Probability Theory
and Mathematical Statistics
Proceedings of the Fourth USSR – Japan Symposium,
held at Tbilisi, USSR, August 23–29, 1982
Edited by K. Itô and J.V. Prokhorov
Springer-Verlag
Berlin Heidelberg New York Tokyo 1983
STATIONARY AND MARKOV POLICIES IN COUNTABLE STATE

DYNAMIC PROGRAMMING

E.A. Fainberg and I.M. Sonin (Moscow)

1. Introduction. Consider a Markov decision model $M = (X, A(\cdot), p, r)$, where (i) $X$ is a countable state space; (ii) $A(x)$ is the set of actions available in the state $x \in X$ and is assumed to be a measurable subset of some set $A$ endowed with a $\sigma$-field containing all one-point sets; (iii) $p(z|x,a)$ is a transition probability, $p(z|x,a) \geq 0$, $\sum_{z \in X} p(z|x,a) = 1$ for all $x,z \in X$ and $a \in A(x)$; (iv) $r(x,a)$ is a reward function, $-\infty \leq r(x,a) < +\infty$ for all $x \in X$ and $a \in A(x)$. The functions $r(x,a)$ and $p(z|x,a)$ are assumed to be measurable in $a$.

Three sets of policies will be distinguished, namely the set $\Pi$ of all (possibly randomized and history dependent) policies satisfying the usual measurability conditions, the set $M$ of all (nonrandomized) Markov policies and the set $S$ of all (nonrandomized) stationary policies. So $S \subseteq M \subseteq \Pi$. For each policy $\pi \in \Pi$ and each initial state $x \in X$ one may define in the usual way a probability measure $P^\pi_x$ on $\mathcal{H} = (X \times A)^\infty$. Expectations with respect to $P^\pi_x$ are denoted by $\mathbb{E}^\pi_x$. An element $h$ of $\mathcal{H}$ is called a trajectory, $h = x_0 a_0 x_1 a_1 \ldots$

The total expected reward $w^\pi(x)$ when the model starts in $x \in X$ and policy $\pi$ is used can be defined

$$w^\pi(x) = \mathbb{E}^\pi_x \sum_{i=0}^{\infty} r(x_i, a_i),$$

whenever the expectation at the right hand side is well-defined. To guarantee this the following usual assumption is made.

CONDITION 1.1 (general convergence condition). For all $x \in X$ and $\pi \in \Pi$

$$u^\pi(x) = \mathbb{E}^\pi_x \sum_{i=0}^{\infty} r^+(x_i, a_i) < \infty,$$

where for any real-valued function $f$

$$f^+(\cdot) = \max \{0, f(\cdot)\}, \quad f^- (\cdot) = \min \{0, f(\cdot)\}.$$

Denote

$$v(x) = \sup_{\pi \in \Pi} w^\pi(x), \quad s(x) = \sup_{\phi \in S} \sup_{\pi \in \Pi} w^\phi(x), \quad u^{+}(x) = \sup_{\pi \in \Pi} u^\pi(x).$$
For any \( f_1, f_2 \) we write \( f_1 > f_2 \) if \( f_1(\cdot) > f_2(\cdot) \) for all arguments. Similarly, if \( > \) is replaced by \( <, =, \geq \) or by operations \( +, -, \min \).

Denote \( \mathcal{I}(\cdot) \) the indicator function of the set \( \{\cdot\} \subseteq \mathbb{H} \) (the argument \( h \in \mathbb{H} \) is omitted).

Obviously \( s \leq v \leq u^* \). Condition 1.1 implies \( u^* < \infty \) [16, th. 2.3].

For \( f: X \rightarrow [\infty, +\infty) \) define operators

\[
\begin{align*}
p^\phi f(x) &= \sum_{z \in X} p(z|x, a)f(z), \\
P_f(x) &= \sup_{a \in A(x)} p^\phi f(x), \\
T^\phi f(x) &= r(x, a) + p^\phi f(x), \\
T_f(x) &= \sup_{a \in A(x)} T^\phi f(x),
\end{align*}
\]

where it is assumed that \( P^\phi f_0(x) \equiv 0 \) for all \( x \in X \), \( a \in A(x) \). Let \( \Phi \) be the set of all functions \( f: X \rightarrow [0, \infty) \) satisfying this condition. It is easy to see that \( u^* \in \Phi \). Hence \( s^+, v^+ \in \Phi \).

A policy \( \pi \) is called optimal if \( w^\pi = v \). A policy \( \pi \) is called \( g \)-optimal (where \( g \) is a nonnegative function on \( X \)) if \( w^\pi \geq v^g \).

The basic known results on the existence of optimal and nearly optimal policies have been given in an interesting survey by Wal and Wessels [19]. Note that the existence of a stationary optimal policy for a model with finite \( X \) and \( A \) has been proved by Krylov [11] (see also [4, ch. 6, sect. 3]). The more general criterion has been studied by Blackwell [2]. Unfortunately the papers [2, 4, 11] have not been referred to in [19]. The existence of Markov nearly optimal policies for wide classes of criteria has been studied in [5, 6].

We systematically use the following result in the paper.

**Theorem W1** (Wal [16], see also Schäl [13]). If for each \( x \in X \) the set \( A(x) \) is finite then \( s = v \).

The following theorems have been proved in [17, 18].

**Theorem W2** (Wal [17]). If in each state \( x \) for which \( v(x) \leq 0 \) there exists a conserving action (i.e. an action \( a \in A(x) \) such that \( T^\phi v(x) = T v(x) \)), then for any \( \varepsilon > 0 \) there exists a stationary \( \varepsilon u^g \)-optimal policy.

**Theorem W3** (Wal [18]). For any \( \varepsilon > 0 \) there exists a Markov \( \varepsilon (u^g + 1) \)-optimal policy.

2. Results. Let \( f: X \rightarrow [\infty, \infty) \), \( f^+ \in \Phi \), \( X' \subseteq X \). Consider the following three conditions on \( \lambda \in \Phi \):

1. \( \lambda(x) > 0 \), \( \lambda(x) \geq P h(x) \) for \( x \in X \setminus X' \),

2. \( \lambda(x) \geq f(x) \) for \( x \in X \setminus X' \),

3. \( \lambda(x) = 0 \) for \( x \in X' \).
The set of functions \( \mathcal{L} \in \Phi \) satisfying conditions (i, ii) will be denoted \( L(f,X') \). Denote \( L(f) = L(f,\Phi) \).

The set of functions \( \mathcal{L} \in \Phi \) satisfying conditions (i - iii) will be denoted \( L_0(f,X') \), \( L_0(f,X') \subseteq L(f,X') \).

The set of functions \( \mathcal{L} \in \Phi \) satisfying condition (i) will be denoted \( \mathcal{K}(X') \), \( \mathcal{K}(X') \supseteq L(f,X') \).

Then (see lemmas 3.1 and 3.2) for all \( X' \subseteq X'' \subseteq X \),

\[
\mathcal{L} + 1 \in L(v) \subseteq L(s) \subseteq L(s,X') \subseteq L(s,X'').
\]

Denote

\[
Z = \{ x \in X : \phi^x(x) = s(x) \text{ for some } \phi^x \in \Phi \},
\]

\[
X^{opt} = \{ x \in X : \pi^x(x) = v(x) \text{ for some } \pi^x \in \Pi \}.
\]

Obviously there exists \( \pi \in \Pi \) such that \( \pi^x(x) = v(x) \) for all \( x \in X^{opt} \). Now we formulate our main result.

**Theorem 2.1.** For any \( \varepsilon > 0 \) and any \( \lambda \in L(s,Z) \) there exists a stationary policy \( \phi \) such that

\[
\phi^x \geq s - \varepsilon L.
\]

Since \( L_0(s,Z) \subseteq L(s,Z) \), the theorem holds for functions \( \lambda \in L_0(s,Z) \).

Further if \( \lambda \in L(s,Z) \), then \( \lambda_0 \in L_0(s,Z) \), where \( \lambda_0(x) = \lambda(x) \) for \( x \in X \setminus Z \) and \( \lambda_0(x) = 0 \) for \( x \in Z \). Hence for any \( \varepsilon > 0 \) and \( \lambda \in L(s,Z) \), the stationary policy \( \phi \) satisfying (2.1) can be chosen so that \( \phi^x(x) = s(x) \) for \( x \in Z \).

The proof of theorem 2.1 consists of three stages. At the first stage theorem 2.1 is proved for \( \lambda \in L(s) \). At the second stage we prove the following result.

**Theorem 2.2.** The function \( s \) satisfies the optimality equation

\[
s = Ts.
\]

At the third stage we prove theorem 2.1 for functions \( \lambda \in L(s,Z) \).

Now we will formulate some corollaries from theorem 2.1. Obviously if \( \lambda_1, \lambda_2 \in L(f,X') \), then \( \lambda = \min(\lambda_1, \lambda_2) \in L(f,X') \). And if \( f \leq K \), where \( K \) is some finite positive constant, then \( \lambda = K \in L(f,X') \). Since \( \varepsilon > 0 \) is arbitrary, theorem 2.1 implies the following result.

**Corollary 2.3** (cp. example in [3]). If \( s \) is bounded from above then for
any \( \varepsilon > 0 \) and any \( \lambda \in L(s, Z) \) there exists a stationary policy \( \phi \) such that
\[
\psi^\lambda \geq s - \varepsilon \min(\lambda, 1)
\]
and, consequently, \( \psi^\lambda \geq s - \varepsilon \).

There are known various sufficient conditions of the existence of stationary \( g \)-optimal (\( g \)-optimal) policies for various concrete functions \( g \). Theorem 2.1 gives the general method of proof of the existence of such policies, namely, it is sufficient to prove that \( s = v \) and \( g \in L(s, Z) \). Using this method one can prove and generalize some known results and prove some new ones. Here we give only two results (corollaries 2.4 and 2.5).

**COROLLARY 2.4** (Ornstein [12], see also [7], [10, th. 13.7]). If \( \tau > 0 \) then for any \( \varepsilon > 0 \) there exists a stationary \( \varepsilon \)-optimal policy.

**PROOF.** In [3] it is proved that if \( \tau > 0 \) then \( s = v \). Denote \( X_0 = \{ x \in X : v(x) = 0 \} \). Then \( \psi^\tau(x) = v(x) = 0 \) for all \( x \in X_0 \) and \( \tau \in \Pi \).

Consequently \( X_0 \subseteq Z \). But \( v = u^\varepsilon \in L(v, X_0) \subseteq L(v, X) \) (Lemma 3.2 (ii)). Hence by lemma 3.1 (ii) \( v \in L(v, Z) \) and we can use theorem 2.1.

**COROLLARY 2.5** (cp. theorem W2). If conditions of theorem W2 are fulfilled, then for any \( \varepsilon > 0 \) and any \( \lambda \in L(v, X^\varepsilon) \), where \( X^\varepsilon = \{ x \in X : u^\varepsilon(x) = 0 \} \), there exists a stationary \( \varepsilon \)-optimal policy.

**PROOF.** By theorem W2, \( s = v \). So to prove corollary 2.5 it is sufficient to prove \( X^\varepsilon \subseteq Z \). Since \( u^\varepsilon \geq 0 \), then \( r(xa) \leq 0 \) and \( p(x \mid x, a) = 0 \) for \( x \in X^\varepsilon \), \( z \in XX^\varepsilon \) and \( a \in A(x) \). Next consider the model \( \tilde{M} = (\bar{x}, \lambda(\cdot), p^\tau, \bar{r}) \) where \( \bar{r} = X^\varepsilon \). Let \( \tilde{v} \) be the function \( v \) for this model. Then \( \tilde{v}(x) = v(x) \) for \( x \in X^\varepsilon \) and there is a conserving action in each state \( x \in X^\varepsilon \). So in the model \( \tilde{M} \) there exists a stationary optimal policy ([15], the proof of theorem 9.1, or theorem W2 for \( u^\varepsilon = 0 \)). Consequently in the model \( M \) there exists a stationary policy \( \psi \) such that \( \psi^\lambda(x) = v(x) \) for all \( x \in X^\varepsilon \). Thus \( X^\varepsilon \subseteq Z \).

Note that since \( u^\varepsilon \in L(v, X^\varepsilon) \) (lemma 3.2 (ii)) then corollary 2.5 extends theorem W2. There are simple examples when the function \( u^\varepsilon \) is unbounded but \( v \leq K \infty \). In such examples
\[
\kappa \in L(v, X^\varepsilon) \quad \text{and} \quad \min(K, u^\varepsilon) \in L(v, X^\varepsilon).
\]

For Markov policies there is the following analog of theorem 2.1.

**THEOREM 2.6** (cp. theorem W3). For any \( \varepsilon > 0 \) and any \( \lambda \in L(v, X^\varepsilon) \) there exists a Markov \( \varepsilon \)-optimal policy.

For functions \( \kappa \in L(v) \) one can obtain the proof of theorem 2.6 using the outline of Wai's proof [18] with some modifications. (The first essential modification is that the right hand side of the inequality proved in lemma 2.1 of [18] must be \( v^\varepsilon - \varepsilon \bar{Z}^\varepsilon \) instead of \( v^\varepsilon - \varepsilon \bar{Z} \)). The second one is the use of corollary 2.5 instead of theorem W2.) When the theorem is proved for \( \kappa \in L(v) \),
the proof for $\pi \in L(v, x^{opt})$ is similar to the third stage of the proof of theorem 2.1. There is also another proof of this theorem which is not based on theorem W2.

If for any $x, z \in X$ and $a \in A(x)$ either $p(z|x,a) = 1$ or $p(z|x,a) = 0$ then the model is called deterministic. The deterministic model is a particular case of the original model. Hence all the results formulated above are valid for deterministic models. But for deterministic models the stronger results are fulfilled. Bertsekas and Shreve [1, th. 2] have proved that if in the deterministic model $r \geq 0$, then for any $\epsilon > 0$ there exists a stationary $\epsilon$-optimal policy (this proposition is wrong if there is no assumption that the model is deterministic).

As it is shown in section 5 all the results formulated above hold for deterministic models not only for functions $l \in L(s, z)$ (or $L(v, x^k)$, $L(v, x^{opt})$) but also for functions $l \in K(z)$ ($K(z^k)$, $K(x^{opt})$ respectively). So, for example, the following analog of corollary 2.4 is valid.

**COROLLARY 2.7** (cp. Bertsekas and Shreve [1, th. 2]). If the model is deterministic and if $r \geq 0$ then for any $\epsilon > 0$ there exists a stationary $\epsilon \min(v, l)$-optimal policy.

3. Preliminary results.

**LEMMA 3.1.** Let $f_1, f_2 \in \Phi$, then

(i) if $f_1 \geq f_2$, then $L(f_1) \leq L(f_2)$;

(ii) if $X' \subseteq X' \subseteq X$, then $L(f) \leq L(f, X') \leq L(f, X''$);

(iii) if $l_1, l_2 \in L(f, X')$, then $l_1 + l_2 \in L(f, X')$, $\min(l_1, l_2) \in L(f, X')$.

**LEMMA 3.2.** (i) $u^* + 1 \in L(v)$;

(ii) $u^* \in L_0(v, x^k)$, where $x^k = \{x \in X: u^*(x) = 0\}$.

**PROOF.** Routine.

Let $Q$ be the set of all Markov times $\tau = \tau(h) \leq \infty$. For $f: X \rightarrow [-\infty, \infty)$, $\tau \in Q$, $x \in X$ and $\pi \in \Pi$ we write $E^\pi_X f(x_\tau)$ instead of $E^\pi_X \mathbf{1}_{\{\tau = \infty\}} f(x_\tau)$.

**LEMMA 3.3** (Shiryayev [14, ch. II, lemma 3]). For any $\pi \in \Pi$, $\tau \in Q$ and any nonnegative excessive function $f$ (i.e., $f \in \Phi$, $f \geq Pf$)

$$f(x) \geq E^\pi_X f(x_\tau), \quad x \in X.$$ 

For $\pi \in \Pi$, $f: X \rightarrow [-\infty, \infty)$ and $\tau \in Q$ denote

$$u^\pi(x, \tau, f) = E^\pi_X \left[ \sum_{i=0}^{\tau-1} r^+(x_{i}, a_i) + f^+(x_{\tau}) \right]$$
and if \( u^\pi(x, \tau, f) \leq \), define

\[
u^\pi(x, \tau, f) = E_X^\pi \left( \sum_{i=0}^{\infty} r(x_i, a_i) + f(x_i) \right).
\]

The following lemma provides the correctness of using the function \( \nu^\pi(x, \tau, f) \) below.

**Lemma 3.4.** For any \( x \in X, \pi \in \Pi \) and any \( \tau \in Q \)

\[ 0 \leq u^\pi(x, \tau, 0) \leq u^\pi(x, \tau, \sigma) \leq u^\pi(x, \tau, \nu) \leq u^\pi(x, \tau, u^\pi) < \infty. \]

**Proof.** Let \( \varepsilon > 0 \). Consider a policy \( \sigma \) such that \( u^\sigma \geq u^\pi - \varepsilon \).

Define the policy \( \gamma \) by

\[
\gamma_i(\cdot|x_0 a_0 \ldots x_i) = \begin{cases} 
\pi_i(\cdot|x_0 a_0 \ldots x_i) & \text{for } i < \tau, \\
\sigma_{i-\tau}(\cdot|x_\tau a_\tau \ldots x_i) & \text{for } i \geq \tau.
\end{cases}
\]

Since \( \gamma \in \Pi \), then \( u^\gamma(x) < \infty \). Further we have

\[ u^\pi(x, \tau, u^\pi) \leq u^\pi(x, \tau, u^\gamma + \varepsilon) \leq u^\pi(x, \tau, u^\sigma) + \varepsilon = u^\gamma(x) + \varepsilon < \infty. \]

**Lemma 3.5.** Let \( \tau_n \in Q, n = 1, 2, \ldots, \) and \( \tau_n(h) \nrightarrow \) as \( n \to \infty \) for all \( h \in H \). Then for any \( \pi \in \Pi \),

\[
\nu^\pi(x) = \lim_{n \to \infty} \nu^\pi(x, \tau_n, 0), \ x \in X. \quad (3.1)
\]

**Proof.** For any \( h \in H \)

\[
\nu^\pi(x, \tau_n, 0) \leq \sum_{i=0}^{\tau_n-1} r(x_i, a_i) + \sum_{i=0}^{\infty} r(x_i, a_i).
\]

Consequently \( u^\pi(x, \tau_n, 0) \leq u^\pi(x) \). Similarly one can consider sums of \( r^- \). In view of condition 1.1 we obtain (3.1).

**Lemma 3.6.** Let \( \tau_n \in Q, n = 1, 2, \ldots, \) and \( \tau_n(h) \nrightarrow \) as \( n \to \infty \) for all \( h \in H \). If for some \( \pi \in \Pi \), \( x \in X \) and some function \( f: X \to [-\infty, +\infty) \) the following conditions hold

(a) \( \lim_{n \to \infty} \inf \nu^\pi(x, \tau_n, f) \geq f(x) - \delta_1 \),

(b) \( \lim_{n \to \infty} \inf \nu^\pi(x, \tau_n, f) \leq \delta_2 \),
then
\[ w^\pi(x) \geq f(x) - (\delta_1 + \delta_2). \]  
(3.2)

**PROOF.** If \( f(x) = -\infty \), then (3.2) is obvious. If \( f(x) > -\infty \), then using lemma 3.5 and conditions \((a, b)\) we have
\[
w^\pi(x) = \lim_{n \to \infty} w^\pi(x, \tau_n, 0) = \lim_{n \to \infty} \left( w^\pi(x, \tau_n, f) - E^{\pi}_x f(x_{\tau_n}) \right) \geq \]
\[ \geq \lim_{n \to \infty} \inf \ w^\pi(x, \tau_n, f) - \lim_{n \to \infty} \inf \ E^{\pi}_x f(x_{\tau_n}) \geq f(x) - (\delta_1 + \delta_2). \]

**LEMMA 3.7.** (See [8].) For any \( x \in X \) and any \( \varepsilon > 0 \) there exists a Markov policy \( \sigma \) such that \( w^\sigma(x) \geq v(x) - \varepsilon \).

**LEMMA 3.8.** Let \( Y \) be a finite subset of \( X \). Then for any \( \varepsilon > 0 \) there exists a stationary policy \( \phi \) such that
\[ w^\phi(y) \geq s(y) - \varepsilon, \quad y \in Y. \]

**PROOF.** Let \( Y = \{y^1, y^2, \ldots, y^n\} \). Consider stationary policies \( \phi_i \), \( i = 1, 2, \ldots, n \), such that
\[ w^\phi_i(y^i) \geq s(y^i) - \varepsilon/2, \quad i = 1, 2, \ldots, n. \]
(3.3)

Introduce a point \( z \notin X \) and for some \( b \in A \) denote \( r(z, b) = 0 \), \( p(y^i | z, b) = 1/n, \quad i = 1, 2, \ldots, n \). Consider the model \( \tilde{M} = \{\tilde{X}, \tilde{A}(*), p, r\} \), where \( \tilde{X} = X \cup \{z\} \), \( \tilde{A}(x) = \{\phi_i(x), i = 1, \ldots, n\} \) for \( x \in X \) and \( \tilde{A}(z) = \{b\} \). Denote \( \tilde{v} \) the function \( v \) for this model. By (3.3) and the definition of the model \( \tilde{M} \),
\[ \tilde{v}(y^i) \geq w^\phi_i(y^i) \geq s(y^i) - \varepsilon/2, \quad i = 1, \ldots, n. \]

Since the sets \( \tilde{A}(\cdot) \) are finite, theorem W1 implies the existence of a stationary policy \( \tilde{\phi} \) such that \( w^\tilde{\phi}(z) \geq \tilde{v}(z) - \varepsilon/2n \). Obviously
\[ w^\tilde{\phi} = \frac{1}{n} \sum_{i=1}^{n} w^\phi_i(y^i), \quad \tilde{v}(z) = \frac{1}{n} \sum_{i=1}^{n} \tilde{v}(y^i). \]

Hence
\[ \frac{1}{n} \sum_{i=1}^{n} (w^\phi_i(y^i) - \tilde{v}(y^i)) \geq -\varepsilon/2n. \]
Since $w^\phi(y_i) \leq \tilde{w}(y_i)$, $i = 1, \ldots, n$, then

$$w^\phi(y_i) \geq \tilde{w}(y_i) - \varepsilon/2 \geq s(y_i) - \varepsilon, \ i = 1, \ldots, n.$$ 

Omitting $z$ one can consider $\phi$ as an element of $S$. \qed

A nonrandomized policy $\pi$ is called finite if for each $x \in X$ there exists a finite subset $B(x)$ of $A(x)$ such that $\pi(x_0, a_0, \ldots, x_i) \in B(x_i)$ for all $i = 0, 1, \ldots$.

**Lemma 3.9.** If $\pi$ is a finite policy, then $w^\pi \leq s$.

**Proof.** Consider the model $\tilde{M} = (X, B(\cdot), p, r)$. Let $\tilde{v}$ and $\tilde{s}$ be the functions $v$ and $s$ for this model. Obviously $\tilde{v} \geq \tilde{w}^\pi$. Theorem 3.1 implies $\tilde{u} = \tilde{v}$. Since $A(\cdot) \supseteq B(\cdot)$, then $s \geq \tilde{s}$. Thus $s \geq \tilde{s} = \tilde{v} \geq \tilde{w}^\pi$. \qed

**Lemma 3.10.** For any $\phi \in S$ and any $\tau \in Q$,

$$w^\phi(x) \leq w^\phi(x, \tau, s), \ x \in X.$$ 

**Proof.** $w^\phi(x) = w^\phi(x, \tau, w^\phi) \leq w^\phi(x, \tau, s)$. \qed

**Lemma 3.11.** (cp. lemma 4.3). Let $x \in X$, $\phi \in S$ and $w^\phi(x) > -\infty$. Let $R \subseteq X \times \{0, 1, 2, \ldots\}$ and $\tau$ be the first hitting time of $R$, $\tau = \min\{i: (x_i, 1) \in R\}$. Then

$$w^\phi(x, \tau, s) \leq s(x).$$

**Proof.** By lemma 3.4, $E^\phi_x s(x_i) < \infty$. The assumption $w^\phi(x) > -\infty$ implies $E^\phi_x s(x_i) > -\infty$. Therefore there exists a finite subset $R'$ of $R$ such that

$$E^\phi_x I(x_i, \tau) \in R' \ s(x_i) > E^\phi_x s(x_i) - \varepsilon/3, \quad (3.4)$$

$$E^\phi_x I(x_i, \tau) \in R \setminus R' \ w^\phi(x_i) \geq -\varepsilon/3. \quad (3.5)$$

Let $D = \{x \in X: (x, 1) \in R' \text{ for some } i = 0, 1, \ldots\}$. Since $D$ is finite, lemma 3.8 implies the existence of $\psi \in S$ such that

$$w^\psi(y) \geq s(y) - \varepsilon/3, \ y \in D.$$ 

Define a nonrandomized policy $\tilde{\pi}$ by
\[
\pi_i(x_0, a_0, \ldots, x_i) = \begin{cases} 
\phi(x_i) & \text{for } i < \tau \text{ and for } i \geq \tau, x_i \in X \setminus D, \\
\psi(x_i) & \text{for } i \geq \tau, x_i \in D.
\end{cases}
\]

Since the policy \( \pi \) is finite, lemma 3.9 implies \( w^\pi \leq s \). Using the definitions of \( \pi \) and \( \psi \), inequalities (3.4), (3.5) and the identity \( \{x_\tau \in D\} = \{- (x_\tau, \tau) \in E'\} \), one can write

\[
s(x) \geq w^\pi(x) = w^\pi(x_\tau, 0) + \mathbb{E}_x^\pi [I(x_\tau \in X \setminus D \cup \mathcal{A}^\phi(x_i) + E_x^\phi(I(x_\tau \in D \cup \mathcal{A}^\psi(x_i) \geq \\
\geq w^\phi(x_\tau, 0) - \epsilon / 3 + \mathbb{E}_x^\phi[I(x_\tau \in D \cup \mathcal{A}^\phi(s(x_\tau) - \epsilon / 3) \geq w^\phi(x_\tau, 0) + \mathbb{E}_x^\phi(s(x_\tau) - \epsilon = \\
= w^\phi(x_\tau, s) - \epsilon.
\]

Since \( \epsilon > 0 \) is arbitrary, the lemma is proved. \( \square \)

**Lemma 3.12.** Let \( Y \) be a finite subset of the set \( \{x \in X: s(x) > -\infty\} \).

Then for any \( \epsilon > 0 \) and any \( N < \infty \) there exist a stationary policy \( \phi \) and an integer \( n > N \) such that

(a) \( w^\phi(y) \geq s(y) - \epsilon, \quad y \in Y \),

(b) \( E_y^\phi s^+(x_n) < \epsilon, \quad y \in Y \).

**Proof.** By lemma 3.8 there exists a stationary policy \( \phi \) such that

\[
w^\phi(y) \geq s(y) - \epsilon / 3, \quad y \in Y.
\]  \hspace{1cm} (3.6)

Let \( n > N \) be an integer such that for all \( y \in Y \)

\[
E_y^\phi \sum_{i=n}^\infty r(x_i, a_i) \leq \epsilon / 3, \quad E_y^\phi \sum_{i=n}^\infty r(x_i, a_i) \geq - \epsilon / 3.
\]  \hspace{1cm} (3.7)

By inequality (3.6) and the first of inequalities (3.7) we have

\[
w^\phi(y, n, 0) = w^\phi(y) - E_y^\phi \sum_{i=n}^\infty r(x_i, a_i) \geq \\
\geq w^\phi(y) - E_y^\phi \sum_{i=n}^\infty r^+(x_i, a_i) \geq s(y) - 2\epsilon / 3.
\]

for all \( y \in Y \). By lemma 3.11 (for \( \tau \in n, R = X \times \{n\} \))
\[ s(y) \geq w^\phi(y, n, s) = w^\phi(y, n, 0) + E^\phi_y s(x_n) \]

for all \( y \in X \). So, with this inequality and (3.8)

\[ E^\phi_y s(x_n) = E^\phi_y s^+(x_n) + E^\phi_y s^-(x_n) \leq 2\varepsilon/3 \]

(3.9)

for all \( y \in Y \). Since

\[ E^\phi_y s^-(x_n) = E^\phi_y \min(0, s(x_n)) \geq E^\phi_y \min(0, \psi^\phi(x_n)) \geq E^\phi_y \sum_{i=n}^\infty r^-(x_i, a_i) \]

using the second of inequalities (3.7) and inequality (3.9) one can obtain (b). □

4. The proofs of theorems 2.1 and 2.2. At first theorem 2.1 will be proved for functions \( \lambda \in L(s) \subseteq L(s, Z) \), i.e. for positive excessive majorants of \( s \).

The proof of theorem 2.1 for such \( \lambda \) is based on lemmas 4.1 and 4.2.

**Lemma 4.1.** Let \( \varepsilon > 0 \), \( \lambda \in L(s) \), \( N < \infty \) and \( Y \) be a finite subset of the set \( \{ x \in X : s(x) > -\infty \} \). Then there exists a stationary policy \( \psi \), an integer \( n > N \) and a finite set \( D \), \( D \supseteq Y \), such that

(a) \( \psi(z) \geq s(z) - \varepsilon\lambda(z) \), \( z \in D \),

(b) \( E^\phi_y s^+(x_n) \leq \varepsilon \), \( y \in Y \),

where \( \tau = \min\{n, \bar{\tau}\} \) and \( \bar{\tau} = \min\{k : x_k \in X \backslash D\} \) is the first exit time from \( D \).

**Proof.** Denote \( \lambda = \min\{1, \min(\lambda(y) : y \in Y)\} \). The set \( Y \) is finite and \( \lambda > 0 \). Hence \( \lambda > 0 \). Let \( \varepsilon < 1 \) and \( \varepsilon' = \varepsilon^2/3 \leq \varepsilon^2/3 < \varepsilon/3 \). Using lemma 3.12 one can select a stationary policy \( \psi \) and an integer \( n > N \) such that

\[ \psi(y) \geq s(y) - \varepsilon', \quad y \in Y; \]

\[ E^\phi_y s^+(x_n) \leq \varepsilon', \quad y \in Y. \]

(4.1)

(4.2)

Define \( G = \{ z : \psi^\phi(z) < s(z) - \varepsilon\lambda(z) \} \) and show that

\[ E^\phi_y s^+(x_0) \leq \varepsilon'/\kappa, \quad y \in Y, \]

(4.3)

where \( \kappa = \min\{k : x_k \in G\} \).

Using the definition of \( G \) one can write
\[ w^*(y) = w^*(y, \rho, 0) + E_y^* w^*(x_\rho) \leq \psi^*(y, \rho, 0) + E_y^* s(x_\rho) - \epsilon E_y^* \lambda(x_\rho). \] (4.4)

By lemma 3.11,
\[ \psi^*(y, \rho, s) - w^*(y, \rho, 0) + E_y^* s(x_\rho) \leq s(y). \] (4.5)

Using (4.1), (4.4) and (4.5) we have
\[ s(y) - \epsilon' \leq w^*(y) \leq s(y) - \epsilon E_y^* \lambda(x_\rho), \quad y \in Y. \]

Hence
\[ E_y^* \lambda(x_\rho) \leq \epsilon' / \epsilon. \]

Since \( \epsilon \in L(s) \), then \( \lambda \geq s \) and \( \lambda > 0 \). So \( \lambda \geq s^+ \) and consequently (4.3) is proved.

Now construct the set \( \mathcal{D} \). Define \( \mathcal{D}_1 = Y \). From (a) and definition of \( G \) we have \( \mathcal{D}_1 \cap G = \phi \). If for some \( i = 2, \ldots, n-1 \) the finite set \( \mathcal{D}_{i-1} \) if constructed and \( \mathcal{D}_{i-1} \cap G = \phi \), then choose a finite set \( \mathcal{D}_i \) so that \( \mathcal{D}_i \supset \mathcal{D}_{i-1} \), \( \mathcal{D}_i \cap G = \phi \) and for any \( y \in \mathcal{D}_{i-1} \) the following inequality is fulfilled:
\[ E_y^* \{ x_\tau \notin \mathcal{D}_i \cup G \} s^+(x_\tau) < \epsilon / 3n. \] (4.6)

Define \( \mathcal{D} = \bigcup_{i=1}^{n-1} \mathcal{D}_i \) and \( \tau = \min \{ n, \tau \} \), where \( \tau = \min \{ \kappa : x_\kappa \in X \setminus \mathcal{D} \} \).

By inequality (4.6),
\[ E_y^* \{ \tau < n, x_\tau \notin \mathcal{D} \cup G \} s^+(x_\tau) < \epsilon / 3, \quad y \in Y. \] (4.7)

Check now that inequalities (a), (b) of lemma 4.1 hold.

Inequality (a) is fulfilled because \( \mathcal{D} \cap G = \phi \) and \( \epsilon' < \epsilon \lambda(y) \) for all \( y \in Y \). For checking (b) we write
\[ E_y^* s^+(x_\tau) = E_y^* \{ \tau = n \} s^+(x_\tau) + \]
\[ + E_y^* \{ \tau < n, x_\tau \notin G \} s^+(x_\tau) + E_y^* \{ \tau < n, x_\tau \in X \setminus G \} s^+(x_\tau). \] (4.8)

Note that \( s^+ \geq 0 \). Using inequalities (4.2), (4.3) and (4.7) for each of expectations in the right hand side of (4.8) we have
$E^\phi_y s^\phi(x) \leq c' + \varepsilon/c + \varepsilon/3 \leq \varepsilon$,

i.e. inequality (b) of lemma 4.1 is fulfilled too. \hfill \Box

For $\phi \in S$ and $Y \subseteq X$ define

$S(\phi, Y) = \{\psi \in S: \psi(x) = \phi(x), x \in Y\},$

$s^{\phi,Y}(x) = \sup\{\psi(x): \psi \in S(\phi, Y)\}, x \in X.$

**Lemma 4.2.** Let a set $Y, Y \subseteq X$ a constant $\varepsilon > 0$, a positive excessive functions $\xi$ and a stationary policy $\phi$ be such that

$w^\phi\gamma (y) \geq s(y) - \varepsilon\xi(y), y \in Y.$

Then

$s^{\phi,Y}(x) \geq s(x) - \varepsilon\xi(x), x \in X.$

**Proof.** Fix some $x$ from $X$. For any $\delta > 0$ there exists a stationary policy $\psi$ such that

$w^\psi(x) \geq s(x) - \delta.$ \hfill (4.9)

Denote $\tau = \min(\kappa: x_\kappa \in Y)$ and define a nonrandomized policy $\pi$ by

$\pi(x_0, a_0, ..., x_\tau) = \begin{cases} 
\psi(x_i) & \text{for } i < \tau, \\
\phi(x_i) & \text{for } i \geq \tau.
\end{cases}$

Then

$w^\pi(x) = w^\psi(x, \tau, 0) + E_x^\psi w^\psi(x_\tau) \geq w^\psi(x, \tau, 0) + E_x^\psi s(x_\tau) - \varepsilon E_x^\psi \xi(x_\tau) \geq$

$\geq w^\psi(x, \tau, 0) + E_x^\psi w^\psi(x_\tau) - \varepsilon E_x^\psi \xi(x_\tau) = w^\psi(x) - \varepsilon E_x^\psi \xi(x_\tau).$

Applying lemma 3.3 and inequality (4.9) we have

$w^\pi(x) \geq s(x) - \delta - \varepsilon\xi(x).$
Consider the model \( \hat{M} = (X, A(\cdot), p, \pi) \), where \( A(x) = \{ \phi(x) \} \) for \( x \in Y \) and \( \hat{A}(x) = A(x) \) for \( x \in X \setminus Y \). The policy \( \pi \) is finite. Applying \( \text{lemma 3.9} \) to the policy \( \pi \) and model \( \hat{M} \) we have

\[
\hat{s}_1 \geq \eta(x) \geq s(x) - \delta - \varepsilon \ominus(x) .
\]

Since \( \delta > 0 \) is arbitrary, \( \text{lemma 4.2} \) is proved. \( \square \)

**THE PROOF OF THEOREM 2.1 FOR FUNCTIONS \( \tilde{\phi} \in L(s) \).** Let \( \varepsilon > 0 \), \( \lambda \in L(s) \), \( X = (x_1, x_2, \ldots) \) and \( \varepsilon \), \( i = 1, 2, \ldots \), be a sequence of positive numbers such that

\[
\sum_{i=1}^{\infty} \varepsilon_i < \varepsilon/2 .
\]

Define \( \hat{Y}_1 = \{ x_1 \} \) and using \( \text{lemma 4.1} \) choose \( \phi_1, \eta_1, \Omega_1 \) such that

\[
\phi_1(x) \geq s(x) - \varepsilon_1 \lambda(x) , \quad x \in \Omega_1 ,
\]

\[
\hat{s}_1(x) \geq \hat{s}(x) - \varepsilon_1 , \quad y \in \hat{Y}_1 ,
\]

where \( \tau_1 = \min \{ n_1, \bar{\tau}_1 \} \), \( \hat{\tau}_1 = \min \{ \kappa : x_\kappa \in \hat{D}_1 \} \).

Let for some \( \kappa = 1, 2, \ldots \) the set \( \hat{D}_\kappa \) and stationary policy \( \hat{\phi}_\kappa \) be constructed and the integer \( n_\kappa \) be determined. Consider the model \( M_\kappa = (X, A_\kappa(\cdot), p, \tau) \), where \( A_\kappa(x) = \{ \phi_\kappa(x) \} \) for \( x \in \hat{D}_\kappa \) and \( A_\kappa(x) = A(x) \) for \( x \in X \setminus \hat{D}_\kappa \).

Let \( s_\kappa(x) = \phi_\kappa(x) \), \( \hat{D}_\kappa \) and let \( \lambda_\kappa \) be the set of all positive excessive majorants of the function \( s_\kappa \) in the model \( M_\kappa \). Since \( \lambda_\kappa \subseteq \lambda_\kappa(\cdot) \), then \( s_\kappa \leq s \) and \( \lambda_\kappa \subseteq \lambda(s) \). Thus \( \lambda \in \lambda_\kappa \).

Define \( Y_{\kappa+1} = \hat{D}_{\kappa+1} \cup \{ x_{\kappa+1} \} \). Applying \( \text{lemma 4.1} \) to the model \( M_\kappa \), set \( Y_{\kappa+1} \), constants \( \varepsilon_{\kappa+1} > 0 \), \( n_\kappa \) and function \( \lambda \) one can see the existence of a stationary policy \( \phi_{\kappa+1} \), an integer \( n_{\kappa+1} > n_\kappa \) and a set \( \hat{D}_{\kappa+1} \), \( \hat{D}_{\kappa+1} \subseteq \hat{D}_{\kappa+1} \), such that

\[
\phi_{\kappa+1}(x) \geq s_{\kappa}(x) - \varepsilon_{\kappa+1} \lambda(x) , \quad x \in \hat{D}_{\kappa+1} ,
\]

\[
\hat{s}_{\kappa+1}(x) \geq \hat{s}(x) - \varepsilon_{\kappa+1} , \quad x \in Y_{\kappa+1} .
\]

where \( \tau_{\kappa+1} = \min \{ n_{\kappa+1}, \bar{\tau}_{\kappa+1} \} \), \( \bar{\tau}_{\kappa+1} = \min \{ i : x_i \in \hat{X} \setminus \hat{D}_{\kappa+1} \} \).

Note that the definition of the model \( M_\kappa \) implies \( \phi_{\kappa+1}(x) = \phi_\kappa(x) \) for
\[ x \in D_K. \]

By induction in \( k = 1, 2, \ldots \) we can prove

\[ s_K(x) \geq s(x) - \left( \varepsilon_1 + \ldots + \varepsilon_K \right) \ell(x), \quad x \in X. \] (4.13)

If \( k = 1 \) then (4.10) and lemma 4.2, applied in the model \( \mathcal{M} \) to the set \( D_1 \), policy \( \phi_1 \), constant \( \varepsilon_1 \) and function \( \ell \), imply inequality (4.13). Let for some \( k = 1, 2, \ldots \) inequality (4.13) be fulfilled. Then by inequality (4.11) and lemma 4.2, applied in the model \( \mathcal{M}_k \) to the set \( D_{k+1} \), policy \( \phi_{k+1} \), constant \( \varepsilon_{k+1} \) and function \( \ell \),

\[ s_{k+1}(x) \geq s_k(x) - \varepsilon_{k+1} \ell(x) \geq s(x) - \left( \varepsilon_1 + \ldots + \varepsilon_{k+1} \right) \ell(x). \]

Define the stationary policy \( \phi : \phi(x) = \phi_k(x) \) for \( x \in D_k \), \( k = 1, 2, \ldots \). This definition is correct because \( \phi_{k+1}(x) = \phi_k(x) \) for \( x \in D_k \), \( D_k \subseteq D_{k+1} \), \( k = 1, 2, \ldots \), and \( \bigcup_{k=1}^{\infty} D_k = X \).

Fix some \( x \in X \) and show that

\[ \omega^\phi(x) \geq s(x) - \varepsilon \ell(x). \] (4.14)

In order to prove (4.14) we check the fulfilment of conditions (a, b) from lemma 3.6 for \( \delta_1 = \delta_2 = \varepsilon \ell(x)/2 \) and for the sequence of Markov times \( \tau_k \) defined above. From \( D_k \subseteq D_{k+1} \), \( n_{k+1} > n_K \), \( k = 1, 2, \ldots \) and \( X = \bigcup_{k=1}^{\infty} D_k \) we have \( \tau_k(h) \leq \tau_h \) when \( k \rightarrow \infty \) for all \( h \in \mathcal{H} \). We can choose \( i \) such that \( x \in D_1 \).

Using in turn the definition of \( \phi \), lemma 4.10 and inequalities (4.11), (4.13), we have for \( k > i \)

\[ \omega^\phi(x, \tau_{k+1}, s) = \omega^\phi_{k+1}(x, \tau_{k+1}, s) \geq \omega^\phi_{k+1}(x) \geq s(x) - \left( \varepsilon_1 + \ldots + \varepsilon_{k+1} \right) \ell(x) \geq s(x) - \varepsilon \ell(x)/2. \]

So condition (a) of lemma 3.6 holds.

For \( k > i \) by (4.13), (4.14) and lemma 3.3 we have

\[ E_x^\phi s^+(x, \tau_{k+1}) \leq E_x^\phi s_k(x, \tau_{k+1}) + \left( \varepsilon_1 + \ldots + \varepsilon_k \right) \ell(x) \leq E_x^\phi s^+(x, \tau_{k+1}) \leq E_x^\phi s^+(x) + \left( \varepsilon_1 + \ldots + \varepsilon_k \right) \ell(x) \leq \varepsilon_{k+1} + \left( \varepsilon_1 + \ldots + \varepsilon_k \right) \ell(x). \]
Consequently

\[ \lim \inf_{\kappa \to \infty} \mathbb{E}_{x}^{\phi} s^{\phi}(x_{\kappa}) \leq \varepsilon \ell(x)/2 \]

and condition (b) of lemma 3.6 also holds. Thus (4.14) is proved. \( \square \)

To prove theorem 2.2 we need the following result.

**Lemma 4.3.** For any \( \phi \in S \) and any \( \tau \in \mathbb{Q} \),

\[ w^{\phi}(x, \tau, s) \leq s(x), \quad x \in X. \]

**Proof.** Fix \( x \in X, \phi \in S, \tau \in \mathbb{Q} \). Consider an arbitrary \( \varepsilon > 0 \) and arbitrary function \( \ell \in L(s) \). Let \( \delta \) be a positive number such that \( \delta \ell(x) \leq \varepsilon \).

The proved part of theorem 2.1 implies the existence of a stationary policy \( \psi \) such that \( w^{\psi} \geq s-\delta \ell \).

Define a nonrandomized policy \( \sigma \) by

\[ \sigma^{i}(x_{0} a_{0} \ldots x_{i}) = \begin{cases} \phi(x_{i})' & \text{for } i < \tau, \\ \psi(x_{i}) & \text{for } i \geq \tau. \end{cases} \]

Then

\[ w^{\phi}(x, \tau, s) = w^{\phi}(x, \tau, w^{\psi} + \delta \ell) = w^{\phi}(x, \tau, w^{\psi}) + \delta \mathbb{E}_{x}^{\phi} \ell(x_{\tau}) \leq w^{\sigma}(x) + \delta \ell(x_{\tau}). \]

But by lemma 3.3 \( \mathbb{E}_{x}^{\phi} \ell(x_{\tau}) \leq \ell(x) \). The policy \( \sigma \) is finite. Therefore by lemma 3.9 \( w^{\sigma}(x) \leq s(x) \). Consequently

\[ w^{\phi}(x, \tau, s) \leq s(x) + \delta \ell(x) \leq s(x) + \varepsilon. \]

Since \( \varepsilon > 0 \) is arbitrary, the lemma is proved. \( \square \)

**The Proof of Theorem 2.2.** For any \( \phi \in S \),

\[ T_{s}(x) \geq T^{\phi}(x) \geq T^{\phi}(x) w^{\phi}(x) = w^{\phi}(x), \quad x \in X. \]

So \( T_{s} \geq s \). Let \( x \in X \) and \( a \in A(x) \) be fixed and let \( \phi \) be a stationary policy such that \( \phi(x) = a \). By lemma 4.3 (for \( \tau \equiv 1 \))

\[ T^{a} s(x) = w^{\phi}(x, 1, s) \leq s(x). \]
Consequently $T_s \leq s$. \qed

To prove theorem 2.1 for functions $\in L(s, Z)$ we need the following lemma.

**Lemma 4.4.** There exists a stationary policy $\psi \in S$ such that

(a) $w^\psi(x) = s(x), \ x \in Z$,

(b) $p(y \mid x, \psi(x)) = 0, \ x \in Z, \ y \in X \setminus Z$.

**Proof.** For arbitrary $x \in X$ and $\phi \in S$ denote

$r^\phi(x) = \{ y \in X : p^\phi(x,y) > 0 \text{ for some } n = 0, 1, \ldots \}$.

Let $Z = \{ x^i, \ i = 1, 2, \ldots \}$. For each $i = 1, 2, \ldots$ consider a policy $\phi^i \in S$ such that $w^{\phi^i}(x^i) = s(x^i)$.

For any $i, n = 1, 2, \ldots$

$s(x^i) = w^\phi(x^i) = w^{\phi^i}(x^i, n, s) \leq w^{\phi^i}(x^i, n, s) \leq s(x^i)$

(the last inequality follows from lemma 4.3). Consequently

$$E^\phi(x^i_n, s(x^i_n)) = 0.$$ But $w^\phi \leq s$ for all $\phi \in S$. Hence $w^\phi(x) = s(x)$ for any $x \in r^{\phi^i}(x^i), i = 1, 2, \ldots$, and thus $r^{\phi^i}(x^i) \subseteq Z$.

Define $D_1 = r^{\phi^1}(x^1), D_i = r^{\phi^i}(x^i) \setminus \bigcup_{k=1}^{i-1} D_k$, where $i = 2, 3, \ldots$.

Then $Z = \bigcup_{i=1}^{\infty} D_i$ and $D_i \cap D_j = \emptyset$, $i \neq j$.

Let $\phi$ be some fixed stationary policy. Define a stationary policy $\psi$ by $\psi(x) = \phi^i(x)$ for $x \in D_i$, $i = 1, 2, \ldots$, and $\psi(x) = \phi(x)$ for $x \in X \setminus Z$.

Obviously equality (b) holds for the policy $\psi$. Show by induction that equality (a) holds too.

For $x \in r^{\phi^1}(x^i)$, equality (a) is proved above. Consider some $i = 2, 3, \ldots$ Let $w^\psi(x) = s(x)$ for all $x \in \bigcup_{i=1}^{\infty} D_k$.

Denote $\tau = \min \{ n : x_n \in \bigcup_{k=1}^{i-1} D_k \}$. Then for any $x \in D_i$

$$w^\psi(x) = w^\phi(x, \tau, 0) + E^\psi_x w^\psi(x_\tau) = w^{\phi^i}(x, \tau, 0) + E^\phi^i x s(x_\tau) \geq$$

$$\geq w^{\phi^i}(x, \tau, 0) + E^\phi x w^{\phi^i}(x_\tau) = w^\phi(x).$$
L(s) and L(s, Z) is used in the proof of theorem 2.1 only to estimate (4.3). Therefore if (4.3) is proved without using this condition, then theorem 2.1 holds for functions l from the class K(Z) ⊇ L(s, Z).

5. Deterministic dynamic programming. In this section we show that if the model is deterministic then theorem 2.1 holds for functions λ ∈ K(Z). According to remark 4.5, it is sufficient to prove that for deterministic models inequality (4.3) always holds.

If a stationary policy φ and an initial state x are fixed then in a deterministic model any Markov time τ is deterministic.

Consider the Markov time ρ defined after the inequality (4.3). If ρ = ∞, then E^ρ_x f(x) = 0 and (4.3) holds.

Let ρ < ∞. Then by definition of the policy φ and the set G from (4.3) and lemma 4.3

s(x) - ε ≤ w^φ(x) = w^φ(x, ρ, w^φ) ≤ w^φ(x, ρ, s-ε) = w^φ(x, ρ, s) - ε ≤ s(x) - ε.

Since ε' < ε, we obtain a contradiction. Hence ρ = ∞ and (4.3) holds.

The proofs of all other results listed in section 2 are based on theorem 2.1. Therefore for deterministic models these results hold for functions l from more general classes introduced in section 2.

REFERENCES

13 Schüli, M., Stationary policies in dynamic programming models under compactness assumptions, Universität Bonn, preprint, 1981.

E.A. Fainberg, Moscow, Moscow Institute of Transport Engineers (MIIT), Department of Applied Mathematics

I.M. Sonin, Moscow, Central Econ.-Math. Institute, Academy of Science of the USSR