   for all $\alpha \in A$ the Markov process P_α is $(A, 1_V)$-recurrent and (A, r_α)-re-current,

ii)   the embedded Markov process Q_α of P_α on A is quasi-compact for all $\alpha \in A$, (Q_α is interpreted as a Markov process on (A, Σ_A)),

iii)   the functions $\sum_{n=0}^{\infty} P_{\alpha B}^n 1_V$ and $\sum_{n=0}^{\infty} P_{\alpha B}^n r_\alpha$ on A, with $B = V \backslash A$, are uniform-ly bounded on A.

Let for all $\alpha \in A$, n_α be the dimension of $N(I - Q_\alpha)$, $E_{\alpha j}$ for $j = 1, \ldots, n_\alpha$ the maximal invariant sets of Q_α, and $\pi_{\alpha j}$ the invariant probabilities of Q_α with support $E_{\alpha j}$. Let

$$E_\alpha := \sum_{j=1}^{n_\alpha} E_{\alpha j} \quad \text{and} \quad S_\alpha := \lim_{n \to \infty} \frac{1}{n} \sum_{\ell=0}^{n-1} Q_\alpha^\ell .$$

By theorem 6 the average costs of (P_α, r_α) starting in u,

$$g_\alpha(u) := \lim_{n \to \infty} \frac{1}{n} \sum_{\ell=0}^{n-1} (P_\alpha^\ell r_\alpha)(u)$$

exist for all $\alpha \in A$ and $u \in V$. In theorem 8 we proved

$$g_\alpha(u) = \frac{(S_\alpha T_\alpha r_\alpha)(u)}{(S_\alpha T_\alpha 1_V)(u)} \quad \text{for } u \in E_\alpha ,$$

where

$$T_\alpha f = \sum_{n=0}^{\infty} P_{\alpha B}^n f .$$

Hence g_α is constant on $E_{\alpha j}$ for $j = 1, \ldots, n_\alpha$. These constants are denoted by $g_{\alpha j}$. Let ρ be a metric on A such that

iv) $\lim_{\rho(\alpha, \alpha_0) \to 0} \| Q_\alpha - Q_{\alpha_0} \| = 0 \quad$ for all $\alpha_0 \in A$,

v) $\lim_{\rho(\alpha, \alpha_0) \to 0} |\pi_{\alpha_0 j}(T_\alpha r_\alpha) - \pi_{\alpha_0 j}(T_{\alpha_0} r_{\alpha_0})| = 0$

for all $\alpha_0 \in A$ and $j = 1, \ldots, n_{\alpha_0}$.

$$\lim_{\rho(\alpha, \alpha_0) \to 0} \left| \pi_{\alpha_0 j}(T_\alpha 1_V) - \pi_{\alpha_0 j}(T_{\alpha_0} 1_V) \right| = 0$$

for all $\alpha_0 \in A$ and $j = 1, \ldots, n_{\alpha_0}$.

For $A = V$ these assumptions are identical to the assumptions made in [5], section 4.

In the next two subsections we consider the continuity of g_α and the existence of an optimal strategy.

4.1. Continuity of g_α

Let A_n be the subset of A with all α such that $n_\alpha = n$. In [5] lemma 2.9, it was shown that the assumption iv) implies the continuity of S_α as operator valued function on A_n. This is used in the next lemma.

Lemma 9. Let $n \in \mathbb{N}$ and $\alpha_0 \in A_n$. Then there is a $\delta > 0$ such that for all $\alpha \in A_n$ with $\rho(\alpha, \alpha_0) < \delta$ and for all $i = 1, 2, \ldots, n$, $\pi_{\alpha_0 i}(E_{\alpha j}) > 0$ for precisely one $j \in \{1, 2, \ldots, n\}$.
Consider the set $A_{n\delta} := \{\alpha \in A_n \mid \rho(\alpha, \alpha_0) < \delta\}$. Let for all $\alpha \in A_{n\delta}$ and $i \in \{1, 2, \ldots, n\}$ the integer i_α be defined by $\pi_{\alpha_0 i}(E_{\alpha i_\alpha}) > 0$. Then for all $i = 1, 2, \ldots, n$ the functions $\pi_{\alpha_0 i}(E_{\alpha_0 i} \backslash E_{\alpha i_\alpha})$, $\| \pi_{\alpha_0 i} - \pi_{\alpha i_\alpha} \|$, and $\left| g_{\alpha_0 i} - g_{\alpha i_\alpha} \right|$ on A_n converge to 0 if $\rho(\alpha, \alpha_0)$ converges to 0.

Proof. The continuity of S_α on A_n implies the existence of a $\delta > 0$ such that $\| S_\alpha - S_{\alpha_0} \| < \frac{1}{4}$ for all $\alpha \in A_n$ with $\rho(\alpha, \alpha_0) < \delta$.
Let $i \in \{1, 2, \ldots, n\}$ and let ν_i be some probability on Σ_A with $\nu_i(E_{\alpha_0 i}) = 1$, then $\pi_{\alpha_0 i}(E_{\alpha_0 i}) = (\nu_i S_{\alpha_0})(E_{\alpha_0 i}) = 1$. Hence $(\nu_i S_\alpha)(E_{\alpha_0 i}) > \frac{3}{4}$ for all $\alpha \in A_n$ with $\rho(\alpha, \alpha_0) < \delta$. But

$$(\nu_i S_\alpha)(E_{\alpha_0 i}) = (\nu_i S_\alpha)(E_{\alpha_0 i} \cap (\bigcup_{j=1}^{n} E_{\alpha j}))$$

and therefore

$$(\nu_i S_{\alpha_0})(E_{\alpha_0 i} \cap (\bigcup_{j=1}^{n} E_{\alpha j})) > \tfrac{1}{2}$$

for all $\alpha \in A_n$ with $\rho(\alpha, \alpha_0) < \delta$. This implies the existence of at least one $j \in \{1, 2, \ldots, n\}$ such that $(\nu_i S_{\alpha_0})(E_{\alpha j}) = \pi_{\alpha_0 i}(E_{\alpha j}) > 0$ for $\alpha \in A_n$ with $\rho(\alpha, \alpha_0) < \delta$.

Suppose that for some $\alpha \in A_n$ with $\rho(\alpha, \alpha_0) < \delta$ there are two j's, j_1 and j_2 such that $\pi_{\alpha_0 i}(E_{\alpha j}) > 0$. Let the probabilities ν_{ij_1} and ν_{ij_2} on Σ_A be given by

$$\nu_{ij_1}(E) = \frac{\pi_{\alpha_0 i}(E \cap E_{\alpha j_1})}{\pi_{\alpha_0 i}(E_{\alpha j_1})} \quad \text{and} \quad \nu_{ij_2}(E) = \frac{\pi_{\alpha_0 i}(E \cap E_{\alpha j_2})}{\pi_{\alpha_0 i}(E_{\alpha j_2})} .$$

Then $\nu_{ij_1} S_{\alpha_0} = \nu_{ij_2} S_{\alpha_0} = \pi_{\alpha_0 i}$. Using $(\nu_{ij_1} S_\alpha)(E_{\alpha j_1}) = 1$ and $(\nu_{ij_2} S_\alpha)(E_{\alpha j_2}) = 1$ it is easy to see that

$$\pi_{\alpha_0 i}(E_{\alpha j_1}) = (\nu_{ij_1} S_{\alpha_0})(E_{\alpha j_1}) > \tfrac{3}{4}$$

and

$$\pi_{\alpha_0 i}(E_{\alpha j_2}) = (\nu_{ij_2} S_{\alpha_0})(E_{\alpha j_2}) > \tfrac{3}{4} .$$

The disjunctness of $E_{\alpha j_1}$ and $E_{\alpha j_2}$ implies $\pi_{\alpha_0 i}(E_{\alpha j_1} \cup E_{\alpha j_2}) > 1\tfrac{1}{2}$ which contradicts the fact that $\pi_{\alpha_0 i}$ is a probability. This completes the proof of the first part of the lemma.

Now let for all $\alpha \in A_{n\delta}$ and $i \in \{1, 2, \ldots, n\}$ the integers i_α be such that $\pi_{\alpha_0 i}(E_{\alpha i_\alpha}) > 0$. The probability ν_{ii_α} on Σ is given by

$$\nu_{ii_\alpha}(E) := \frac{\pi_{\alpha_0 i}(E \cap E_{\alpha i_\alpha})}{\pi_{\alpha_0 i}(E_{\alpha i_\alpha})} .$$

Then $\pi_{\alpha i_\alpha} = \nu_{ii_\alpha} S_\alpha$ and $\pi_{\alpha_0 i} = \nu_{ii_\alpha} S_{\alpha_0}$. Hence $\| \pi_{\alpha i_\alpha} - \pi_{\alpha_0 i} \| \leq \| S_\alpha - S_{\alpha_0} \|$ and therefore

$$(1) \qquad \lim_{\rho(\alpha, \alpha_0) \to 0} \| \pi_{\alpha i_\alpha} - \pi_{\alpha_0 i} \| = 0 .$$

Furthermore $\pi_{\alpha i_\alpha}(E_{\alpha_0 i} \backslash E_{\alpha i_\alpha}) = 0$ and hence

$$\lim_{\rho(\alpha,\alpha_0)\to 0} \pi_{\alpha_0 i}(E_{\alpha_0 i} \setminus E_{\alpha i_\alpha}) = 0 \ .$$

For $j = 1,\ldots,n$ we have

$$g_{\alpha j} = \frac{(S_\alpha T_\alpha r_\alpha)(u)}{(S_\alpha T_\alpha 1_V)(u)} \quad \text{for } u \in E_{\alpha j} \ ,$$

$(S_\alpha T_\alpha r_\alpha)(u) = \pi_{\alpha j}(T_\alpha r_\alpha)$ for $u \in E_{\alpha j}$, and $(S_\alpha T_\alpha 1_V)(u) = \pi_{\alpha j}(T_\alpha 1_V)$. But

$$\left| \pi_{\alpha i_\alpha}(T_\alpha r_\alpha) - \pi_{\alpha_0 i}(T_{\alpha_0} r_{\alpha_0}) \right| \leq \left| (\pi_{\alpha i_\alpha} - \pi_{\alpha_0 i}) T_\alpha r_\alpha \right| +$$

$$+ \left| \pi_{\alpha_0 i}(T_\alpha r_\alpha - T_{\alpha_0} r_{\alpha_0}) \right|$$

and

$$\left| \pi_{\alpha i_\alpha}(T_\alpha 1_V) - \pi_{\alpha_0 i}(T_{\alpha_0} 1_V) \right| \leq \left| (\pi_{\alpha i_\alpha} - \pi_{\alpha_0 i}) T_\alpha 1_V \right| +$$

$$+ \left| \pi_{\alpha_0 i}(T_\alpha 1_V - T_{\alpha_0} 1_V) \right| \ .$$

Using (1), the uniform boundedness of $T_\alpha r_\alpha$ and $T_\alpha 1_V$, and the continuity assumptions made at the beginning of this section, we get

$$\lim_{\rho(\alpha,\alpha_0)\to 0} \left| g_{\alpha i_\alpha} - g_{\alpha_0 i} \right| = 0 \ . \qquad \square$$

This result implies the continuity of $g_{\alpha 1}$ on A_1. However, the condition

$$\lim_{\rho(\alpha,\alpha_0)\to 0} \| Q_\alpha - Q_{\alpha_0} \| = 0$$

is unnecessarily strong. It can be replaced by (see [4])

$$\lim_{\rho(\alpha,\alpha_0)\to 0} \| Q_\alpha^k(Q_\alpha - Q_{\alpha_0}) \| = 0 \quad \text{for some } k \geq 1 \ .$$

Now we shall prove that μg_α is continuous on A_n for each nonnegative measure μ.

Lemma 10. Let μ be a nonnegative measure on Σ_A. Then the function μg_α is continuous on A_n for all $n \in \mathbb{N}$.

Proof. We have $g_\alpha = P_\alpha g_\alpha$, hence by lemma 2, $g_\alpha = Q_\alpha g_\alpha$. Therefore

$$\int_A g_\alpha(u)\mu(du) = \int_A (Q_\alpha g_\alpha)(u)\mu(du) = \int_A (S_\alpha g_\alpha)\mu(du) =$$

$$= \int_A g_\alpha(u)(\mu S_\alpha)(du) .$$

The measure μS_α is a linear combination of the $\pi_{\alpha j}$, $j = 1,\ldots,n_\alpha$. So $(\mu S_\alpha)(A\backslash E_\alpha) = 0$ and

$$\int_A g_\alpha(u)(\mu S_\alpha)(du) = \int_{E_\alpha} g_\alpha(u)(\mu S_\alpha)(du) .$$

Let $n \in \mathbb{N}$ and $\alpha_0,\alpha \in A_n$. Then

$$\int_A g_\alpha(u)\mu(du) - \int_A g_{\alpha_0}(u)\mu(du) = \int_{E_\alpha} g_\alpha(u)\mu(S_\alpha - S_{\alpha_0})(du) +$$

$$+ \int_{E_\alpha} (g_\alpha(u) - g_{\alpha_0}(u))(\mu S_{\alpha_0})(du) + \int_{E_\alpha} g_{\alpha_0}(u)(\mu S_{\alpha_0})(du) -$$

$$- \int_{E_{\alpha_0}} g_{\alpha_0}(u)(\mu S_{\alpha_0})(du) .$$

The continuity of S_α as an operator valued function on A_n and the uniform boundedness of g_α imply

$$\lim_{\rho(\alpha,\alpha_0)\to 0} \left| \int_{E_\alpha} g_\alpha(u)\mu(S_\alpha - S_{\alpha_0})(du) \right| = 0 .$$

Using $(\mu S_{\alpha_0})(V\backslash E_{\alpha_0}) = 0$ we get

$$\int_{E_\alpha} (g_\alpha(u) - g_{\alpha_0}(u))(\mu S_{\alpha_0})(du) = \int_{E_\alpha \cap E_{\alpha_0}} (g_\alpha(u) - g_{\alpha_0}(u))(\mu S_{\alpha_0})(du)$$

and

$$\int_{E_\alpha} g_{\alpha_0}(u)(\mu S_{\alpha_0})(du) = \int_{E_\alpha \cap E_{\alpha_0}} g_{\alpha_0}(u)(\mu S_{\alpha_0})(du) .$$

Hence

$$(1) \qquad \int_{E_\alpha} (g_\alpha(u) - g_{\alpha_0}(u))(\mu S_{\alpha_0})(du) = \sum_{j=1}^{n} \int_{E_{\alpha_0 j} \cap E_\alpha} (g_\alpha(u) - g_{\alpha_0}(u))(\mu S_{\alpha_0})(du)$$

and

$$(2) \qquad \int_{E_\alpha} g_{\alpha_0}(u)(\mu S_{\alpha_0})(du) - \int_{E_{\alpha_0}} g_{\alpha_0}(u)(\mu S_{\alpha_0})(du) = - \int_{E_{\alpha_0} \setminus E_\alpha} g_{\alpha_0}(u)(\mu S_{\alpha_0})(du) \ .$$

We complete the proof by application of lemma 9 on (1) and (2), using that μS_{α_0} is a linear combination of the $\pi_{\alpha_0 j}$. $\qquad\qquad\qquad\square$

4.2. Existence of optimal strategies

In lemma 9 we proved the continuity of $g_{\alpha 1}$ on A_1. So, if $A = A_1$ and A is compact then an optimal strategy exists. If $A_1 \neq A$ but if A_1 is dominating A (see [5]) we may restrict our attention to the set A_1, which is easier to analyse since g_α is constant on V for $\alpha \in A_1$. To formulate conditions sufficient for A_1 to dominate A, some new concepts are needed.

Definition 11. The SMD is called A-*communicative* if for all $\alpha \in A$ and $j = 1,\ldots,n_\alpha$ there is an $\alpha_1 \in A_1$ such that $\pi_{\alpha_1 1}(E_{\alpha j}) > 0$.

Notice that A-communicativeness is equivalent to communicativeness as defined in [5] if $A = V$.

Definition 12. Let $\alpha \in A$ and $i \in \{1,2,\ldots,n_\alpha\}$. The set
$$\bar{E}_{\alpha i} := \{u \in V \mid Q_\alpha(u,E_{\alpha i}) = 1\}$$
is called the *extension* of $E_{\alpha i}$.
Notice that $E_{\alpha i} \subset \bar{E}_{\alpha i}$ and that by lemma 2, $\bar{E}_{\alpha i}$ is an invariant set of P_α.

Lemma 13. Let the SMD be complete (see [5]) and A-communicative. Then A_1 dominates A.

Proof. Let $\alpha \in A$. Choose j_0 such that $g_{\alpha j_0} = \min\limits_{j=1,2,\ldots,n_\alpha} g_{\alpha j}$. The A-communi-

cativeness implies the existence of a strategy $\alpha_1 \in A_1$ such that $\pi_{\alpha_1 1}(E_{\alpha j_0}) > 0$.

Let $C := \bar{E}_{\alpha j_0}$ and $\alpha_2 := \alpha C \alpha_1$ (apply strategy on C and strategy α_1 outside of

C). Since $\pi_{\alpha_1 1}(C) > 0$, under strategy $\alpha_2 = \alpha C \alpha_1$ the system will never reach

set C and the induced sub-Markov process Q'_{α_2} of P_{α_2} on C is a Markov process.

Further $\alpha_2 \in A_1$ and by lemma 2, $g_{\alpha_2} := Q'_{\alpha_2} g_{\alpha_2}$. The invariance of $C = \bar{E}_{\alpha j_0}$ un-

der P_α implies $g_{\alpha_2}(u) = g_{\alpha j_0}$ for $u \in C$. Hence

$$g_{\alpha_2}(u) = (Q'_{\alpha_2} g_{\alpha_2})(u) = \sum_{n=0}^{\infty} (P^n_{\alpha_2 V \backslash C} P_{\alpha_2 C} g_{\alpha_2})(u) =$$

$$= g_{\alpha j_0} \cdot \sum_{n=0}^{\infty} (P^n_{\alpha_2 V \backslash C} P_{\alpha_2 C} 1_V)(u) = g_{\alpha j_0} \cdot 1_V(u) \text{ for all } u \in V ,$$

which completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ □

The following theorem is an extension of theorem 4.9 of [5].

Theorem 14. Let the SMD be complete and A-communicative. If A is compact
then an optimal strategy exists.

Proof. Let $g := \inf\limits_{\alpha \in A_1} g_{\alpha 1}$. The compactness of A implies the existence of a

sequence $\{\alpha_k\}$ in A_1 converging to $\alpha_0 \in A$ such that $\lim\limits_{k \to \infty} g_{\alpha_k 1} = g$. Without loss

of generality we may assume that $\pi_j := \lim\limits_{k \to \infty} \pi_{\alpha_k 1}(E_{\alpha_0 j})$ exists for all

$j = 1,\ldots,n_{\alpha_0}$. We have

$$g_{\alpha_k 1} = \frac{\pi_{\alpha_k 1}(T_{\alpha_k} r_{\alpha_k})}{\pi_{\alpha_k 1}(T_{\alpha_k} 1_V)} \quad \text{for all } k = 1,2,3,\ldots .$$

As in the proof of theorem 4.9 of [5] we can show that

$$\lim_{k \to \infty} \pi_{\alpha_k 1}(T_{\alpha_k} r_{\alpha_k}) = \sum_{j=1}^{n_{\alpha_0}} \pi_j \cdot r_j \ ,$$

where $r_j := \pi_{\alpha_0 j}(T_{\alpha_0} r_{\alpha_0})$ and

$$\lim_{k \to \infty} \pi_{\alpha_k 1}(T_{\alpha_k} 1_V) = \sum_{j=1}^{n_{\alpha_0}} \pi_j \cdot t_j \ ,$$

where $t_j := \pi_{\alpha_0 j}(T_{\alpha_0} 1_V)$. Hence

$$g = \frac{\displaystyle\sum_{j=1}^{n_{\alpha_0}} \pi_j \cdot r_j}{\displaystyle\sum_{j=1}^{n_{\alpha_0}} \pi_j \cdot t_j}$$

and therefore

$$\min_{j=1,\ldots,n_{\alpha_0}} \{\frac{r_j}{t_j}\} \leq g \ .$$

But $g_{\alpha_0 j} = \frac{r_j}{t_j}$ for $j = 1,2,\ldots,n_{\alpha_0}$, which implies that

$$\min_{j=1,\ldots,n_{\alpha_0}} \{g_{\alpha_0 j}\} \leq g \ .$$

By lemma 13 there is an $\alpha \in A_1$ such that

$$g_\alpha(u) \leq \min_{j=1,\ldots,n_{\alpha_0}} \{g_{\alpha_0 j}\} \quad \text{for all } u \in V \ .$$

The strategy α is optimal. $\square$

4.3. Extensions and remarks

In subsection 4.2 we derived conditions for the existence of an optimal stra-
tegy. Optimality of some strategy α_0 implies of course the μ-optimality of
this strategy for each nonnegative measure μ on Σ. Conversely, if α_0 is μ-
optimal and the SMD is complete, then α_0 is optimal μ-almost everywhere, that

means $g_{\alpha_0}(u) \leq g_\alpha(u)$, μ-almost everywhere on V for all $\alpha \in A$. See [4] for a proof.

To verify the conditions i - v of this section (section 4), it can be usefull to introduce the spaces B_ω and M_ω.

Let ω be a positive measurable function on V with $\inf_{u \in V} \omega(u) > 0$. The space B_ω is the space of all complex valued measurable functions f on V such that $\frac{f}{\omega} \in B$. With the norm $\| f \|_\omega := \frac{f}{\omega}$ this space is a Banach space.

The space M_ω is the space of all measures μ such that the measure μ_ω defined by

$$\mu_\omega(E) = \int\limits_E \omega(u)\mu(du), \quad E \in \Sigma$$

is an element of M. With the norm $\| \mu \|_\omega := \| \mu_\omega \|$ this space is a Banach space. For an application of this idea to inventory problems, see [4].

The methods described in this section can also be applied to semi-Markovian decision problems. It is sufficient that the average costs can be written as the quotient of the expected recurrence costs and the expected recurrence time to a fixed set A.

5. Countable state space

In this section some results of the preceding one are applied to the case where V is countable and Σ is the σ-field of all subsets of V.

In the next lemma it will be shown that the conditions i), ii), iii), iv), and v), stated in section 4, are implied by some simpler ones.

Lemma 15. Let the following conditions be satisfied.

a) The functions r_α are bounded on V for all $\alpha \in A$ and the boundedness is uniform on A.

b) There is a metric ρ on A such that $P_\alpha(u,v)$ and $r_\alpha(u)$ are continuous in α for all $u,v \in V$. (Instead of $P_\alpha(u,\{v\})$ we write $P_\alpha(u,v)$.)

c) There is a finite subset A of V such that the sum

$$\sum\limits_{n=0}^{\infty} (P_{\alpha B}^n 1_V)(u)$$

with B := V\A, exists for all u $\in$ V, $\alpha \in$ A, and the convergence is uniform on A for all u $\in$ A.

Then the conditions i), ii), iii), iv), and v) are satisfied.

For the proof of this lemma we need the following result.

<u>Lemma 16</u>. Let ρ be a metric on A such that $P_\alpha(u,v)$ is continuous as function on A, for all u,v $\in$ V. Let $\{f_\alpha\}$, $\alpha \in$ A be a set of complex valued functions, bounded on V uniform on A and let $f_\alpha(u)$ be continuous in α for all u $\in$ V. Then $(P_{\alpha G} f_\alpha)(u)$ is continuous in α for all u $\in$ V, G $\in \Sigma$.

<u>Proof</u>. Choose u $\in$ V, G $\in \Sigma$, $\alpha_0 \in$ A. Let $\varepsilon > 0$. There is a finite set F such that $P_{\alpha_0}(u, F_\varepsilon) > 1 - \varepsilon$. The continuity of $P_\alpha(u, F_\varepsilon)$ implies the existence of a $\delta > 0$ such that $P_\alpha(u, V \backslash F_\varepsilon) < 2\varepsilon$ for all $\alpha \in$ A with $\rho(\alpha, \alpha_0) < \delta$. We have

$$(P_{\alpha G} f_\alpha)(u) = \int_{G \backslash F_\varepsilon} P_\alpha(u,ds) f_\alpha(s) + \int_{F_\varepsilon \cap G} P_\alpha(u,ds) f_\alpha(s)$$

and

$$(P_{\alpha G} f_\alpha)(u) - (P_{\alpha_0 G} f_{\alpha_0})(u) = \int_{G \backslash F_\varepsilon} P_\alpha(u,ds) f_\alpha(s) - \int_{G \backslash F_\varepsilon} P_{\alpha_0}(u,ds) f_{\alpha_0}(s) +$$

$$+ \int_{F_\varepsilon \cap G} (P_\alpha(u,ds) - P_{\alpha_0}(u,ds)) f_\alpha(s) + \int_{F_\varepsilon \cap G} P_{\alpha_0}(u,ds)(f_\alpha(s) - f_{\alpha_0}(s)) .$$

The rest of the proof is obvious. $\qquad \square$

Now we can give the proof of lemma 15.

<u>Proof of lemma 15</u>. The conditions i) and iii) are direct consequences of the conditions a) and c), condition ii) is implied by the finiteness of the set A (Q_α is even compact). To prove iv) it is sufficient to prove the continuity of $Q_\alpha(u,E)$ in α for all u $\in$ A, E $\in \Sigma_A$. This is easily done by using the expression

$$Q_\alpha(u,E) = \sum_{n=0}^\infty (P_{\alpha B}^n P_{\alpha A} 1_E)(u) .$$

Namely, condition c) implies that for all $\epsilon > 0$ there is an integer N such that

$$\sum_{n=N}^{\infty} (P_{\alpha B}^n P_{\alpha A} 1_E)(u) < \epsilon$$

for all $u \in A$, $\alpha \in A$, $E \in \Sigma_A$. The continuity of

$$\sum_{n=0}^{N-1} (P_{\alpha B}^n P_{\alpha A} 1_E)(u)$$

in α follows from lemma 16. The rest of the proof of iv) is straightforward. That condition v) is also satisfied can be shown similarly, using the continuity of $r_\alpha(u)$ in α. $\quad\square$

The following theorem is a direct consequence of lemma 15 and theorem 14.

Theorem 17. Let the conditions a), b), and c) of lemma 15 be satisfied and let the SMD be complete and A-communicative. If A is compact then an optimal strategy exists.

For the case of a countable V we shall relate our results to those of some others. Ross [3] and Hordijk [1] investigate the existence of a stationary policy which is average optimal in the class of all policies (Ross) or in the class of all Markov policies (Hordijk) of a Markov decision process. If only the stationary policies are allowed the Markov decision process corresponds to a complete SMD $\{(P_\alpha, r_\alpha)\}$, $\alpha \in A$, where A is the set of all stationary policies. The existence of an optimal strategy of this SMD implies the existence of a stationary policy which is optimal only in the class of all stationary policies. It is important to be consicous of this fact in relating our results to those of Hordijk and Ross.
The conditions of Hordijk, in our terminology, are as follows:

1) the functions $r_\alpha(\cdot)$ are bounded on V, uniform on A;
2) the simultaneous Doeblin condition is satisfied: there is a finite set A, a pos. number c, and an integer n such that $P_\alpha^n(u,A) \geq c$ for all $u \in V$, $\alpha \in A$;
3) there is a metric ρ on A such that A is compact and

4) for all $u, v \in V$ the functions $r_\alpha(u)$ and $P_\alpha(u, v)$ are continuous in α;
5) the SMD is comminicative. (The simultaneous Doeblin condition implies the quasi-compactness of P_α for all $\alpha \in A$, so we may speak indeed about communicativeness.)

The most striking difference with the conditions of theorem 17 is the simultaneous Doeblin condition. Instead of this condition we require condition c) of lemma 15: there is a finite set $A \subset V$ such that the sum

$$\sum_{n=0}^{\infty} (P_{\alpha B}^n 1_V)(u) ,$$

where $B := V \backslash A$, exists for all $u \in V$ and $\alpha \in A$, and the convergence is uniform on A for all $u \in A$.

The simultaneous Doeblin condition implies the convergence of

$$\sum_{n=0}^{\infty} (P_{\alpha B}^n 1_V)(u), \quad \text{uniform on } V \times A .$$

Ross [3] gives the following conditions:

- for all $u \in V$ the set $A(u)$ of all possible actions in u is finite;
- the functions $r_\alpha(\cdot)$ are bounded on V, uniform on A;
- there exists a state $v \in V$, an integer $N > 0$, and a sequence of discount factors $\{\beta_n\}$, $0 < \beta_n < 1$, such that $\lim_{n \to \infty} \beta_n = 1$ and $M_{uv}(R_{\beta_n}) < N$ for all $u \in V$, $n \in \mathbb{N}$, where $M_{uv}(R_{\beta_n})$ is the mean time to go from state u to state v when using the β_n-discounted optimal policy R_{β_n}.

The finiteness of $A(u)$ makes the compactness and continuity conditions superfluous.

The last condition of Ross states a very strong recurrency, (recurrency to a point $v \in V$) for a subset $\{R_{\beta_n}\}$, $n = 1, 2, \ldots$ of the set of all stationary deterministic policies. This condition quarantees the quasi-compactness of the Markov process under policy R_{β_n} and also $R_{\beta_n} \in A_1$ (only one invariant probability). In condition c) of lemma 15 a weaker recurrency is stated (recurrency to a set A), but for all strategies $\alpha \in A$. The A-communicativeness assumed in theorem 17 implies that A_1 dominates A.

In a set of conditions different from the just mentioned one Hordijk [1] section 5, also requires recurrency to a point. This set of conditions is more directly related to the conditions i - v of section 4 of this paper, with V countable and A consisting of one point. The conditions guarantee the continuity of the recurrence costs to A and the recurrence time to A as function of α. The boundedness of r_α and the quasi-compactness of P_α is not required.

References

[1] Hordijk, A. (1974): Dynamic programming and Markov potential theory. Math. Centre Tracts, no. 51, Amsterdam.

[2] De Leve, G. (1964): Generalized Markovian Decision Processes, I and II. Math. Centre Tracts, no. 3 and 4, Amsterdam.

[3] Ross, S.M. (1968): Nondiscounted denumerable Markovian decision models. Ann. Math. Statist. 39, 412-423.

[4] Wijngaard, J. (1975): Stationary Markovian decision problems. Dissertation Technological University, Eindhoven.

[5] Wijngaard, J. (1975): Stationary Markovian decision problems I. Memorandum-COSOR 75-14, Technological University Eindhoven, The Netherlands.