EXTENSION OF THE CLASS OF MARKOV MODELS

V.I. Arkin

CP-84-8
March 1984

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

PREFACE

In a recent book, the author proposed a new method of solving stochastic control problems, which, unlike the traditional approach, is not based on dynamic programming techniques. The main features of the new method are the extension of the Markov controls and the use of non-Markov controls which depend on the complete history of the process.

In this extended control domain the optimal control problem becomes a mathematical programming problem in the space of functions and can be studied using convex analysis. The author first generalizes the Markov control extension theorem for problems with constraints which depend on future time, and then obtains a method for finding the optimal control in convex problems through the solution of the auxiliary mathematical programming problem.

EXTENSION OF THE CLASS OF MARKOV CONTROLS

V.I. Arkin
Central Economics and Mathematics Institute (CEMI),
Moscow, USSR

INTRODUCTION

In control theory, for example when deriving existence theorems or optimality criteria, it is often necessary to extend the class of controls without changing the value of the problem. There are a number of well-known methods for doing this which are based on the convexity of integrals of measurable multifunctions and which are related to randomized and relaxed controls.

This paper is devoted to some new theorems of this kind for control problems involving stochastic difference equations with mixed constraints on phase coordinates and controls.

The results presented here are generalizations and extensions of earlier results obtained by the author [1].

1. STATEMENT OF THE PROBLEM

Let s_t be a Markov process defined on a measurable space $(S,\mathcal{E})$. Assume that s_t has transition function $P_t(s_t,ds_{t+1})$, $t = 0,1,\ldots$ and initial distribution $P_0(ds_0)$.

Consider the following problem:

$$\sum_{t=0}^{\tau-1} E\phi^{t+1}(s_t,s_{t+1},y_t,u_t) \rightarrow \max \qquad (1)$$

subject to

$$y_{t+1} = f^{t+1}(s_t, s_{t+1}, y_t, u_t) \quad , \quad y_0 = \tilde{y}_0(s_0) \tag{2}$$

$$u_t = u_t(s^t) \in U_t(s_t) \tag{3}$$

$$g^{t+1}(s_t, s_{t+1}, y_t, u_t) \geq 0 \tag{4}$$

and

$$u_t = V_t(s_t, y_t) \tag{5}$$

for some measurable function $V_t(s,y)$, i.e., u_t is a Markov control. Relations (2)-(5) hold almost surely (a.s.).

Here $s^t = (s_0, s_1, \ldots, s_t)$ is the "history" of the process s_t up to time t and $U_t(s_t)$ is a measurable multifunction with values in a Polish space U with Borel σ-algebra $\mathcal{B}$ such that graph Gr $U_t(s_t) \in \mathcal{B} \times \mathcal{E}$, $y_t \in R^n$, $g_t \in R^m$, and ϕ^t, f^t, g^t are measurable. Controls which have the form $u_t = u_t(s^t)$ we call *non-anticipatory*.

THEOREM 1 (Sufficiency of Markov controls). *Let* $\{\tilde{u}_t\}_0^{\tau-1}$ *be a non-anticipatory control and* $\{\tilde{y}_t\}_0^{\tau-1}$ *a trajectory such that constraints (2)-(4) are satisfied. Then there exists a Markov control* $\{u_t\}_0^{\tau-1}$ *and a trajectory* $\{y_t\}_0^{\tau}$ *which satisfy both constraints (2)-(5) and the following inequality:*

$$\sum_{t=0}^{\tau-1} E\phi^{t+1}(s_t, s_{t+1}, y_t, u_t) \geq \sum_{t=0}^{\tau-1} E\phi^{t+1}(s_t, s_{t+1}, \tilde{y}_t, \tilde{u}_t) \quad . \tag{6}$$

In other words, it is sufficient to consider only the class of Markov controls when searching for a solution of problem (1)-(4). Thus the problems (1)-(4) and (1)-(5) are equivalent.

2. APPLICATIONS

(1)-(4) is a mathematical programming problem in the space of variables $\{u_t(s^t), y_t(s^t)\}_{t=0}^{t=\tau-1}$. Under certain assumptions, both a maximum principle and an existence theorem can be derived for this problem (see, e.g., [1]). By contrast, (1)-(5) is a dynamic programming problem which can be solved only by applying Bellman's equation, and this can be very complicated. The theorem given above states that all results obtained for problem (1)-(4) are also valid for problem (1)-(5).

3. PRELIMINARY RESULTS

The following measurable selection theorem will be used in the proof of Theorem 1.

THEOREM 2 (Sant-Bev). *Let* $(X, \mathcal{B})$ *be a Polish space with Borel σ-algebra and* $(\Omega, \mathcal{F})$ *be an arbitrary measurable space. Then for each* $\Gamma \in \mathcal{F} \times \mathcal{B}$ *there exists an* $\bar{\mathcal{F}}$*-measurable selection* $\xi(w)$ *(such that* $(w, \xi(w)) \in \Gamma$*), where* $\bar{\mathcal{F}}$ *is the universal completion of* $\mathcal{F}$.

The following corollary is also helpful.

COROLLARY. *The projection of* Γ *on* Ω *is such that* $\mathrm{Proj}_{\Omega}\, \Gamma \in \bar{\mathcal{F}}$.

LEMMA 1. *Let* u_t *be a Markov control and* $u_t \in U_t(s_t)$ (a.s.). *Then there exists an* $(\mathcal{F} \times \mathcal{B})$*-measurable* $V_t(s,y)$ *such that:*

(i) $u_t = V_t(s_t, y_t)$ (a.s.)

(ii) $P\{V_t(s_t, y) \in U_t(s_t)\ \forall\ y\} = 1$.

PROOF. Since u_t is a Markov control, then there exists a $V_t(s,y)$ with property (i). We define the set $\mathcal{D}$ as follows:

$$\mathcal{D} = \{(s,y) : V_t(s,y) \in U_t(s)\} .$$

$\mathcal{D}$ is measurable, since $\mathcal{D} = \{(s,y) : (s, V_t(s,y)) \in \mathrm{Gr}\, U_t(s)\}$.

Let $\mathcal{Q}$ be the image of the measure P in the space $S \times R^n$ under the mapping $s^t \to (s_t, y_t)$. Then $\mathcal{Q}(\mathcal{D}) = 1$, $\mathrm{Proj}_S \mathcal{D} \in \bar{\mathcal{E}}$ (the universal completion of $\mathcal{E}$), and $\mathcal{Q}_S(\mathrm{Proj}_S \mathcal{D}) = 1$, where $\mathcal{Q}_S$ is the projection of measure $\mathcal{Q}$ on S.

From the measurable selection theorem, there exists a measurable function $\hat{u}(s) \in U_t(s)$ ($\mathcal{Q}_S$-a.s.). The function

$$V_t^1(s,y) = \begin{cases} V_t(s,y) & , \ (s,y) \in \mathcal{D} \\ \hat{u}(s) & , \ (s,y) \bar{\in} \mathcal{D} \end{cases}$$

then satisfies conditions (i) and (ii) of the lemma.

LEMMA 2. *Let* (Ω,F,P) *be a probability space with σ-algebra* $F_0 \subseteq F$ *and* (U,B) *be a Polish space. Take* $\Phi(w,u)$ *to be* $(F_0 \times B)$*-measurable and let* $w \to \Gamma(w)$ *be a multifunction with graph* $\Gamma = \{w,u{:}u \in \Gamma(w)\} \in F_0 \times B$. *Assume that* $u(w) \in \Gamma(w)$ (a.s.) *and that* $u(w)$ *is* F*-measurable,* $E|\Phi(w,u(w))| < \infty$. *Then there exists an* F_0*-measurable function* $v(w) \in \Gamma(w)$ (a.s.), *such that*

$$E\,\Phi(w,v(w)) \geq E\,\Phi\,(w,u(w)) \quad .$$

PROOF. Let $\Psi(w) = E[\Phi(w,u(w))|F_0]$ and set

$$A = \{(w,u):\Phi(w,u) \geq \Psi(w) \ , \ u \in \Gamma(w)\} \quad ,$$

so that $A \in F_0 \times B$. Denote $\mathcal{D} = \mathrm{Proj}_\Omega A \in \bar{F}_0$, where $\bar{F}_0$ is the universal completion of F_0. Let us show that $P(\mathcal{D}) = 1$.

If this is *not* true, then

$$P\{B \equiv \Omega - \mathcal{D}\} = \{w{:}\Phi(w,u) < \Psi(w), \quad \forall\ u \in \Gamma(w)\} > 0$$

and since $\Phi(w,u) < \Psi(w)$ for each $w \in B$, we have $E\ \chi_B\Phi(w,u(w)) < E\ \chi_B\Psi(w)$, which contradicts the definition of $\Psi(w)$.

From the measurable selection theorem there exists an $\bar{F}_0$-measurable function $v(w)$, $(w,v(w)) \in A$. This means that $v(w) \in \Gamma(w)$ (a.s.) and

$$\Phi(w,v(w)) \geq \Psi(w) \ \text{(a.s.)} \Rightarrow E\ \Phi(w,v(w)) \geq E\ \Psi(w) = E\ \Phi(w,u(w)) \quad .$$

4. PROOF OF THEOREM 1

The proof will be divided into three parts and carried out by induction.

4.1. *Inductive assumptions.* Assume that we have constructed random vectors $y_{k+1}^k,\ldots,y_\tau^k$ and measurable functions $V_k(s,y),\ldots,V_{\tau-1}(s,y)$ with the following properties:

1. $P\{V_t(s_t,y) \in U_t(s_t) \;,\; \forall y \in R^n\} = 1$ $\qquad t = k,\dots,\tau-1$

2. $y_t^k = f^t(s_{t-1},s_t,y_{t-1}^k,V_{t-1}(s_{t-1},y_{t-1}^k)), y_t^k = \tilde{y}_t$ (a.s.) $\qquad t = k+1,\dots,\tau$

3. $g^{t+1}(s_t,s_{t+1},y_t^k,V_t(s_t,y_t^k)) \geq 0$ (a.s.) $\qquad t = k,\dots,\tau-1$

4. $E \sum_{t=k}^{\tau-1} \phi^{t+1}(s_t,s_{t+1},y_t^k,V_t(s_t,y_t^k)) \geq E \sum_{t=k}^{\tau-1} \phi^{t+1}(s_t,s_{t+1},\tilde{y}_t,\tilde{u}_t)$.

The theorem will be proved if it can be established that k can be replaced by k-1 in these four relations.

4.2. Preliminaries. Let $Y_t(u)$ $(k \leq t \leq \tau)$ be a sequence of random variables which depend on the parameter $u \in U_{k-1}(s_{k-1})$:

$$Y_t(u) = f^t(s_{t-1},s_t,Y_{t-1}(u) \;,\; V_{t-1}(s_{t-1},Y_{t-1}(u))) \quad , \quad t > k$$

$$Y_k(u) = f^k(s_{k-1},s_k,\tilde{y}_{k-1},u) \quad .$$

It is easily seen that the $Y_t(u)$ are measurable with respect to the σ-algebra $F_{s_{k-1},\dots,s_t,\tilde{y}_{k-1}} \times B$ and that $Y_t(\tilde{u}_{k-1}) = y_t^k$.

Put

$$G^t(s_{k-1},\dots,s_t,\tilde{y}_{k-1},u) = g^t(s_{t-1},s_t,Y_{t-1}(u),V_{t-1}(s_{t-1},Y_{t-1}(u))) \;,$$

$$t = k,\dots,\tau-1$$

and consider the sets

$$\Gamma^t = \{w,u: G^t(s_{k-1},\dots,s_t,\tilde{y}_{k-1},u) \geq 0 \quad , \quad \pi(s_{k-1},ds_k,\dots,ds_t)\text{- (a.s.)}\},$$

$$k \leq t \leq \tau \;,$$

where $\pi(s_{k-1},ds_k,\dots,ds_t)$ is the conditional distribution of random parameters $s_k,\dots,s_t$, given s_{k-1}.

Since

$$\Gamma^t = \{w,u: \int G_-^t(s_{k-1},\dots,s_t,\tilde{y}_{k-1},u)\ \pi(s_{k-1},ds_k,\dots,ds_t) = 0\}\ ,$$

$G_-^t = \min\ (G^t,0)$, we have $\Gamma^t \in F_{s_{k-1},\tilde{y}_{k-1}} \times \mathcal{B}$.

Define:

$$\Gamma = \bigcap_{t=k}^{\tau} \Gamma^t \cap \{w,u: u \in U_{k-1}(s_{k-1})\} \tag{7}$$

$$F(s_{k-1},\dots,s_\tau,\tilde{y}_{k-1},u) = \phi^k\ (s_{k-1},s_k,\tilde{y}_{k-1},u) + \\ + \sum_{t=k}^{\tau-1} \phi^{t+1}\ (s_t,s_{t+1},Y_t(u),V_t(s_t,Y_t(u))) \tag{8}$$

$$\Phi(s_{k-1},\tilde{y}_{k-1},u) = \int F(s_{k-1},\dots,s_{\tau-1},\tilde{y}_{k-1},u)\ \pi(s_{k-1},ds_k,\dots,ds_\tau). \tag{9}$$

4.3. Use of Lemma 2. Let us apply Lemma 2 to the set $\Gamma(w) = \{u:(w,u) \in \Gamma\}$ defined by (7), to the function Φ defined by (9), and to the σ-algebra $F_0 = F_{s_{k-1},\tilde{y}_{k-1}}$. This shows that there exists a measurable function $V(s,y)$ such that

$$E\Phi(s_{k-1},\tilde{y}_{k-1},V(s_{k-1},\tilde{y}_{k-1})) \geq E\Phi(s_{k-1},\tilde{y}_{k-1},\tilde{u}_{k-1}) \tag{10}$$

and with probability 1:

$$V(s_{k-1},\tilde{y}_{k-1}) \in U_{k-1}(s_{k-1}) \tag{11}$$

$$G^t(s_{k-1},\dots,s_t,\tilde{y}_{k-1},V(s_{k-1},\tilde{y}_{k-1})) \geq 0 - \pi(s_{k-1},ds_k,\dots,ds_t)\ \text{(a.s.)}. \tag{12}$$

This last relation is equivalent to

$$G^t(s_{k-1},\dots,s_t,\tilde{y}_{k-1},V(s_{k-1},\tilde{y}_{k-1})) \geq 0 \quad \text{(a.s.)}\ . \tag{13}$$

4.4. Completion of the proof. From Lemma 1, there exists a measurable function $V_{k-1}(s,y)$ such that

$$V_{k-1}(s_{k-1},\tilde{y}_{k-1}) = V(s_{k-1},\tilde{y}_{k-1}) \quad \text{(a.s.)}$$

$$P\{V_{k-1}(s_{k-1},y) \in U_{k-1}(s_{k-1}) \ , \ \forall y\} = 1 \quad .$$

It is clear that the relations (10)-(13) remain valid if we replace V by V_{k-1}.

Now define

$$y_{k-1}^{k-1} = \tilde{y}_{k-1} \ ; \ y_t^{k-1} = f^t(s_{t-1},s_t,y_{t-1}^{k-1},V_{t-1}(s_{t-1},y_{t-1}^{k-1})) \ ,$$

noting that

$$Y_t(V_{k-1}(s_{k-1},\tilde{y}_{k-1})) = y_t^{k-1} \qquad (t \geq k-1) \quad .$$

Then from (13) we obtain

$$g^t(s_{t-1},s_t,y_{t-1}^{k-1},V_{t-1}(s_{t-1},y_{t-1}^{k-1})) =$$

$$= G^t(s_{k-1},\ldots,s_t,\tilde{y}_{k-1},V_{k-1}(s_{k-1},\tilde{y}_{k-1})) \geq 0 \quad \text{(a.s.)}$$

and from (8) we get

$$E \sum_{t=k-1}^{\tau-1} \phi^{t+1}(s_t,s_{t+1},y_t^{k-1},V_t(s_t,y_t^{k-1})) =$$

$$= E\,F(s_{k-1},\ldots,s_\tau,\tilde{y}_{k-1},V_{k-1}(s_{k-1},\tilde{y}_{k-1})) =$$

$$= E\,\Phi(s_{k-1},\tilde{y}_{k-1},V_{k-1}(s_{k-1},\tilde{y}_{k-1})) \geq$$

$$\geq E\,\Phi(s_{k-1},\tilde{y}_{k-1},\tilde{u}_{k-1}) = E\ \phi^k(s_{k-1},s_k,\tilde{y}_{k-1},\tilde{u}_{k-1}) +$$

$$+ E \sum_{t=k}^{\tau-1} \phi^{t+1}(s_t,s_{t+1},y_t^k,V_t(s_t,y_t^k)) \geq$$

$$\geq E \sum_{t=k-1}^{\tau-1} \phi^{t+1}(s_t,s_{t+1},y_t,u_t) \quad ,$$

using the inductive assumptions. This completes the proof.

REMARK. *The case of independent* s_t. Let the random elements s_t, $t = 0,1...$ be independent and assume that the mappings ϕ^t, f^{t+1}, U_t, g^{t+1} do not depend on s_t. Then for each non-anticipatory control one can choose a special kind of Markov control which depends only on the values y_t of the controlled process

$$v_t = v_t(y_t) \quad . \tag{14}$$

This implies the Blackwell-Strauch-Ryll-Nardzewski theorem on the sufficiency of simple strategies for controlled Markov processes.

5. CONSTRUCTION OF MARKOV CONTROLS

5.1. Preliminaries. Suppose now that the convexity conditions stated below are satisfied for problem (1)-(5). In this case, it is possible to construct (quite efficiently) the majorizing Markov pair (y_t, u_t) for every non-anticipatory pair $(\tilde{y}_t, \tilde{u}_t)$ which satisfies constraints (2)-(4). (Note that the time moment τ is not necessarily finite.)

CONVEXITY CONDITIONS. *For any collection* $(s_t, y^1, y^2, u^1, u^2, \alpha)$, $y^1, y^2 \in R^n$, $u^1, u^2 \in U_t(s_t)$, $0 \leq \alpha \leq 1$, *there exists a* $u \in U_t(s_t)$ *such that the following conditions are satisfied* $P_t(s_t, ds_{t+1})$-a.s.:

$$\alpha\phi^{t+1}(s_t, s_{t+1}, y^1, u^1) + (1-\alpha)\phi^{t+1}(s_t, s_{t+1}, y^2, u^2) \leq$$
$$\leq \phi^{t+1}(s_t, s_{t+1}, \alpha y^1 + (1-\alpha)y^2, u) \tag{15}$$

$$\alpha f^{t+1}(s_t, s_{t+1}, y^1, u^1) + (1-\alpha) f^{t+1}(s_t, s_{t+1}, y^2, u^2) =$$
$$= f^{t+1}(s_t, s_{t+1}, \alpha y^1 + (1-\alpha)y^2, u) \tag{16}$$

$$\alpha g^{t+1}(s_t, s_{t+1}, y^1, u^1) + (1-\alpha) g^{t+1}(s_t, s_{t+1}, y^2, u^2) \leq$$
$$\leq g^{t+1}(s_t, s_{t+1}, \alpha y^1 + (1-\alpha)y^2, u) \quad . \tag{17}$$

In order to simplify the proof we shall also assume that the sets $U_t(s_t)$ are compact and that the functions ϕ^t, f^t, g^t are both continuous with respect to (y,u) and bounded with respect to y on any bounded set $C \subseteq R^n$:

$|\phi^t| + |f^t| + |g^t| \leq K_C$, $y \in C$ for some constant $K_C > 0$. Assume also that $y_0(s_0)$ is a bounded function.

THEOREM 3.

1. *Let sequences* $\{\tilde{u}_t\}$, $\{\tilde{y}_t\}$ *satisfy the conditions of Theorem 1. Then there exists a Markov pair* $\{u_t\}$, $\{y_t\}$ *which satisfies constraints* (2)-(5) *and is such that the process* y_t *is defined by the following equations:*

$$y_{t+1} = E[\tilde{y}_{t+1}/s_t, s_{t+1}, y_t] \tag{18}$$

and

$$E\phi^{t+1}(s_t, s_{t+1}, y_t, u_t) \geq E\phi^{t+1}(s_t, s_{t+1}, \tilde{y}_t, \tilde{u}_t) \quad , \quad t = 0,1\ldots \tag{19}$$

2. *If the elements* s_t *are independent and the mappings* ϕ^{t+1}, f^{t+1}, g^{t+1}, U_t *do not depend on* s_t, *then it is possible to choose Markov controls of the form* $u_t = u_t(y_t)$, *where the process* y_t *is defined by the process* $\tilde{y}_t$ *as follows:*

$$y_{t+1} = E[\tilde{y}_{t+1}/s_{t+1}, y_t] \ . \tag{20}$$

The pair $\{u_t\}$, $\{y_t\}$ *satisfies both* (2)-(5) *and inequality* (19).

We shall now formulate two auxiliary results which will be used in the proof of Theorem 3.

LEMMA 3. *Let* U *be a Polish space,* $u(s^t)$ *be a measurable function defined on* U, $\alpha(s^t)$ *be another measurable function, and* $\pi(s_t, \alpha, du)$ *be the conditional distribution of* $u(s^t)$ *for fixed* s_t *and* $\alpha(s^t)$. *Then for any measurable function* $\beta(s_t, s_{t+1}, u)$ *such that the function* $\beta(s_t, s_{t+1}, u(s^t))$ *is summable, the following equality is satisfied:*

$$E[\beta(s_t, s_{t+1}, u(s^t))/s_t, s_{t+1}, \alpha] = \int_U \pi(s_t, \alpha(s^t), du)\, \beta(s_t, s_{t+1}, u) \quad (a.s.) \ .$$

Let U be a metric compact set, Y be a compact set in R^n, S be a measurable space with probabilistic measure ν, and function $\psi(y,u,s)$ be continuous with respect to (y,u), measurable with respect to s, and with values in finite-dimensional space.

Assume that the following convexity condition is satisfied:

for all $y^1, y^2 \in Y$, $u^1, u^2 \in U$, $0 \leq \alpha \leq 1$

there exists a $u \in U$ *such that* ν-a.s.

$$\alpha\psi(y^1,u^1,s) + (1-\alpha)\psi(y^2,u^2,s) \leq \psi((1-\alpha)y^1 + \alpha y^2,u,s) .$$

LEMMA 4. *For any probabilistic measure* μ *on* $Y \times U$ *there exists a* $u \in U$ *such that*

$$\int_{Y\times U} \psi(y,u,s)\,\mu(dy \times du) \leq \psi\Big(\int_{Y\times U} y\,\mu(dy \times du,u,s)\Big) \quad (\nu\text{-a.s.}) .$$

The proof of these simple results can be found in [1].

5.2. Proof of Theorem 3. We shall prove only the first part of the theorem since the proof of the second part is analogous to that of the first. We shall first verify that there exists a measurable function $\bar{u}_t = \bar{u}_t(s_{t-1},s_t,y_{t-1})$ such that the following relations are satisfied:

$$E\phi^{t+1}(s_t,s_{t+1},y_t,\bar{u}_t) \geq E\phi^t(s_t,s_{t+1},\tilde{y}_t,\tilde{u}_t) \tag{21}$$

$$E[\tilde{y}_{t+1}/s_{t-1},s_t,s_{t+1},y_{t-1}] = f^{t+1}(s_t,s_{t+1},y_t,\bar{u}_t) \tag{22}$$

$$g_{t+1}(s_t,s_{t+1},y_t,\bar{u}_t) \geq 0 \tag{23}$$

$$\bar{u}_t \in U_t(s_t) \quad . \tag{24}$$

We shall denote by $\pi(s_{t-1},s_t,\tilde{y}_{t-1},dy \times du)$ the conditional distribution of the element $(\tilde{y}_t,\tilde{u}_t)$ for fixed values of the element $(s_{t-1},s_t,\tilde{y}_{t-1})$. Take

$$J_1(s_{t-1},s_t,s_{t+1},y_{t-1}) = E[\phi^{t+1}(s_t,s_{t+1},\tilde{y}_t,\tilde{u}_t)/s_{t-1},s_t,s_{t+1},y_{t-1}] =$$

$$= \int_{R^n\times U_t(s_t)} \pi(s_{t-1},s_t,y_{t-1},dy \times du)\, \phi^{t+1}(s_{t-1},s_t,s_{t+1},y,u) \tag{25}$$

$$J_2(s_{t-1},s_t,s_{t+1},y_{t-1}) = E[f^{t+1}(s_t,s_{t+1},\tilde{y}_t,\tilde{u}_t)/s_{t-1},s_t,s_{t+1},y_{t-1}] =$$

$$= \int_{R^n \times U_t(s_t)} \pi(s_{t-1},s_t,y_{t-1},dy \times du)\ f^{t+1}(s_t,s_{t+1},y,u) \qquad (26)$$

$$J_3(s_{t-1},s_t,s_{t+1},y_{t-1}) = E[g^{t+1}(s_t,s_{t+1},\tilde{y}_t,\tilde{u}_t)/s_{t-1},s_t,s_{t+1},y_{t-1}] =$$

$$= \int_{R^n \times U_t(s_t)} \pi(s_{t-1},s_t,y_{t-1},dy \times du)\ g^{t+1}(s_t,s_{t+1},y,u)\ . \qquad (27)$$

The equalities (25)-(27) are due to Lemma 3. Lemma 4 and the convexity condition imply that for every value of parameters (s_{t-1},s_t,y_{t-1}) there exists an element $u \in U_t(s_t)$ such that the following relations are satisfied $P_t(s_t,ds_{t+1})$-a.s.:

$$J_1(s_{t-1},s_t,s_{t+1},y_{t-1}) \leq \phi^{t+1}(s_t,s_{t+1},E[\tilde{y}_t/s_{t-1},s_t,y_{t-1}],u) =$$

$$= \phi^{t+1}(s_t,s_{t+1},y_t,u) \qquad (28)$$

$$J_2(s_{t-1},s_t,s_{t+1},y_{t-1}) = f^{t+1}(s_t,s_{t+1},E[\tilde{y}_t/s_{t-1},s_t,y_{t-1}],u) =$$

$$= f^{t+1}(s_t,s_{t+1},y_t,u) \qquad (29)$$

$$J_3(s_{t-1},s_t,s_{t+1},y_{t-1}) \leq g^{t+1}(s_t,s_{t+1},E[\tilde{y}_t/s_{t-1},s_t,y_{t-1}],u) =$$

$$= g^{t+1}(s_t,s_{t+1},y_t,u) \quad . \qquad (30)$$

According to the measurable selection theorem there exists a measurable function $\bar{u}_t = \bar{u}_t(s_{t-1},s_t,y_{t-1})$ for which (28)-(30) are satisfied. Relations (28)-(30) immediately lead to (21)-(24).

The second part of the proof is similar to the first. It is necessary only to take the conditional mathematical expectation with respect to (s_t, s_{t+1},y_t) in (21)-(24) and apply Lemmas 3 and 4, and the measurable selection theorem, making use of the fact that y_t depends measurably on (s_{t-1},s_t,y_{t-1}) (see (16)).

REFERENCE

[1] V.I. Arkin and I.V. Evstigneev. *Stochastic Models of Control and Economic Dynamics*. Nauka, Moscow, 1978.