Z. Wahrscheinlichkeitstheorie verw. Gebiete 32, 179–196 (1975)

Conditions for Optimality in Dynamic Programming and for the Limit of n-Stage Optimal Policies to Be Optimal

Manfred Schäl

0. Introduction

The present work deals with a stationary decision model allowing the discount factor to depend on the states of the system and of the selected actions. The model includes as special case stationary models in the sense of Blackwell [2] and Strauch [21]. However, the results of this paper generalize to non-stationary non-Markovian decision models in the sense of Hinderer [6] (cp. [17]).

We impose a rather weak convergence condition (condition (C)) on the expected total rewards thus including the negative (unbounded) case and the discounted case.

The main purpose of the present paper is to give sufficient conditions for the existence of an optimal policy and to interrelate the optimal total expected rewards as well as the optimal actions of the model with infinite horizon and those of the model with finite horizon N as N tends to infinity. The results of the paper may be regarded as generalizations of results by Blackwell [2] Theorem 7b, Strauch [21] Theorem 9.1, Maitra [11], Hinderer [6] Theorem 17.12, Hinderer [7] Theorem 4.2, Furukawa [5] Theorem 4.2, [16] Theorem 7.2. In particular, it turns out that the results by Maitra and Furukawa proved for the discounted case carry over to the negative case.

The analysis of this paper is based on some results on set-valued mappings, upper semi-continuous functions, measurable selections, and topologies on spaces of probability measures presented in Sections 9–12 and 14.

1. The Decision Model

The background for a theory of dynamic programming may be provided by a *decision model* which is given by a tupel $((S, \mathfrak{S}), (A, \mathfrak{A}), D, q, \beta, r)$ of the following meaning:

(i) $(S, \mathfrak{S})$ stands for the *state space* and is assumed to be a standard Borel space, i.e. S is a non-empty Borel subset of a Polish (complete, separable, metric) space and $\mathfrak{S}$ is the system of all Borel subsets of S.

(ii) $(A, \mathfrak{A})$ is the *action space* and is assumed to be a standard Borel space.

(iii) D: $S \to \mathfrak{P}'(A)$, where $\mathfrak{P}'(A)$ denotes the set of all non-empty subsets of A, specifies the *set of all admissible actions* $D(s)$ if the system is in state s. We assume

$$K = \{(s, a) \in S \times A; a \in D(s)\} \tag{2.1}$$

is an $\mathfrak{S} \otimes \mathfrak{A}$-measurable set and contains the graph of a measurable map of S into A [1].

(iv) The so-called *transition law* q is a transition probability q: $K \to \mathscr{P}(S)$ [2]. $q(s, a; \cdot)$ is the distribution of the state next visited by the system if the system is in state s and the action a is taken.

(v) β is a bounded measurable function of $K \times S$ into the set of the non-negative real numbers and can be interpreted as a *discount factor*.

(vi) The *reward function* r: $K \times S \to \underline{\mathbb{R}}$ [3] is a measurable function bounded from above.

Given that we have experienced the history $(s_1, a_1, \ldots, s_{n+1})$, we will receive for period n the *discounted reward*

$$r_n(s_1, a_1, \ldots, s_{n+1}) = \beta(s_1, a_1, s_2) \ldots \beta(s_{n-1}, a_{n-1}, s_n)\, r(s_n, a_n, s_{n+1})$$

(especially $r_1 = r$). Models where the discount factor is not constant arise from semi-Markov processes as well as from stopping problems. When dealing with stopping problems, we may write $A = A_c \cup A_s$ where $\beta(s, a, s') = 0$, s, $s' \in S$, if and only if $a \in A_s$, i.e. A_s is the set of terminal actions.

We write $H_1 = S$, $H_{n+1} = K \times H_n$, $n \in \mathbb{N}$ [4]. As usual, a *randomized policy* $\pi = (\pi_n)$ is defined as a sequence of transition probabilities π_n: $H_n \to \mathscr{P}(A)$ such that $\pi_n(s_1, a_1, \ldots, s_n; \cdot)$ assigns probability one to $D(s_n)$ for any $(s_1, a_1, \ldots, s_n) \in H_n$, $n \in \mathbb{N}$. We write Δ for the set of all randomized policies and D^S for the set of all decision functions, i.e.

$$D^S = \{f; f: (S, \mathfrak{S}) \to (A, \mathfrak{A}), f(s) \in D(s) \text{ for } s \in S\}.$$

A *Markov policy* is a sequence (f_n) where $f_n \in D^S$, $n \in \mathbb{N}$. A *stationary policy* is a Markov policy (f_n) where $f_n = f$ is independent of n. For such a policy we write f^∞. We may look at the set of all Markov policies as a subset of Δ. When dealing with models with finite horizon n we have to specify only $f_1, \ldots, f_n$ and we will call $(f_1, \ldots, f_n)$ where $f_i \in D^S$, $1 \leq i \leq n$, an *n-stage Markov policy*. Given some initial distribution $p \in \mathscr{P}(S)$, the transition law and a policy π define a probability measure P_π on the product space $S \times A \times S \times A \times \cdots$ endowed with the product σ-algebra and thus a random process $(\zeta_1, \alpha_1, \zeta_2, \alpha_2, \ldots)$ (cp. Hinderer [6] p. 80) where ζ_n and α_n denote the projection from $S \times A \times \cdots$ onto the n-th state space and the n-th action space, respectively, i.e. the random variables ζ_n and α_n describe the state of the system and the action at time n. In this paper we are concerned only with the conditional distributions $P_\pi(\cdot \mid \zeta_1)$ which are given by q and π and are independent of p. Therefore we can dispense with an initial distribution p.

We remark that a Bayesian decision model may be reduced to a decision model as defined above (cp. Rieder [13]).

---

1 cp. Remark 13.1.

2 For any measurable space $(S, \mathfrak{S})$, let $\mathscr{P}(S)$ denote the set of all probability measures on $\mathfrak{S}$.

3 We use $\mathbb{R}$ [$\underline{\mathbb{R}}$] to denote the set of all real numbers [augmented by the point $-\infty$, respectively].

4 Let $\mathbb{N}$ denote the set of the positive integers.

2. The Total Expected Rewards

In order that the total expected rewards are well defined, we have to impose some convergence assumption. Define

$$I_{+}(\pi)=E_{\pi}\left[\sum_{i=1}^{\infty} r_i^{+} \mid \zeta_1\right]^{5}$$

and

$$u_{+}=\sup_{\pi\in\Delta} I_{+}(\pi).$$

Throughout the paper we make the following

General assumption (GA). $u_{+}(s)<\infty,\ s\in S.$

As has been shown in Hinderer [6], the total rewards

$$R_m^n=\sum_{i=m}^{n} r_i, \qquad m\in\mathbb{N}, \quad n\in\mathbb{N} \text{ or } n=\infty,$$

where $R_m^n=0$ for $n<m$, exist a.s. with respect to $P_\pi(\cdot\mid\zeta_1)$ for any policy π and the following functions are well defined

$$I(\pi)=E_\pi[R_1^\infty\mid\zeta_1], \qquad I_n(\pi)=E_\pi[R_1^n\mid\zeta_1],\ \pi\in\Delta,\ n\in\mathbb{N},$$

$$u^*=\sup_{\pi\in\Delta} I(\pi), \qquad u_n=\sup_{\pi\in\Delta} I_n(\pi),\ n\in\mathbb{N},\ u_0=0,$$

$$u_\infty=\varliminf_{n\to\infty} u_n.$$

For the main results of this paper, it is necessary to impose a stronger convergence condition (condition (C) below) than the general assumption. For the formulation of this condition we have to generalize the notions of $I_n(\pi)$ and u_n and define

$$I_{mn}(\pi)=E_\pi[R_m^n\mid\zeta_1], \quad \pi\in\Delta, \qquad u_{mn}=\sup_{\pi\in\Delta} I_{mn}(\pi),$$

$$u_{m\infty}=\varliminf_{n\to\infty} u_{mn}, \qquad z_m=\sup_n u_{m+1,n}.$$

Then $I_{1,n}(\pi)=I_n(\pi)$, $u_{1,n}=u_n$, $u_{1,\infty}=u_\infty$. Further, it is to be noticed that $z_m\geqq u_{m+1,m}=0$. With these preparations we are now in a position to introduce

Condition (C). $z_m(s)\to 0$ *as* $m\to\infty,\ s\in S.$

Let us discuss the general assumption (GA) and condition (C) in the following cases.

Negative case. $r\leqq 0.$

Discounted case. $\bar\beta=\|\beta\|<1,\ \bar r=\|r\|<\infty$ [6].

Positive case. $r\geqq 0.$

In the negative case, we have $u_{+}=0=z_n$ for $n\in\mathbb{N}$. Thus (GA) and (C) hold. In the discounted case, we have $u_{+}\leqq\bar r/(1-\bar\beta)$ and $z_n\leqq\bar\beta^n\cdot\bar r/(1-\bar\beta)$. Thus (GA) and (C) hold. We remark that (GA) and (C) still hold if $\bar\beta<1$ and $\|r^{+}\|<\infty$. In the positive case, (C) coincides with condition (C^{+}) in Hinderer [7] (and with condition (B) in [16] as well as with condition (N_4) in [15]), i.e. (GA) and (C) hold if and

[5] For an extended real number c we write $c^{+}=\max(0,+c)$.

[6] We write $\|u\|=\sup_{x\in X}|u(x)|$ for any extended real valued function u defined on any set X.

only if

$$\sup_{\pi\in\Delta} E_\pi[R_n^\infty|\zeta_1]\to 0 \quad \text{as } n\to\infty.$$

In general however, (C) may be derived from (C^+).

The following inequalities are obvious from the definitions.

$$u^*\leqq u_+,\quad u_n\leqq u_+. \tag{2.1}$$

$$u_{n+1,\infty}\leqq z_n\leqq u_+. \tag{2.2}$$

$$u_n\leqq u_m+u_{m+1,n}\leqq u_m+z_m,\quad m\leqq n. \tag{2.3}$$

For later use we note that

$$I_n(\pi)\to I(\pi)\quad \text{as } n\to\infty, \tag{2.4}$$

$$u^*\leqq u_\infty. \tag{2.5}$$

The relation (2.4) is a consequence of the general assumption. From (2.4) and the relation $I_n(\pi)\leqq u_n$ we conclude that $I(\pi)\leqq u_\infty$ for any $\pi\in\Delta$ and (2.5) is proved.

3. The Operators L and U

For any $u: S\to\underline{\mathbb{R}}$ such that the following expressions are defined, we set

$$\mathbf{L}u(s,a)=\int q(s,a;ds')[r(s,a,s')+\beta(s,a,s')\,u(s')],\quad a\in D(s),$$

$$\mathbf{L}_f u(s)=\mathbf{L}u(s,f(s)),\quad f\in D^S,$$

$$\mathbf{U}u(s)=\sup_{a\in D(s)}\mathbf{L}u(s,a),\quad s\in S.$$

If r is replaced by 0, we write $\tilde{\mathbf{L}}$ and $\tilde{\mathbf{U}}$ instead of $\mathbf{L}$ and $\mathbf{U}$, respectively. Further, if r is replaced by r^+, we write $\mathbf{U}_+$ instead of $\mathbf{U}$. We use $\mathscr{I}$ to denote the set of all universally measurable functions $u: S\to\mathbb{R}$ such that $u\leqq u_+$. It is easily seen that

$$\mathbf{L}_f^n(u+v)=\mathbf{L}_f^n u+\tilde{\mathbf{L}}_f^n v,\quad f\in D^S,\ u,v\in\mathscr{I},\ n\in\mathbb{N}, \tag{3.1}$$

$$\tilde{\mathbf{L}}_f^n v=E_{f^\infty}[\beta(\zeta_1,\alpha_1,\zeta_2)\dots\beta(\zeta_n,\alpha_n,\zeta_{n+1})\,v(\zeta_{n+1})|\zeta_1]. \tag{3.2}$$

If $\mathbf{L}_f^m u$ and $\tilde{\mathbf{L}}_f^m v$ are not universally measurable for any $1\leqq m<n$, then we cannot define $\mathbf{L}_f^n u$ and $\tilde{\mathbf{L}}_f^n v$ inductively. In this case we define $\tilde{\mathbf{L}}_f^n v$ through (3.2) and $\mathbf{L}_f^n u=\mathbf{L}_f^n 0+\tilde{\mathbf{L}}_f^n u$. By use of the operator $\mathbf{L}$ we can write

$$\mathbf{L}_f I(f^\infty)=I(f^\infty), \tag{3.3}$$

$$\mathbf{L}_f^n 0=I_n(f^\infty),\quad f\in D^S. \tag{3.4}$$

Now we list some *optimality equations* which are easily derived from the results in Hinderer [6, 7].

$$u_+\in\mathscr{I},\quad u_+=\mathbf{U}_+u_+, \tag{3.5}$$

$$u^*\in\mathscr{I},\quad u^*=\mathbf{U}u^*, \tag{3.6}$$

$$u_n\in\mathscr{I},\quad u_n=\mathbf{U}u_{n-1}=\mathbf{U}^n 0, \tag{3.7}$$

$$u_{mn}\in\mathscr{I},\quad u_{m+1,n+1}=\tilde{\mathbf{U}}u_{mn}. \tag{3.8}$$

From (3.5) we obtain

$$\tilde{\mathbf{L}}_f u \leqq \tilde{\mathbf{U}} u \leqq \mathbf{U}_+ u \leqq u_+, \tag{3.9}$$

$$\mathbf{L}_f u \leqq \mathbf{U} u \leqq \mathbf{U}_+ u \leqq u_+, \quad \text{for } u \in \mathscr{J},\ f \in D^S. \tag{3.10}$$

4. The Function u_∞

Lemma 4.1. *Let (c_n) be a sequence of extended real numbers and (ε_n) a sequence of non-negative numbers with $\varepsilon_n \to 0$. If $c_n \leqq c_m + \varepsilon_m$ for $n \geqq m$, then* $\lim c_n$ *exists.*

Proof. From $\sup_{n \geqq m} c_n \leqq c_m + \varepsilon_m$ we conclude that

$$\lim_m \sup_{n \geqq m} c_n \leqq \underline{\lim}\, c_m + \overline{\lim}\, \varepsilon_m.$$

Hence $\overline{\lim}\, c_n \leqq \underline{\lim}\, c_n$. ▯

As a consequence of Lemma 4.1 and (2.3) we obtain

Theorem 4.2. *Assume* (C). *Then* $\lim_n u_n = u_\infty$ *exists.*

Thus, in view of (3.7), we have under condition (C)

$$u_\infty = \lim_n \mathbf{U}^n 0. \tag{4.1}$$

We need the following generalization of the dominated and monotone convergence theorems.

Lemma 4.3. *Suppose there is given a measure μ on $(S, \mathfrak{S})$, a sequence of extended real valued measurable functions v_n on S and a sequence of non-negative functions ε_n on S with $\varepsilon_n \to 0$. If $v_n \leqq g$, $\varepsilon_n \leqq g$ for some μ-integrable function g and $v_n \leqq v_m + \varepsilon_m$ for $n \geqq m$, then* $\lim_n \int v_n \, d\mu$ *and* $\lim_n v_n$ *exist and*

$$\lim_n \int v_n \, d\mu = \int \lim_n v_n \, d\mu.$$

Proof. From the dominated convergence theorem we conclude that $\int \varepsilon_n \, d\mu \to 0$. Thus $v_n(s)$, $s \in S$, and $\int v_n \, d\mu$ satisfy the condition of Lemma 4.1 and the existence of the limits is proved. Upon setting $\bar{v}_n = \sup_{m \geqq n} v_m$, we obtain $v_n \leqq \bar{v}_n \leqq v_n + \varepsilon_n$ and hence $\int v_n \, d\mu \leqq \int \bar{v}_n \, d\mu \leqq \int v_n \, d\mu + \int \varepsilon_n \, d\mu$. The passage to the limit yields $\lim \int v_n \, d\mu = \lim \int \bar{v}_n \, d\mu$. Finally the monotone convergence theorem implies $\lim \int \bar{v}_n \, d\mu = \int \lim \bar{v}_n \, d\mu = \int \lim v_n \, d\mu$. ▯

Theorem 4.4. *Assume* (C). *Then*

4.4.1. $\lim_n \mathbf{L} u_n = \mathbf{L} u_\infty$,
4.4.2. $\mathbf{U} u_\infty \leqq u_\infty$,
4.4.3. $\tilde{\mathbf{L}}_f^n u_\infty \leqq u_{n+1,\infty}$.

Proof. The first assertion can be derived from Lemma 4.3 where $\mu = q(s, a; \cdot)$, $v_n = r(s, a, \cdot) + \beta(s, a, \cdot) u_n$, $\varepsilon_n = \beta(s, a, \cdot) z_n$, $g = r^+(s, a, \cdot) + \beta(s, a, \cdot) u_+$ for some fixed $(s, a) \in K$. By (2.2), (2.3), and (3.5), the conditions of Lemma 4.3 are satisfied. Now

$$\begin{aligned} \mathbf{U} u_\infty(s) &= \sup_{a \in D(s)} \lim_n \mathbf{L} u_n(s, a) \\ &\leqq \underline{\lim}_n \sup_{a \in D(s)} \mathbf{L} u_n(s, a) = \underline{\lim}_n \mathbf{U} u_n(s) \\ &= \lim_n u_{n+1}(s) = u_\infty(s), \end{aligned}$$

where use is made of (3.7).

The third relation may be proved by similar arguments. We remark that $\tilde{\mathbf{L}}_f^n u_m$ may be regarded as the optimal reward in a certain non-stationary decision model and hence is universally measurable. Consequently $\tilde{\mathbf{L}}_f^n u_\infty = \lim_m \tilde{\mathbf{L}}_f^n u_m \in \mathscr{I}$. □

5. Criterion of Optimality

In the present paper we are concerned with the following concept of optimality: A policy $\pi \in \Delta$ will be called *optimal* if $I(\pi) = u^*$.

Lemma 5.1. *Let be $u \in \mathscr{I}$ and $f \in D^S$ such that $u = \mathbf{L}_f u$. Then*

$$u = I_n(f^\infty) + \tilde{\mathbf{L}}_f^n u.$$

Proof. The assertion follows from (3.1) and (3.4). □

The following theorem yields a slight modification of the criteria of optimality given by Hordijk [8] and Rieder [14].

Theorem 5.2. *Let be $f \in D^S$.*

5.2.1. *The following statements are equivalent.*

(i) *f^∞ is optimal and $u_\infty = u^*$.*

(ii) *$\mathbf{L}_f u_\infty = u_\infty$ and $\overline{\lim}_n \tilde{\mathbf{L}}_f^n u_\infty \leqq 0$.*

(iii) *$\mathbf{L}_f u_\infty = u_\infty$ and $\lim_n \tilde{\mathbf{L}}_f^n u_\infty = 0$ on $\{s; u_\infty > -\infty\}$.*

5.2.2. *The following statements are equivalent.*

(i) *f^∞ is optimal.*

(ii) *$\mathbf{L}_f u^* = u^*$ and $\overline{\lim}_n \tilde{\mathbf{L}}_f^n u^* \leqq 0$.*

(iii) *$\mathbf{L}_f u^* = u^*$ and $\lim_n \tilde{\mathbf{L}}_f^n u^* = 0$ on $\{s; u^* > -\infty\}$.*

Proof. Since the two parts of this theorem have similar proofs, only that of Part 1 is given here. The implication "(i) ⇒ (ii)" is a consequence of (3.3) and the fact that $\tilde{\mathbf{L}}_f^n I(f^\infty) \leqq \tilde{\mathbf{L}}_f^n I_+(f^\infty) \to 0$. The implication "(i) ⇒ (iii)" follows from (3.3), Lemma 5.1 and (2.4). For a proof of "(ii) ⇒ (i)" we again use Lemma 5.1 and (2.4) and we obtain $u_\infty \leqq I(f^\infty)$. On the other hand we have by (2.5) $I(f^\infty) \leqq u^* \leqq u_\infty$. For a proof of "(iii) ⇒ (i)" we use the same arguments and obtain $I(f^\infty) = u^* = u_\infty$ on $\{s; u_\infty > -\infty\}$. On $\{s; u_\infty = -\infty\}$ these identities trivially hold. □

As a consequence of Theorem 5.2 and Theorem 4.4.3 combined with (2.2) and (2.5) we obtain

Theorem 5.3. *Assume* (C) *and let be $f \in D^S$.*

5.3.1. *f^∞ is optimal and $u_\infty = u^*$ if and only if $\mathbf{L}_f u_\infty = u_\infty$.*

5.3.2. *f^∞ is optimal if and only if $\mathbf{L}_f u^* = u^*$.*

Remark 5.4. By Theorem 4.4.2 and (3.6) we have for any $f \in D^S$ (at least if (C) hold): $\mathbf{L}_f u_\infty \leqq u_\infty$, $\mathbf{L}_f u^* \leqq u^*$.

Remark 5.5. Dubins and Savage in Chapter 3 of [4] gave necessary and sufficient conditions that a strategy be optimal. In their terminology optimality is equivalent to being "*thrifty*" and "*equalizing*". Their results were applied by Blackwell to dynamic programming in [3] and were extended somewhat by Sudderth in [22]. Roughly, a strategy or policy is "thrifty" if it (almost) always selects actions which achieve the supremum in the optimality equation. It is "equalizing" if it ultimately forces the system into states from which little future gain can be made. Condition (C) of the present paper guarantees that all policies are equalizing. Thus, in view of (3.2), Theorem 5.2 and Theorem 5.3 are close cousins to Theorem 3.9.6 in [4].

6. The Sets of Optimal Actions

Let us write

$$\Gamma^*(s)=\{a\in D(s);\, \mathbf{L}u^*(s,a)=\sup_{a'\in D(s)} \mathbf{L}u^*(s,a')\}.$$

Then $a\in\Gamma^*(s)$ if and only if $\mathbf{L}u^*(s,a)=\mathbf{U}u^*(s)=u^*(s)$. Thus we can rewrite Theorem 5.3.2 as

Corollary 6.1. *Assume* (C) *and let* $f\in D^S$. *Then* f^∞ *is optimal if and only if* $f(s)\in\Gamma^*(s)$, $s\in S$.

The corresponding result for a model with finite horizon is the following where we write

$$\Gamma_n(s)=\{a\in D(s);\, \mathbf{L}u_{n-1}(s,a)=\sup_{a'\in D(s)} \mathbf{L}u_{n-1}(s,a')\}.$$

Corollary 6.2. *Let be* $f_1,\dots,f_n\in D^S$, $n\in\mathbb{N}$. *Then*

$$I_m((f_m,f_{m-1},\dots,f_1))=u_m,\qquad m=1,\dots,n$$

if and only if

$$f_m(s)\in\Gamma_m(s),\qquad s\in S,\ m=1,\dots,n.$$

Corollary 6.2 can be proved by induction or, since any model with finite horizon is a special non-stationary model, it may be derived from Theorem 5.3 by use of the transformation described in § 8 below.

$\Gamma^*(s)$ may be regarded as the set of optimal actions for the model with infinite horizon if the system is in state s. $\Gamma_n(s)$ may be interpreted as the set of optimal actions if the system is in state s and we terminate after n periods. In order to describe the behaviour of Γ_n as $n\to\infty$, let us define for any sequence of subsets C_n of A (cp. Kuratowski [9])

(6.1) $\mathrm{Ls}_n\, C_n=\{a\in A;\, a$ is accumulation point of some sequence (a_n) with $a_n\in C_n,\ n\in\mathbb{N}\}$ [7]

and

(6.2) $$\Gamma_\infty(s)=\mathrm{Ls}_n\,\Gamma_n(s),\qquad s\in S.$$

In particular, $\mathrm{Ls}_n\{a_n\}$ is the set of all accumulation points of the sequence (a_n) where $a_n\in A$, $n\in\mathbb{N}$.

7. The Basic Statements

The purpose of the present paper is to give sufficient conditions for the following statements.

Statement I $u_\infty(s)=u^*(s)$, $s\in S$.

Statement II $\Gamma_\infty(s)\subset\Gamma^*(s)$, $s\in S$.

[7] This definition makes sense for an arbitrary topological space A.

Statement III *There is a stationary optimal policy* f^∞ *with*

$$f(s) \in \Gamma_\infty(s), \qquad s \in S.$$

Referring to (4.1), Statement I implies that under condition (C)

$$u^* = \lim_n \mathbf{U}^n 0,$$

i.e. u^* may be calculated by value iteration. Certainly, another consequence of Statement I is that the optimal total expected reward for the model with finite horizon n tends to the optimal total expected reward for the infinite horizon model. It is known that Statement I is true in the discounted and in the positive case (or more generally in case (C^-) in the terminology of Hinderer [7], cp. ibid. Theorem 3.5). In the negative case, however, Statement I may fail as has been shown by Strauch [21] (cp. ibid. Example 6.1 or Example 7.1 below).

Statement II implies that the actions which are optimal for the model with finite horizon n may approximately be regarded as optimal for the infinite horizon model if n is large. If $D(s)$ is compact, then $\Gamma_\infty(s) \subset \Gamma^*(s)$ is equivalent to the following statement. For any neighbourhood G of $\Gamma^*(s)$, there is some $n_0 \in \mathbb{N}$ such that $\Gamma_n(s) \subset G$ for $n \geqq n_0$.

Now suppose that for any $s \in S$ $D(s)$ is compact and for any $m \in \mathbb{N}$ there is some $f_m \in D^S$ such that the n-stage Markov policy $(f_n, f_{n-1}, \ldots, f_1)$ is optimal with respect to the model with horizon n for $n \in \mathbb{N}$, i.e. $f_m(s) \in \Gamma_m(s)$, $s \in S$, $m \in \mathbb{N}$. If A is a subset of the real line, then we can conclude from Statement II that f^∞ where $f = \underline{\lim}\, f_n$ or $f = \overline{\lim}\, f_n$ is optimal for the infinite horizon model. If there is one and only one optimal policy f^∞ for the infinite horizon model, then $f(s)$ is the only element of $\Gamma^*(s)$ and Statement II implies that $\lim_n f_n(s) = f(s)$, $s \in S$.

Example 7.1. Let there be given a sequence of functions (δ_n) on A such that $\delta_0 = 0 = \delta_1 \geqq \delta_2 \geqq \delta_3 \geqq \cdots$ and choose $S = A \times \mathbb{N}$, $D(s) = A$, $s \in S$, $q((t,1), a; \{(a,2)\}) = 1$, $q((t,k), a; \{(t, k+1)\}) = 1$, $t \in A$, $k \geqq 2$, $\beta = 1$,

$$r(s, a, s') = r(s) = \delta_k(t) - \delta_{k-1}(t) \qquad \text{for } s = (t,k) \in S.$$

It is clear from these definitions that $\Gamma_n(s) = \Gamma^*(s) = A$ and $u_\infty(s) = u^*(s)$ for $s = (t,k)$ where $k \geqq 2$. Hence, the only interesting states are $S_1 = \{(t,1); t \in A\}$. A little consideration shows that for $s \in S_1$

$$u_n(s) = \sup_{a \in A} \delta_n(a), \qquad \Gamma_n(s) = \{a; \delta_n(a) = \sup_{a' \in A} \delta_n(a')\},$$

$$u^*(s) = \sup_{a \in A} \delta_\infty(a), \qquad \Gamma^*(s) = \{a; \delta_\infty(a) = \sup_{a' \in A} \delta_\infty(a')\}$$

where $\delta_\infty = \lim \delta_n$.

Now set $A = [0,2]$, $\delta_n(a) = b$ for $a = 0$ and $1/n < a < 1$, $\delta_n(a) = c_n$ for $0 < a \leqq 1/n$, $\delta_n(a) = d$ for $1 \leqq a \leqq 2$, $n \geqq 2$, where b, c_n, d are real numbers with $0 \geqq c_n \geqq c_{n+1} > d \geqq b$. Set $c_\infty = \lim c_n$. Then for $s \in S_1$ $u_n(s) = c_n$, $u_\infty(s) = c_\infty$, $\Gamma_n(s) = (0, 1/n]$, $\Gamma_\infty(s) = \{0\}$, $\delta_\infty(a) = b$ for $0 \leqq a < 1$, $\delta_\infty(a) = d$ for $1 \leqq a \leqq 2$, $u^*(s) = d$. If $b < d$ then $\Gamma^*(s) = [1,2]$, if $b = d$ then $\Gamma^*(s) = [0,2]$. Further, $u_\infty = u^*$ if and only if $c_\infty = d$. Thus the following four cases are possible.

Case 1: Statement I fails, Statement II fails ($c_\infty > d > b$).

Case 2: Statement I holds, Statement II fails ($c_\infty = d > b$).

Case 3: Statement I fails, Statement II holds ($c_\infty > d = b$).

Case 4: Statement I holds, Statement II holds ($c_\infty = d = b$).

The next remark concerns Statement III. If the infinite horizon model is used as approximation of a model with finite horizon n where n is large, then only those optimal policies f^∞ may be admitted that satisfy the condition $f(s) \in \Gamma_\infty(s)$, $s \in S$. Finally we note that Statement III implies that u^* is Borel-measurable since $I(\pi)$ is Borel-measurable for any $\pi \in \Delta$.

8. Non-Stationary Decision Models

A non-stationary Markovian decision model is given by a tupel $((S_n, \mathfrak{S}_n), (A_n, \mathfrak{A}_n), D_n, q_n, \beta_n, r_n; n \in \mathbb{N})$ of the following meaning: S_n and A_n stand for the state space and the action space at time n, respectively. D_n specifies the set of all admissible actions at time n. $q_n(s, a; \cdot)$ is the distribution of ζ_{n+1} given $\zeta_n = s$, $\alpha_n = a$. β_n and r_n are the discount factor and the reward function for period n, respectively. Given the history $(s_1, a_1, \ldots, s_{n+1}) \in S_1 \times A_1 \times \cdots \times S_{n+1}$, one will receive for period n the discounted reward $\beta_1(s_1, a_1, s_2) \ldots \beta_{n-1}(s_{n-1}, a_{n-1}, s_n) \, r_n(s_n, a_n, s_{n+1})$.

The purpose of this section is to show that every non-stationary problem can be reformulated so as to be stationary (cp. also Dubins and Savage [4] Chap. 12.2).

For a given non-stationary model define $S = \{(s, n); s \in S_n, n \in \mathbb{N}\}$, $\mathfrak{S} = \{B \subset S; \{s; (s, n) \in B\} \in \mathfrak{S}_n$ for all $n \in \mathbb{N}\}$, $A = \{(a, n); a \in A_n, n \in \mathbb{N}\}$, $\mathfrak{A}$ through $(\mathfrak{A}_n)$ as $\mathfrak{S}$ through $(\mathfrak{S}_n)$, $D((s, n)) = D_n(s) \times \{n\}$, $\beta((s, n)), (a, n), (s', n')) = \beta_n(s, a, s')$ for $n' = n + 1$ and $= 0$ otherwise, $r((s, n), (a, n), (s', n')) = r_n(s, a, s')$ for $n' = n + 1$ and $= 0$ otherwise, and let $q((s, n), (a, n); \cdot)$ assign probability one to $S_{n+1} \times \{n+1\}$ with the marginal distribution of the first coordinate being $q_n(s, a; \cdot)$ for $(s, a) \in K$.

Stationary policies (resp. optimal stationary policies) in this stationary model correspond in an obvious way to Markov policies (resp. strongly optimal Markov policies in the sense of Hinderer [6] p. 132) in the original model.

For readers which are particularly interested in non-stationary decision models, non-stationary versions of the present paper are available for distribution.

9. Set-Valued Mappings

Throughout this section, $(S, \mathfrak{S})$ is allowed to be any measurable space and (A, ρ) may be any separable metric space. We use $\mathfrak{C}(A)$ to denote the set of all non-empty compact subsets of A. One may introduce a metric on $\mathfrak{C}(A)$ – the Hausdorff metric – that is for any $C, C' \in \mathfrak{C}(A)$

$$d(C, C') = \max(\sup_{c \in C} \rho(c, C'), \sup_{c' \in C'} \rho(c', C))$$

(cp. Kuratowski [9], 21 VII). We write $C_n \to C$ if $d(C_n, C) \to 0$. The Hausdorff metric was probably first used in stochastic optimization problems by Dubins and Savage in Section 2.16 of [4].

Proposition 9.1. 9.1.1. $(\mathfrak{C}(A), d)$ *is a separable metric space. (If A' is a countable dense subset of A, then the set of all finite subsets of A' is dense in $\mathfrak{C}(A)$.)*

9.1.2. *If (A, ρ) is locally compact, so is $(\mathfrak{C}(A), d)$.*

The *proof* follows from Michael [12] Theorems 3.3, 3.6, 4.5, 4.9.

A mapping $\varphi: S \to \mathfrak{C}(A)$ is called *measurable* if it is measurable with respect to the σ-algebra of Borel subsets of $(\mathfrak{C}(A), d)$.

A mapping $\varphi: S \to \mathfrak{C}(A)$ is called *separable* (with separating set A') if

(i) A contains a countable dense subset A' such that $A' \cap \varphi(s)$ is dense in $\varphi(s)$, $s \in S$.

(ii) $\{s;\ \varphi(s) \ni a\} \in \mathfrak{S}, \quad a \in A$.

Proposition 9.2. *If $\varphi: S \to \mathfrak{C}(A)$ is measurable, then $\{(s, a);\ a \in \varphi(s)\}$ is a (product-) measurable subset of $S \times A$.*

The *proof* can be found in Furukawa [5] Lemma 3.1.

Proposition 9.3. *Every separable mapping $\varphi: S \to \mathfrak{C}(A)$ (with separating set A') is measurable.*

For a *proof* it is sufficient to observe that

$$\{s;\ d(\varphi(s), C) < \delta\} = \bigcup_{a \in A',\ \rho(a, C) < \delta} \{s, a \in \varphi(s)\} \quad \text{for } C \in \mathfrak{C}(A), \quad \delta > 0.$$

Proposition 9.4. *$\varphi: S \to \mathfrak{C}(A)$ is measurable if and only if φ is the limit of a sequence of separable mappings.*

Proof. The "if" direction can be deduced from Proposition 9.3 above and from Kuratowski [9] 31 VIII Theorem 1. Now suppose that φ is measurable. From Proposition 9.1.1 and Kuratowski [9] 31 VIII Theorem 3, we conclude that there is a sequence of measurable mappings $\varphi_n: S \to \mathfrak{C}(A)$ taking on only countably many values with $\varphi_n(s) \to \varphi(s)$, $s \in S$. By virtue of the separability of A and Proposition 9.2, it is easily seen that the mappings φ_n are separable. □

For the remainder of this section, suppose that S is a topological space and $\mathfrak{S}$ is the σ-algebra of Borel subsets of S.

A mapping $\varphi: S \to \mathfrak{C}(A)$ is called *upper semi-continuous* (*u.s.c.*), if for each open $G \subset A$ the set $\{s;\ \varphi(s) \subset G\}$ is open in S[8] (cp. Kuratowski [9] 18 I).

We note that φ is u.s.c. if and only if for any $s \in S$ and for any open G containing $\varphi(s)$ there is a neighbourhood H of s such that $\varphi(H) \subset G$ (cp. Kuratowski [9] 18 III Theorem 3).

Proposition 9.5. *Suppose that A is a locally compact separable metric space and $\varphi: S \to \mathfrak{C}(A)$ is u.s.c. Then φ is measurable.*

Proof. By the one-point-compactification theorem of Alexandrov, we may assume A to be an open subset of some compact metric space $\hat{A}$. Then $\mathfrak{C}(A)$ is an open subset of $\mathfrak{C}(\hat{A})$ and φ may be regarded as an u.s.c. mapping $\varphi: S \to \mathfrak{C}(\hat{A})$. From Kuratowski [10] 43 VII Theorem 1 we know that $\varphi: S \to \mathfrak{C}(\hat{A})$ is measurable. Hence $\varphi: S \to \mathfrak{C}(A)$ is measurable. □

Remark 9.6. Proposition 9.5 remains true if the condition that A is locally compact is replaced by the following: S admits a measurable partition $S = \bigcup S_n$ such that $\bigcup_{s \in S_n} \varphi(s)$ is relatively compact.

---

[8] This definition makes sense for an arbitrary topological space A.

10. Upper Semi-Continuous Functions

Throughout this section, S and A are allowed to be any topological spaces. We use $\mathscr{C}(A)$ to denote the set of all bounded continuous functions $u: A \to \mathbb{R}$ and $\hat{\mathscr{C}}(A)$ to denote the set of all upper semicontinuous functions $u: A \to \underline{\mathbb{R}}$ which are bounded from above. If $v \in \hat{\mathscr{C}}(A)$ and A is countably compact[9], then v attains its supremum on A. If A is a metric space, then we know from the theorem of Baire that $v \in \hat{\mathscr{C}}(A)$ if and only if v is the limit of some non-increasing sequence of functions $v_n \in \mathscr{C}(A)$. The following result is basic for this paper.

Proposition 10.1. *Let (w_n) be a sequence of functions $w_n \in \hat{\mathscr{C}}(A)$ and let (ε_n) be a sequence of non-negative numbers converging to zero such that $w_n(a) \leqq w_m(a) + \varepsilon_m$, $m \leqq n$, $a \in A$.*[10] *Then*

10.1.1. $w_\infty = \lim w_n$ *exists and* $w_\infty \in \hat{\mathscr{C}}(A)$,

10.1.2. $\mathrm{Ls}_n\, W_n \subset W_\infty$ *where* $W_n = \{a \in A;\ w_n(a) = \sup_A w_n\}$, $n \in \mathbb{N}$ or $n = \infty$,

10.1.3. $\lim_n \sup_A w_n = \sup_A \lim_n w_n$ *provided that A is countably compact or more generally that* $\mathrm{Ls}\, W_n$ *is not empty.*

Proof. The existence of w_∞ follows from Lemma 4.1. From the relation

$$\{a;\ w_\infty(a) < r\} = \bigcup_{n \in \mathbb{N}} \{a;\ w_n(a) + \varepsilon_n < r\}, \qquad r \in \mathbb{R},$$

we infer that $w_\infty \in \hat{\mathscr{C}}(A)$. Since $\sup w_n \leqq \sup w_m + \varepsilon_m$, $m \leqq n$, we further know that $\lim_n \sup_A w_n$ exists and $\lim_n \sup_A w_n \geqq \sup_A \lim_n w_n = \sup_A w_\infty$. Now suppose that $a_0 \in \mathrm{Ls}\, W_n$, i.e. there is some sequence (a_n) with $a_n \in W_n$ such that a_0 is an accumulation point of (a_n). Without loss of generality we may assume that $\sup_A w_\infty > -\infty$. Otherwise we have $W_\infty = A$. Now choose some $M < \lim_n \sup_A w_n$. Then $\sup_A w_n > M$, $n \geqq n_0$, for some $n_0 \in \mathbb{N}$. Hence $W_n \subset \{a;\ w_n(a) \geqq M\}$, $n \geqq n_0$. For $n \geqq m \geqq n_0$ we have $w_m(a_n) \geqq w_n(a_n) - \varepsilon_m \geqq M - \varepsilon_m$. Since $\{a \in A;\ w_m(a) \geqq M - \varepsilon_m\}$ is closed, we infer that $w_m(a_0) \geqq M - \varepsilon_m$. From the obvious passage to the limit we obtain $w_\infty(a_0) \geqq M$ for any $M < \lim_n \sup_A w_n$. Thus $w_\infty(a_0) \geqq \lim_n \sup_A w_n \geqq \sup_A w_\infty \geqq w_\infty(a_0)$ and we realize that equality holds throughout. □

We need the following generalization of results by Dubins and Savage [4] Lemma 2.16.7, Maitra [11] Lemma 3.4, Hinderer [6] Lemma 5.10.

Proposition 10.2. *Let S and A be any topological spaces and $\varphi: S \to \mathfrak{C}(A)$ be u.s.c. Let be $v \in \hat{\mathscr{C}}(\Phi)$ where $\Phi = \{(s, a);\ a \in \varphi(s)\}$ is endowed with the relativization of the product topology and set $v^*(s) = \max_{a \in \varphi(s)} v(s, a)$. Then $v^* \in \hat{\mathscr{C}}(S)$.*

Proof. We have to show that for a fixed $r \in \mathbb{R}$ and a fixed $s \in V = \{s';\ v^*(s') < r\}$ there is a neighbourhood X of s such that $X \subset V$. Since $v(s, \cdot)$ attains its supremum on $\varphi(s)$, we can rewrite $s \in V$ as $v(s, a) < r$ for $a \in \varphi(s)$. Thus $s \in V$ implies that there are open neighbourhoods H_a and G_a of s and $a \in \varphi(s)$, respectively, such that $v < r$ on $H_a \times G_a \cap \Phi$. As $\varphi(s)$ is compact there is a finite subset F of $\varphi(s)$ such that $\varphi(s) \subset \bigcup_{a \in F} G_a$. Set $G = \bigcup_{a \in F} G_a$, $H_F = \bigcap_{a \in F} H_a$, $H_G = \{s';\ \varphi(s') \subset G\}$, and $X = H_F \cap H_G$. From the semi-continuity of φ we conclude that H_G and hence X are open. Now $v < r$ on $X \times G \cap \Phi = X \times A \cap \Phi$. Thus $X \subset V$. □

---

[9] For a definition cp. Ash [1] p. 384.

[10] If (w_n) is non-increasing, we may choose $\varepsilon_n = 0$. If (w_n) is uniformly convergent, we may choose $\varepsilon_n = \sup_{m \geqq n} \|u_m - u_n\|$.

11. The Set of Functions $\mathscr{Q}(S \times A)$

Throughout this section, $(S, \mathfrak{S})$ may be any measurable space and (A, ρ) may be any separable metric space endowed with the σ-algebra of Borels subsets of A. We use $\mathscr{B}(S)$ to denote the set of all bounded measurable functions u: $S \to \mathbb{R}$ and $\hat{\mathscr{B}}(S)$ to denote the set of all measurable functions u: $S \to \mathbb{R}$ which are bounded from above. Further, we introduce the following sets of functions

(11.1) $\mathscr{Q}(S \times A) = \{v \in \mathscr{B}(S \times A);\ v(s, \cdot) \in \mathscr{C}(A),\ s \in S\}$,

(11.2) $\hat{\mathscr{Q}}(S \times A) = \{v \in \hat{\mathscr{B}}(S \times A)$; v is the limit of some non-increasing sequence of functions $v_n \in \mathscr{Q}(S \times A)\}$.

Remark 11.1. If $u \in \hat{\mathscr{Q}}(S \times A)$, then $u(s, \cdot) \in \hat{\mathscr{C}}(A)$, $s \in S$.

Remark 11.2. If S is a metric space then the theorem of Baire implies that $\hat{\mathscr{C}}(S \times A) \subset \hat{\mathscr{Q}}(S \times A)$.

Remark 11.3. If (v_n) is a non-increasing sequence of functions $v_n \in \hat{\mathscr{Q}}(S \times A)$, then $\lim v_n \in \hat{\mathscr{Q}}(S \times A)$.

Remark 11.4. Since A is separable, it may be shown that $u \in \mathscr{Q}(S \times A)$ if and only if u is bounded and $u(\cdot, a) \in \mathscr{B}(S)$, $a \in A$, $u(s, \cdot) \in \mathscr{C}(A)$, $s \in S$.

Remark 11.5. If u: $S \times A \to \underline{\mathbb{R}}$ is bounded from above and if

$$u(\cdot, a) \in \hat{\mathscr{B}}(S), \qquad a \in A, \tag{11.3}$$

$$\overline{\lim}_{a' \to a} u(s, a') = \overline{\lim}_{a' \to a,\, a' \in A'} u(s, a') = u(s, a), \qquad s \in S,\ a \in A, \tag{11.4}$$

where A' is some countable dense subset of A, then $u \in \hat{\mathscr{Q}}(S \times A)$. This fact follows from the proof of

Proposition 11.6. *Let φ: $S \to \mathfrak{C}(A)$ be a separable mapping with separating set A' and set $\Phi = \{(s, a) \in S \times A;\ a \in \varphi(s)\}$. Let u: $\Phi \to \underline{\mathbb{R}}$ be bounded from above. Suppose that $u(\cdot, a)$ is measurable on $\{s;\ \varphi(s) \ni a\}$, $a \in A$, and*

$$\overline{\lim}_{\substack{a' \to a \\ a' \in \varphi(s)}} u(s, a') = \overline{\lim}_{\substack{a' \to a \\ a' \in A' \cap \varphi(s)}} u(s, a') = u(s, a), \qquad a \in \varphi(s),\ s \in S.$$

Then u admits an extension $\hat{u} \in \hat{\mathscr{Q}}(S \times A)$. Such an extension is given by setting $\hat{u} = -\infty$ on $S \times A - \Phi$.

Proof. Set $\hat{u} = u$ on Φ and $\hat{u} = -\infty$ on $S \times A - \Phi$. Then $\hat{u}$ has the properties (11.3) and (11.4). As there is an order-preserving homeomorphism of $\underline{\mathbb{R}}$ and $[0, 1)$, it is sufficient to show that any function $\hat{u}$: $S \times A \to [0, 1)$ enjoying the properties (11.3) and (11.4) is an element of $\hat{\mathscr{Q}}(S \times A)$. Set $v_n(s, a) = \sup_{a' \in A} \{\hat{u}(s, a') - n\rho(a', a)\}$. As, by (11.4), $\hat{u}(s, \cdot) \in \hat{\mathscr{C}}(A)$, $s \in S$, we know from the proof of the theorem of Baire (cp. Ash [1] p. 390) that $v_n \geqq v_{n+1}$, $v_n(s, \cdot) \in \mathscr{C}(A)$, $\lim v_n = \hat{u}$. From (11.4) we conclude that $v_n(s, a) = \sup_{a' \in A'} \{\hat{u}(s, a') - n\rho(a', a)\}$. Thus, by (11.3), $v_n \in \mathscr{B}(S \times A)$ which implies that $v_n \in \mathscr{Q}(S \times A)$. □

12. Selection Theorem

In this section, we shall list some results proved in [18]. Again, $(S, \mathfrak{S})$ may be any measurable space and A may be any separable metric space.

Theorem 12.1. *Let be $u \in \hat{\mathfrak{Q}}(S \times A)$ and let φ: $S \to \mathfrak{C}(A)$ be measurable. Then there is a measurable mapping f: $S \to A$ such that $f(s) \in \varphi(s)$ and*

$$u(s, f(s)) = \max_{a \in \varphi(s)} u(s, a), \qquad s \in S.$$

In view of Proposition 9.4, this result is Theorem 2 in [18].

Proposition 12.2. *Let (f_n) be a sequence of measurable mappings f_n: $S \to A$ and φ: $S \to \mathfrak{C}(A)$ any mapping with $f_n(s) \in \varphi(s)$, $s \in S$. Then there is measurable mapping f: $S \to A$ such that*

$$f(s) \in \mathrm{Ls}_{n \to \infty} \{f_n(s)\} \subset \varphi(s), \qquad s \in S.$$

This is Lemma 4 in [18].

13. Proof of the Basic Statements

The proof of Statements I–III can be carried through with aid of

Condition (A).

(A 1) *$D(S) \subset \mathfrak{C}(A)$ and D: $S \to \mathfrak{C}(A)$ is measurable,*

(A 2) *$\mathbf{L} u_{n-1}$ admits an extension $v_n \in \hat{\mathfrak{Q}}(S \times A)$, $n \in \mathbb{N}$.*

Remark 13.1. The measurability condition in (A 1) implies that the condition (2.1) imposed in the definition of the decision model is satisfied. This may be seen from Proposition 9.2 and Theorem 12.1.

In Section 15 and 16 we shall give sufficient conditions (condition (S) and condition (W)) for (A) about D, q, β, and r.

Lemma 13.2. *Assume* (A). *There exist $f_n \in D^S$, $n \in \mathbb{N}$, such that $f_n(s) \in \Gamma_n(s)$, $s \in S$, $n \in \mathbb{N}$, i.e. $(f_n, \ldots, f_1)$ is an optimal n-stage Markov policy for the decision model with horizon n, $n \in \mathbb{N}$.*

Proof. From Theorem 12.1 we know that there are $f_n \in D^S$ such that $v_n(s, f_n(s)) = \max_{a \in D(s)} v_n(s, a)$. Because of $v_n|_K = \mathbf{L} u_{n-1}$, it will be found that $f_n(s) \in \Gamma_n(s)$, $s \in S$, $n \in \mathbb{N}$. In view of Corollary 6.2, the proof is complete. □

Theorem 13.3. *The basic Statements* I, II, III *are valid under conditions* (C) *and* (A).

Proof. For an arbitrary fixed $s \in S$, set $w_n(a) = \mathbf{L} u_{n-1}(s, a)$, $a \in D(s)$. Then by condition (A 2), $w_n \in \hat{\mathscr{C}}(D(s))$. Further, we infer that $w_n \leqq w_m + z_m(s)$, $m \leqq n$, from the following inequalities

$$\begin{aligned} \mathbf{L} u_{n-1} &\leqq \mathbf{L}(u_{m-1} + u_{m,n-1}) \leqq \mathbf{L} u_{m-1} + \tilde{\mathbf{U}} u_{m,n-1} \\ &= \mathbf{L} u_{m-1} + u_{m+1,n} \leqq \mathbf{L} u_{m-1} + z_m \end{aligned}$$

where use is made of (2.3), (3.1), (3.9), and (3.8). Now Proposition 10.1.3 applies and we obtain

$$\sup_{a \in D(s)} \lim_n w_n(a) = \lim_n \sup_{a \in D(s)} w_n(a). \tag{13.1}$$

From (3.7) and Theorem 4.4.1 we know that

$$\sup_{a \in D(s)} w_n(a) = \mathbf{U} u_{n-1}(s) = u_n(s),$$

$$\lim_n w_n(a) = \mathbf{L} u_\infty(s, a), \qquad a \in D(s).$$

Now we can rewrite (13.1) as

$$\mathbf{U}u_\infty(s)=u_\infty(s). \tag{13.2}$$

Hence an appeal to Proposition 10.1.2 shows that

$$\mathrm{Ls}_n\,\Gamma_n(s)\subset\{a\in D(s);\,\mathbf{L}u_\infty(s,a)=u_\infty(s)\}. \tag{13.3}$$

Let $\Gamma'(s)$ denote the set on the right-hand side of (13.3). By Lemma 13.2, there are $f_n\in D^S$, $n\in\mathbb{N}$, with $f_n(s)\in\Gamma_n(s)$, $s\in S$, $n\in\mathbb{N}$. In view of Proposition 12.2, there exists some $f\in D^S$ with

$$f(s)\in\mathrm{Ls}_n\{f_n(s)\}\subset\mathrm{Ls}_n\,\Gamma_n(s)=\Gamma_\infty(s),\qquad s\in S. \tag{13.4}$$

Now, appealing to (13.3), we have $\mathbf{L}_f u_\infty=u_\infty$. Combining this result with Theorem 5.3.1, Statement I and the first part of Statement III are proved. The second part follows from (13.4). From Statement I we conclude that $\Gamma'=\Gamma^*$. Now Statement II is a consequence of (13.3). □

14. Topologies on Spaces of Probabilities

Throughout this section, $(S,\mathfrak{S})$ may be any measurable space and A may be any topological space, unless otherwise indicated. $\mathscr{P}(S)$, the set of probability measures on $\mathfrak{S}$, may be endowed with the s-topology or with the w-topology (weak topology) in the following ways.

A mapping μ: $A\to\mathscr{P}(S)$ will be called s-continuous if $a\to\int v(s)\,\mu(a;ds)$ is continuous for any $v\in\mathscr{B}(S)$.

Proposition 14.1. *Assume A to be a separable metric space. If $v\in\mathscr{Q}(S\times A)$ (resp. $\hat{\mathscr{Q}}(S\times A)$) and if μ: $A\to\mathscr{P}(S)$ is s-continuous, then the function*

$$a\to\int v(s,a)\,\mu(a;ds)$$

is an element of $\mathscr{C}(A)$ (resp. $\hat{\mathscr{C}}(A)$).

Now suppose that S is a topological space and $\mathfrak{S}$ is the system of Borel subsets of S.

A mapping μ: $A\to\mathscr{P}(S)$ will be called w-continuous if $a\to\int v(s)\,\mu(a;ds)$ is continuous for any $v\in\mathscr{C}(S)$.

Proposition 14.2. *Assume S and A to be separable metric spaces. If $v\in\mathscr{C}(S\times A)$ (resp. $\hat{\mathscr{C}}(S\times A)$) and if μ: $A\to\mathscr{P}(S)$ is w-continuous, then the function*

$$a\to\int v(s,a)\,\mu(a,ds)$$

is an element of $\mathscr{C}(A)$ (resp. $\hat{\mathscr{C}}(A)$).

Proofs of Proposition 14.1 and 14.2 can be found in [19].

15. The Conditions (S)

We define $\mathscr{Q}(S\times A\times S)$ and $\hat{\mathscr{Q}}(S\times A\times S)$ through (11.1) and (11.2) on replacing S with $S\times S$. For functions as r that are defined on $K\times S$ only, we had to modify this definition as follows.

$\mathscr{Q}(K\times S)=\{v\in\mathscr{B}(K\times S);\ v(s,\cdot,s')\in\mathscr{C}(D(s)),\ s,s'\in S\}$

$\hat{\mathscr{Q}}(K\times S)=\{v\in\hat{\mathscr{B}}(K\times S);\ v_n\downarrow v$ for some sequence (v_n) with $v_n\in\mathscr{Q}(K\times S),\ n\in\mathbb{N}\}$. Now we can introduce

Condition (S).

(S 1) *$D(S)\subset\mathfrak{C}(A)$, $D\colon S\to\mathfrak{C}(A)$ is separable,*

(S 2) *$q(s,\cdot)\colon D(s)\to\mathscr{P}(S)$ is s-continuous, $s\in S$,*

(S 3) *$r\in\hat{\mathscr{Q}}(K\times S)$, $\beta\in\mathscr{Q}(K\times S)$.*

Condition (S)′ (cp. *Furukawa* [5]).

(S 1)′ (A 1) *is satisfied.*

(S 2)′ *q admits an extension $\hat{q}\colon S\times A\to\mathscr{P}(S)$ such that $\hat{q}(s,\cdot)\colon A\to\mathscr{P}(S)$ is s-continuous $s\in S$.*

(S 3)′ *r admits an extension $\hat{r}\in\hat{\mathscr{Q}}(S\times A\times S)$ and β admits an extension $\hat{\beta}\in\mathscr{Q}(S\times A\times S)$ with $\hat{\beta}\geq 0$.*

Lemma 15.1. *Suppose that either condition* (S) *or condition* (S)′ *is satisfied and that $u\in\hat{\mathscr{B}}(S)$. Then*

15.1.1. $\mathbf{L}u$ *admits an extension $v\in\hat{\mathscr{Q}}(S\times A)$,*

15.1.2. $\mathbf{U}u\in\hat{\mathscr{B}}(S)$.

Proof. Assume (S). By hypothesis there are functions $r_m\in\mathscr{Q}(K\times S)$ with $r_m\downarrow r$ as $m\to\infty$. Upon setting $w_m=\max(u,-m)$, we have $w_m\in\mathscr{B}(S)$ and $w_m\downarrow u$ as $m\to\infty$. Define now

$$\mathbf{L}_m w_m(s,a)=\int q(s,a;ds')[r_m(s,a,s')+\beta(s,a,s')\,w_m(s')],\qquad (s,a)\in K.$$

Then $\mathbf{L}_m w_m\in\mathscr{B}(K)$. From Proposition 14.1 we know that $\mathbf{L}_m w_m(s,\cdot)\in\mathscr{C}(D(s))$, $s\in S$. Upon setting $v_m=\mathbf{L}_m w_m$ on K and $v_m=-\infty$ on $S\times A-K$, we may infer from Proposition 11.6 that $v_m\in\hat{\mathscr{Q}}(S\times A)$. Now $v_m\downarrow v$ for some $v\in\hat{\mathscr{Q}}(S\times A)$ (cp. Remark 11.3) where $v|_K=\mathbf{L}u$. From these arguments it is clear that under condition (S)′ $\hat{\mathbf{L}}u=\int d\hat{q}[\hat{r}+\hat{\beta}u]$ is an extension of $\mathbf{L}u$ and that $\hat{\mathbf{L}}u\in\hat{\mathscr{Q}}(S\times A)$. In this case no appeal to Proposition 11.6 has to be made. Finally, 15.1.2 is a special consequence of Theorem 12.1. □

Theorem 15.2. *If condition* (C) *and either condition* (S) *or condition* (S)′ *are satisfied, then the basic Statements* I, II, III *are valid.*

Proof. In view on Theorem 13.3, we have to show that condition (A) is satisfied. By Proposition 9.3 it is clear that (A 1) is satisfied. From (3.7) and Lemma 15.1.2 it may be shown inductively that $u_n\in\hat{\mathscr{B}}(S)$. Now condition (A 2) follows from Lemma 15.1.1. □

16. The Condition (W)

In this section we shall make use of the following

Condition (W) (cp. *Maitra* [11], *Hinderer* [6] *Theorem* 17.12).

(W 1) *$D(S)\subset\mathfrak{C}(A)$, $D\colon S\to\mathfrak{C}(A)$ is u.s.c.,* (W 2) *$q: K \to \mathscr{P}(S)$ is w-continuous,*

(W 3) *$r \in \hat{\mathscr{C}}(K \times S)$, $\beta \in \mathscr{C}(K \times S)$,*

(W 4) *A is locally compact.*[11]

From (W 1) *it follows that* (cp. *Kuratowski* [9] 18 III *Theorem* 1)

(16.1) $$K \text{ is closed in } S \times A.$$

Lemma 16.1. *Assume* (W) *and let be $u \in \hat{\mathscr{C}}(S)$. Then*

16.1.1. $\mathbf{L}u$ *admits an extension $v \in \hat{\mathscr{C}}(S \times A)$,*

16.1.2. $\mathbf{U}u \in \hat{\mathscr{C}}(S)$.

Proof. From Proposition 14.2 it will be seen that $\mathbf{L}u \in \hat{\mathscr{C}}(K)$. Upon setting $v = \mathbf{L}u$ on K and $v = -\infty$ on $S \times A - K$, we have by (16.1) $v \in \hat{\mathscr{C}}(S \times A)$. Finally, 16.1.2 is a consequence of Theorem 10.2. □

Theorem 16.2. *If condition* (C) *and condition* (W) *are satisfied, then the basic Statements* I, II, III *are valid.*

Proof. We have to verify condition (A). (A 1) derives from Proposition 9.5. In view of (3.7) and Lemma 16.1.2, it is easily established inductively that $u_n \in \hat{\mathscr{C}}(S)$. Now an appeal to Lemma 16.1.1 and Remark 11.2 confirms condition (A 2). □

Corollary 16.3. *If $\|z_n\| \to 0$ as $n \to \infty$ and if condition* (W) *is satisfied, then $u^* \in \hat{\mathscr{C}}(S)$.*

Proof. As has been shown in the proof of Theorem 16.2, we have $u_n \in \hat{\mathscr{C}}(S)$. By (2.3) we obtain $u_n \leqq u_m + \|z_m\|$, $m \leqq n$. Now the assertion follows from Proposition 10.1.1 and Statement I. □

17. Non-Compact Sets of Admissible Actions

Whereas the continuity conditions (S 2), (S 3), (S 2)′, (S 3)′, (W 2), (W 3) are likely to include all cases arising from any practical situation, the compactness condition $D(s) \in \mathfrak{C}(A)$ sometimes turns out to be somewhat restrictive. When dealing with single item dynamic inventory models, for instance, we are concerned with sets of admissible actions $D(s) = [s, \infty)$ where s is the stock on hand plus on order. In this section we are going to show that the condition $D(s) \in \mathfrak{C}(A)$ may be replaced by the condition that the "good" actions are contained in some compact set $\underline{D}(s) \subset D(s)$. More exactly, we shall utilize the following

Condition (B). *There is a measurable mapping $\underline{D}: S \to \mathfrak{C}(A)$ such that*

(17.1) $$\underline{D}(s) \subset D(s), \qquad s \in S,$$

(17.2) $$\sup_{a \in D(s) - \underline{D}(s)} \mathbf{L}u_{n-1}(s, a) < \sup_{a \in D(s)} \mathbf{L}u_{n-1}(s, a), \qquad s \in S, \ n \in \mathbb{N}.$$

In view of Remark 13.1, we obtain a new decision model if we replace D with $\underline{D}$. Certainly, this model inherits the general assumption (GA) from the original

[11] cp. Remark 9.6.

model. All quantities referring to the modified decision model will be underlined, e.g.

$$\underline{\mathbf{U}}u = \sup_{a \in \underline{D}(s)} \mathbf{L}u(\cdot, a), \qquad u \in \mathscr{J}.$$

From (17.1) we conclude that $\underline{\Delta} \subset \Delta$, $\underline{D}^S \subset D^S$. Thus

$$\underline{u}^* \leqq u^*. \tag{17.3}$$

Further (17.2) implies that

$$\mathbf{U}u_{n-1} = \underline{\mathbf{U}}u_{n-1}, \qquad n \in \mathbb{N}, \tag{17.4}$$

$$\Gamma_n(s) \subset \underline{D}(s), \qquad n \in \mathbb{N},\ s \in S. \tag{17.5}$$

By induction, we obtain from (17.4) and (3.7)

$$\underline{u}_n = u_n, \qquad n \in \mathbb{N}, \tag{17.6}$$

$$\underline{u}_\infty = u_\infty. \tag{17.7}$$

Finally, by (17.5) and (17.6), it is easily seen that

$$\underline{\Gamma}_n(s) = \Gamma_n(s), \qquad n \in \mathbb{N},\ s \in S, \tag{17.8}$$

$$\underline{\Gamma}_\infty(s) = \Gamma_\infty(s), \qquad s \in S. \tag{17.9}$$

Theorem 17.1. *Assume* (B). *If the Statements* I, II, III *are valid for the modified model, then they are valid for the original model.*

Proof. On making use of (17.3), (17.7) and (2.5), we obtain $u_\infty = \underline{u}_\infty = \underline{u}^* \leqq u^* \leqq u_\infty$. Hence equality holds throughout and we know that Statement I holds for the original model and

$$\underline{u}^* = u^*. \tag{17.10}$$

Further, by (17.9), $\Gamma_\infty(s) = \underline{\Gamma}_\infty(s) \subset \underline{\Gamma}^*(s) \subset \Gamma^*(s)$, where the last inclusion is a consequence of (17.1) and (17.10). Hence Statement II holds for the original model. Finally, for any $f \in \underline{D}^S$ with $I(f^\infty) = \underline{u}^*$, $f(s) \in \underline{\Gamma}_\infty(s)$, $s \in S$, we obtain by (17.9) and (17.10) $I(f^\infty) = u^*, f(s) \in \Gamma_\infty(s), s \in S$. This completes the domonstration. □

For the modified model, we may replace condition (C) by the weaker

Condition ($\underline{\mathbf{C}}$). $\underline{z}_n \to 0$ as $n \to \infty$.

In correspondence with conditions (S 1) and (W 1), we define **Condition (SB 1)** and **Condition (WB 1)** through condition (B) upon replacing "measurable" by "separable" and "u.s.c.", respectively. Now, the following theorem is a corollary of Theorems 15.2, 16.2, and 17.1.

Theorem 17.2. *Assume* ($\underline{\mathrm{C}}$). *The Statements* I, II, III *are valid provided that either conditions* (SB 1), (S 2), (S 3) *or conditions* (B), (S 2)′, (S 3)′ *or* (WB 1), (W 2), (W 3), (W 4) *are satisfied.*

Acknowledgements. The present paper was stimulated by the lectures and papers of Professor K. Hinderer to whom I want to express my gratitude for his encouragement and many valuable discussions. Thanks are also due to the referee for Remark 5.5, for the transformation described in § 8 and for suggesting the title of this paper.

References

1. Ash, R.B.: Real Analysis and Probability. New York: Academic Press 1972
2. Blackwell, D.: Discounted dynamic programming. Ann. Math. Statist. **36**, 226–235 (1965)
3. Blackwell, D.: On stationary policies. J. Royal Statist. Soc. **133**, 33–37 (1971)
4. Dubins, L.E., Savage, L.J.: How to gamble if you must. New York: McGraw-Hill 1965
5. Furukawa, N.: Markovian decision processes with compact action spaces. Ann. Math. Statist. **43**, 1612–1622 (1972)
6. Hinderer, K.: Foundations of non-stationary dynamic programming with discrete time-parameter. Lecture Notes in Operations Research and Mathematical Systems, vol. 33. Berlin-Heidelberg-New York: Springer 1970
7. Hinderer, K.: Instationäre dynamische Optimierung bei schwachen Voraussetzungen über die Gewinnfunktionen. Abh. math. Sem. Univ. Hamburg **36**, 208–223 (1971)
8. Hordijk, A.: Dynamic Programming and Markov Potential Theory. Amsterdam: Mathematical Centre Tracts 51, 1974
9. Kuratowski, K.: Topology I. New York: Academic Press 1966
10. Kuratowski, K.: Topology II. New York: Academic Press 1968
11. Maitra, A.: Discounted dynamic programming on compact metric spaces. Sankhya **30**, Ser. A, 211–216 (1968)
12. Michael, E.: Topologies on spaces of subsets. Trans. Amer. Math. Soc. **71**, 152–182 (1951)
13. Rieder, U.: Bayesian dynamic programming (To be published)
14. Rieder, U.: On stopped decision processes with discrete time parameter (To be published)
15. Schäl, M.: Ein verallgemeinertes stationäres Entscheidungsmodell der dynamischen Optimierung. Vol. X, 145–162: Methods of operations research, ed. R. Henn. Meisenheim: Anton Hain 1971
16. Schäl, M.: On continuous dynamic programming with discrete time-parameter. Z. Wahrscheinlichkeitstheorie verw. Geb. **21**, 279–288 (1972)
17. Schäl, M.: Dynamische Optimierung unter Stetigkeits- und Kompaktheitsbedingungen. Habilitationsschrift, Univ. Hamburg 1972
18. Schäl, M.: A selection theorem for optimization problems. Arch. Math. XXV, 219–224 (1974)
19. Schäl, M.: On dynamic programming: compactness of the space of policies (To appear in Stochastic Processes Appl. A summary of this paper may be found in [20])
20. Schäl, M.: Dynamic programming under continuity and compactness assumptions. Advances Appl. Probability **5**, 28–29 (1973)
21. Strauch, R.E.: Negative dynamic programming. Ann. Math. Statist. **37**, 871–890 (1966)
22. Sudderth, W.D.: On the Dubins and Savage characterization of optimal strategies. Ann. Math. Statist. **43**, 498–507 (1972)

Manfred Schäl
Institut für Angewandte Mathematik
und Informatik der Universität Bonn
Abtlg. Wahrscheinlichkeitstheorie
und Mathem. Statistik
D-5300 Bonn, Wegelerstraße 6
Federal Republic of Germany

(Received September 10, 1973; in revised form December 17, 1974)