JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 55, 207–238 (1976)

Estimates for Finite-Stage Dynamic Programs*

K. HINDERER†

Institut für Mathematische Stochastik, Universität Hamburg, Rothenbaumchaussee 45, 2 Hamburg 13

Submitted by Karl J. Åström

1. INTRODUCTION

The computational burden in solving (stationary) infinite-stage dynamic programs has been and still is a challenge for the search of approximate solutions, i.e. for estimates of the value function V^∞ and for "good" policies. Apart from the longknown estimate for V^∞, obtainable from the fixed-point theorem for contractions, the work of MacQueen [8] seems to have been the first step. For a recent contribution and review of previous work cf. Schellhaas [12].

There seems to be known very little concerning the approximate solution of *finite-stage* dynamic programs. This is somewhat surprising since there is much evidence that most—if not all—real world problems in dynamic programming should be formulated as finite-stage problems. The value N of the horizon might not be known exactly, but often one will have some realistic estimate of it, or one will know bounds for it. When we look at Howard's well known automobile replacement problem, e.g., where a person buys its first car at the age of x years, we may well assume that he will hold cars for his expected residual life time of e_x years. This implies that we should take for the horizon the number $e_x \cdot 4$, since the length of one period in this example is a quarter of a year. As another example we mention the stock control problem where one usually can estimate the time at which the sale of the stocked item will be discontinued, since, e.g., the item is replaced by a technically imporved one or since the item has become oldfashioned. Examples for which the exact value of the horizon is known, are those production problems which have a fixed time of delivery [4].

Nevertheless, finite-stage dynamic programs are not intensively studied in the literature. Most books on the subject of dynamic programming rest

* Research sponsored by the National Research Council of Canada under Grant A-7223.

† Present address: Institut für Mathematische Statistik, Universität Karlsruhe, Englerstrasse, 7500 Karlsruhe 1.

409/55/1-14

content in that context with the value iteration technique and with some statement, saying that except for small horizon the computational load of that procedure becomes prohibiting. And then the authors usually pass quickly to the infinite-stage problem, which has the following two well-known advantages over the finite-stage problem: (i) One has to compute only one value-function V^∞ instead of a whole sequence $(V^n, 1 \leqslant n \leqslant N)$; (ii) whenever there exists an optimal policy at all, then usually there exists even a stationary optimal policy. As a consequence, there is much emphasis today on infinite-stage models, to such an extent, that usually not the infinite-stage model is considered as approximation of the finite-stage model, but vice versa; cf. however, Beckmann [1, p. 21].

In the present paper we take a radically different point of view: We regard the finite-stage (say N-stage) model with value function V^N as the primary object of interest, and the infinite-stage model as—sometimes very useful, although not always available—approximation. Our approach has the following implications:

(i) Estimates are required for the "goodness" of the approximation of the N-stage model by the infinite-stage model, e.g. upper and lower bounds for $V^N - V^\infty$ are desired. It seems that this problem has hardly been dealt with in the literature explicitly. However, some—though not all—of the known estimates for the infinite-stage model may be used in our sense by just giving them a new interpretation. Section 3 is a short review of such estimates, as far as they are known to us. Nearly all of them are improved and/or generalized in Sections 4 and 5.

(ii) If the discount factor β is not less than one, then the infinite-stage model is in general not defined, though the N-stage model makes real sense: $\beta > 1$ may be interpreted as the case of "inflation". We overcome this difficulty by a *method of extrapolation* from the k-stage model (where k in practice will be "small") to the N-stage model. This is done in Section 4. We exhibit an infinite set of possible estimates, show that all but a finite number may be discarded, construct among classes of computationally simple estimates "best" elements, and consider the convergence of error bounds. The derivation of our estimates depends heavily on a simple, but very useful estimate (cf. Lemma 4.4), which was already used by Porteus [10] for deriving bounds for V^∞.

(iii) We shall not make explicit use of the infinite-stage model, due to the following observation: The minimal requirement for the infinite-stage model to be useful as an approximation is the pointwise convergence of (V^n) to V^∞. This is a nontrivial problem and, in fact, an example given by Strauch [14] shows that (if $\beta \geqslant 1$) both V^∞ and $V := \lim V^n$ may exist without being identical. Now, from the conceptual point of view, the "limit" of the finite-stage model is the important object, and hence we do not care at all about V^∞, but only about V, whenever it exists. The fact, that under most of our assumptions both V^∞ and V exist and coincide, is neither for our problem nor for our method of proof of any relevance.

Finally, we should mention that in Section 5 we give a method for the comparison of V^N not only with V, but with an arbitrary *auxiliary function* w. Taking for w either V (or more generally, a fixed point of the operator U) or—in case $\beta = 1$—a solution of the equation $w + g = Uw$ for some constant g, or finally the function identically zero, yields three different types of estimates.

In Section 6 we show, how the results in Sections 4 and 5 may be used to find lower estimates for "good" policies. We consider two types of policies: The first one is of the "extrapolation" type; it is useable even if $\beta \geqslant 1$; the other type uses in a sense optimal stationary policies of the infinite-stage model.

In the final Section 7 we present some numerical results for Howard's toymaker example. Though being very simple, this example gives not too bad an idea of the goodness of the estimates. We compute several estimates for different values of β, N and the parameter k, and the relative error. The latter can be done, since the simple example admits an explicit solution.

2. The Mathematical Model

A *stationary dynamic program* is a tupel $((S, \mathfrak{S}), (A, \mathfrak{A}), D, q, r, \beta, V^0)$ of the following meaning:

(i) $(S, \mathfrak{S})$ and $(A, \mathfrak{A})$ are Borel spaces[1]; S is called the *state space* and A is called the *action space.*

(ii) D is a measurable subset of $S \times A$ which contains the graph of a measurable map from S into A. D is called the set of admissible state-action pairs; and for any state $s \in S$ the (nonempty and measurable) s-section $D(s)$ of D is called the *set of admissible actions* when the system is in state s.

(iii) q is a transition probability (otherwise called a Markov kernel) from D into S, the so-called *transition law* during a single stage. In the probability model—to be constructed from the data—$q(s, a, B)$ will be the (conditional) probability that the system moves during the nth stage from s to some state in B under the influence of action a, independent of previous states and actions.

(iv) r is a bounded measurable map from D into the set $\mathbb{R}$ of real

[1] In the sense of Blackwell [2]; in Hinderer [3] they are called standard Borel-spaces.

numbers, called the reward function, and $\beta \in \mathbb{R}^+ := (0, \infty)$ is the *discount factor*. This means that one gets the reward $\beta^n r(s, a)$, if the system is in state s at the nth stage, and if then action a is taken. The case $\beta > 1$ may be interpreted as inflation.

(v) V^0 is a bounded measurable map from S to $\mathbb{R}$ and is called the *terminal reward function*. This means that we get the reward $\beta^N V^0(s)$ if N is the horizon and if s is then the final state. (When using the finite-stage model as an approximation to the infinite-stage model, one has a free choice for V^0, and often $V^0 \equiv 0$ is selected. In finite-stage models it is essential for applications to admit $V^0 \not\equiv 0$.)

Let F denote the set of *decision functions*, i.e. of measurable maps $f: S \to A$, whose graphs belong to D. Denote by Δ_N the set of (deterministic Markovian) N-stage *policies*, i.e. of sequences of N decision functions.

For any N-stage policy π there is defined $V_\pi^N(s)$, the (conditional) *expected discounted total reward* (including the terminal reward) when policy π is used and when the system starts in state s. (For a discussion of the underlying probability model cf. eg. Blackwell [2] or Hinderer [3].) As usual, a policy $\pi \in \Delta_N$ is called *optimal* iff

$$V_\pi^N(s) = V^N(s) := \sup_{\sigma \in \Delta_N} V_\sigma^N(s), \qquad s \in S. \tag{2.1}$$

The function V^N will be called the *value function* of the N-stage model.

Denote by $\mathfrak{S}_m$ the set of bounded and universally measurable functions from the state space into the reals. Formulations and proofs simplify considerably by the use of the isotone operator $U: \mathfrak{S}_m \to \mathbb{R}^S$, defined by

$$(Uv)(s) := \sup_{a \in D(s)} \left[r(s, a) + \beta \int q(s, a, dt)\, v(t) \right], \qquad s \in S. \tag{2.2}$$

It is easy to see that Uv is bounded, but example (48) in Blackwell, Freedman and Orkin [2a] shows, that Uv need not be universally measurable. Therefore we use the following device: Let $\mathfrak{S}_U$ be the set of those functions $v \in \mathfrak{S}_m$, for which $Uv \in \mathfrak{S}_m$, $U^2 v \in \mathfrak{S}_m$,.... The set $\mathfrak{S}_U$ has the following properties:

(i) $\mathfrak{S}_U$ is the largest among those subsets of $\mathfrak{S}_m$ which are mapped by U into itself.

(ii) $\mathfrak{S}_U$ is under the metric $d(v, w) := \| v - w \|$ a closed subset of the Banach space $\mathfrak{S}_m$, hence a complete metric space.

(iii) $\mathfrak{S}_U$ contains the set of all bounded and $\mathfrak{S}$-measurable functions from S into the reals. This can be proved, using results of Strauch [14], as follows: At first Theorem 13.2 in Hinderer [3], applied for $r_n := \beta^{n-1} r$ for

$n < m$, $r_m := \beta^{m-1}V^0$ and $r_n = 0$ for $n > m$, shows that V^m for all $m \in \mathbb{N}$ is universally measurable. Moreover, Theorem 14.4 in Hinderer [3] proves the value iteration procedure

$$V^{n+1} = UV^n, \qquad n \in \mathbb{N}_0. \tag{2.3}$$

Combining both results yields assertion (iii).

Of course, equation (2.3) is fundamental for all our investigations, and we shall use it tacitly whenever it is necessary.

3. A Short Review of Known Results

As far as we know, the estimates stated in this section were originally devised as estimates for V^∞ (in cases where $V^\infty = V := \lim_n V^n$); but we can reinterpret them as estimates for V^N, assuming that V is available to us. (If S and A are finite, and $\beta < 1$, V may be computed e.g. by policy iteration.) Note that some of the known estimates for V, e.g. the bounds

$$(1-\beta)^{-1}\inf(V^k - V^{k-1}) \leqslant V - V^{k-1} \leqslant (1-\beta)^{-1}\sup(V^k - V^{k-1}), \qquad k \in \mathbb{N}, \tag{3.1}$$

derived by MacQueen [8] (for the case $\beta < 1$), cannot be used for our purpose.

(i) Since $\mathfrak{S}_U$ is a complete metric space under the sup-metric (cf. Section 2), it follows from the fixed point theorem for contractions, that in case $\beta < 1$ the sequence (V^n) converges in the sup-norm to some function $V \in \mathfrak{S}_U$ and that the following well known estimate holds:

$$\| V^N - V \| \leqslant (1-\beta)^{-1} \cdot \beta^{N-k} \| V^{k+1} - V^k \|, \qquad 0 \leqslant k < N-1. \tag{3.2}$$

This estimate will be improved and generalized by (5.3).

(ii) From the definition of V_π^N one easily infers that

$$\| V_\pi^N \| \leqslant \| r \| \sum_0^{N-1} \beta^\nu + \| V^0 \| \beta^N. \tag{3.3}$$

Since $|\sup v| \leqslant \sup |v|$ for any real-valued function v, we easily get the estimate

$$\| V^N \| \leqslant \| r \| \sum_0^{N-1} \beta^\nu + \| V^0 \| \beta^N, \tag{3.4}$$

derived by Martin [9, Lemma 3.3.1], for finite state space and action space and $\beta < 1$, but within the more general setting of Bayesian dynamic programs. Formula (3.4), which is improved and generalized by our estimate (4.6), may also be derived from the value iteration (2.3) and the simple inequality

$$\| Uv \| \leqslant \| r \| + \beta \| v \|, \qquad v \in \mathfrak{S}_U . \tag{3.5}$$

(iii) In Martin [9, Theorem 3.4.3] it is shown (under the conditions stated in (ii)), that

$$\| V - V^N \| \leqslant \beta^N \max\{\sup r/(1-\beta) - \inf V^0, \sup V^0 - \inf r/(1-\beta)\}. \tag{3.6}$$

If V^0 is identically zero, then the bounds in (3.2), $k = 0$, and (3.6), multiplied by $(1-\beta)\beta^{-N}$, reduce to $\sup_{s\in S} | \sup_{a\in D(s)} r(s, a)| \leqslant \| r \|$ and $\| r \|$, respectively; hence, if $V^0 \equiv 0$, then (3.2), $k = 0$, is at least as good as (3.6), but in general the two estimates are not comparable.

(iv) In Martin [9, Corollary 3.4.4], the following estimates are given (under the conditions stated in (ii)):
If $\sup V^0 - \beta \inf V^0 \leqslant \inf r$, then

$$0 \leqslant V - V^N \leqslant \beta^N((1-\beta)^{-1} \sup r - \inf V^0); \tag{3.7}$$

and if $\inf V^0 - \beta \sup V^0 \geqslant \sup r$, then

$$\beta^N((1-\beta)^{-1} \inf r - \sup V^0) \leqslant V - V^N \leqslant 0. \tag{3.8}$$

These estimates are contained in (5.16) and Corollary 5.1.

(v) In Rieder [11, Satz 6.3], an estimate is proved for general Bayesian models with countable action space and $\beta < 1$. When specialized to our type of dynamic programs it reads as follows: Let $\mathfrak{M}$ be the set of $\mathfrak{S}$-measurable and bounded functions from S into $\mathbb{R}$. Define the operator T: $\mathfrak{M} \to \mathbb{R}^S$ by

$$(Tv)(s) := \sup_{a\in D(s)} \int q(s, a, dt)\, v(t), \qquad s \in S. \tag{3.9}$$

It can be shown that, since A is countable, T as well as U map $\mathfrak{M}$ into itself. Put

$$r'(s) := \sup_{a\in D(s)} r(s, a), \qquad s \in S. \tag{3.10}$$

Then

$$u_n := \max\{U^n(\sup r'/(1-\beta)) - V^0, V^0 - U^n(\inf r'/(1-\beta))\} \tag{3.11}$$

belongs to $\mathfrak{M}$ and

$$| V - V^N | \leqslant \beta^N \inf_{n \in \mathbb{N}} T^N u_n . \tag{3.12}$$

(As shown by Rieder, the sequence $(T^N u_n, n \in \mathbb{N})$ is nonincreasing.)

As noted by Rieder, estimate (3.12) is an improvement of Martin's estimate (3.6); this will also become evident in connection with our Theorem 4.2.

(vi) Porteus [10] derived bounds for the infinite-stage model within the framework of the monotone contraction models of Denardo. From Lemma 5 in Porteus [10] one can derive the following important estimate:

$$\inf(V^1 - V^0) \sum_{\nu=1}^{N-1} \beta^\nu \leqslant V^N - V^1 \leqslant \sup(V^1 - V^0) \sum_{\nu=1}^{N-1} \beta^\nu. \tag{3.13}$$

We will obtain a generalization of (3.13) in (4.13) and (4.24).

4. Estimates for V^N Which are Obtained by Extrapolation

A. In this section we fix, unless something else is stated, the horizon $N \in \mathbb{N}$ and the number $k \in \mathbb{N}_0$, $k < N$, of stages, for which the value functions are already known to us. We are looking for estimates for V^N in terms of $V^0, V^1, \ldots, V^k$, and we do not assume $\beta < 1$.

Our estimates will be of the form

$$v_1 \leqslant V^N \leqslant v_2 , \tag{4.1}$$

for some functions v_1, v_2. We shall say, that the estimate $v_1' \leqslant V^N \leqslant v_2'$ is an *improvement* of (4.1) iff $v_1 \leqslant v_1' \leqslant v_2' \leqslant v_2$. In our numerical examples we shall also evaluate the (maximal absolute) relative errors $\rho(v_i)$, $i = 1, 2$, defined by

$$\rho(v_i) := \|(V^N - v_i)/V^N\| . \tag{4.2}$$

Since the computation of $\rho(v_i)$ needs the exact value of V^N, which usually will not be available, the easily derived upper bound (4.3) may be useful.

Lemma 4.1. *If none on the intervals $\langle v_1(s), v_2(s) \rangle$ contains the origin, then*

$$\rho(v_i) \leqslant \|(v_2 - v_1)/\min(| v_1 |, | v_2 |)\| , \qquad i = 1, 2. \tag{4.3}$$

The right-hand side of (4.3) will be called the *error bound* of estimate (4.1). Moreover, if v_1 and v_2 depend on N, and if the error bound of (4.1) converges to b for $N \to \infty$, then we call b the *asymptotic error bound* of (4.1). When using an estimate for "very large" horizon N, it may be useful to compute first the asymptotic error bound (if it exists), in order to obtain some idea about the accuracy one can expect.

B. In the sequel we shall use the *abbreviations*

$$\beta_{m,n} := \sum_{\nu=m}^{n} \beta^{\nu}, \qquad \beta_n := \beta_{0,n}, \qquad \beta \in (0, \infty), \tag{4.4}$$

and

$$\operatorname{sp} v := \sup v - \inf v, \qquad v \in \mathbb{R}^S \text{ bounded}. \tag{4.5}$$

We call $\operatorname{sp} v$ then *span* of the function v.

Our first and easily obtainable estimate is contained in the following theorem. (The function r' is defined in (3.10).)

THEOREM 4.2. (i) *The estimate*

$$\beta_{N-k-1} \inf r' + \beta^{N-k} \inf V^k \leqslant V^N \leqslant \beta_{N-k-1} \sup r' + \beta^{N-k} \sup V^k \tag{4.6}$$

holds and is improving with increasing k *and fixed* $N > k$.

(ii) *The estimate* (4.6) *has the asymptotic error bound*[2]

$$b_k := \frac{\operatorname{sp} r' + (\beta - 1)^+ \operatorname{sp} V^k}{\min(|\sup r' + (\beta - 1)^+ \sup V^k|, |\inf r' + (\beta - 1)^+ \inf V^k|)}, \tag{4.7}$$

unless $r' \equiv 0$, *in which case the error bound of* (4.6) *is equal to*

$$\operatorname{sp} V^k / [\min(|\sup V^k|, |\inf V^k|)] \qquad \text{for all } N \text{ and } k < N. \tag{4.8}$$

Proof. (i) It is easy to see that

$$\inf r' + \beta \inf v \leqslant Uv \leqslant \sup r' + \beta \sup v, \qquad v \in \mathfrak{S}_U .$$

Now the assertion follows from (2.3) by induction on k.

(ii) Formula (4.7) is obtained by a more or less straightforward computation, which we omit; in the proof of Theorem 4.5 a similar, but more complicated formula will be proved in detail. ∎

Remark 1. Using the fact that $c \leqslant d \leqslant e$ implies $|d| \leqslant \max(|c|, |e|)$, one easily shows, that (4.6) is an improvement of Martin's estimate (3.4). On the other hand, we shall improve on (4.6) in (5.18).

[2] We use the convention $c/0 := \infty$ for $c > 0$, and $0/0 := 0$.

Remark 2. From (4.6) we get for $k = 0$, using $V^0 := (1 - \beta)^{-1} \sup r'$ and $V^0 := (1 - \beta)^{-1} \inf r'$, respectively, that for all $n \in \mathbb{N}$

$$\inf r'/(1 - \beta) \leqslant U^n(\inf r'/(1 - \beta)) \leqslant U^n(\sup r'/(1 - \beta)) \leqslant \sup r'/(1 - \beta).$$

As a consequence we obtain from Rieder's estimate (3.12) the following easily computable improvement of Martin's estimate (3.6): If $\beta < 1$, then

$$\| V^N - V \| \leqslant \beta^N \max\{\sup r'/(1 - \beta) - \inf V^0, \sup V^0 - \inf r'/(1 - \beta)\}. \tag{4.9}$$

An improvement of (4.9) is contained in (5.10).

Sometimes one can assure that (V^n) is nondecreasing or nonincreasing, which yields the useful bounds

$$V^N \geqslant V^k \quad \text{or} \quad V^N \leqslant V^k, \quad 0 \leqslant k < N, \tag{4.10}$$

respectively. Since the operator U is isotone, a sufficient condition for (4.10) to hold is $V^1 \geqslant V^0$ or $V^1 \leqslant V^0$, respectively. Other sufficient conditions which are stronger, but which do not require the computation of V^1, are contained as the special case $c = 0$, $m = 1$ in the following corollary to Theorem 4.2.

COROLLARY 4.3. *For $c \in \mathbb{N}_0$ and $m \in \mathbb{N}$ holds:*

(i) *If* $\sup V^c \leqslant \beta_{m-1} \inf r' + \beta^m \inf V^c$, *then every sequence* $(V^{c+\mu+nm}, n \in \mathbb{N}_0)$ $0 \leqslant \mu < m$, *is nondecreasing.*

(ii) *If* $\inf V^c \geqslant \beta_{m-1} \sup r' + \beta^m \sup V^c$, *then every sequence* $(V^{c+\mu+nm}, n \in \mathbb{N}_0)$, $0 \leqslant \mu < m$, *is nonincreasing.*

Proof. The assumption and Theorem 4.2, part (i), applied for $N := c + m$, imply $V^{c+m} \geqslant V^c$, hence

$$V^{c+\mu+nm+m} = U^{\mu+nm} V^{c+m} \leqslant U^{\mu+nm} V^c = V^{c+\mu+nm},$$

and the first assertion follows. The second assertion is proved in the same way. ∎

C. An important role will play the following lemma, which has been used more or less explicitly (and under other assumptions for the underlying dynamic program) by several authors for the derivation of estimates for infinite-stage models; cf. Beckmann [1, p. 53], Porteus [10, Lemma 4], Hübner [7].

LEMMA 4.4. *If v and w belong to $\mathfrak{S}_U$, then for $n \in \mathbb{N}_0$ holds*

$$\beta^n \inf(v - w) \leqslant U^n v - U^n w \leqslant \beta^n \sup(v - w), \tag{4.11}$$

and

$$\operatorname{sp}(U^n v - U^n w) \leqslant \beta^n \operatorname{sp}(v - w). \tag{4.12}$$

Here is a simple proof: From $v - w \leqslant \sup(v - w)$ we get $v \leqslant w + \sup(v - w)$. Then we deduce from

$$U(v + c) = Uv + \beta c, \qquad v \in \mathfrak{S}_U, \quad c \in \mathbb{R},$$

and the monotonicity of U, that

$$Uv \leqslant U(w + \sup(v - w)) = Uw + \beta \sup(v - w).$$

Now the second inequality in (4.11) follows by induction on n, and the first one is obtained from the second one by interchanging v and w. Finally, (4.12) is an immediate consequence of (4.11). ∎

We start our investigation with the following generalization of Porteus' estimate (3.13).

THEOREM 4.5. (i) *For $1 \leqslant k < N$ the estimate*

$$\beta_{1,N-k} \inf(V^k - V^{k-1}) \leqslant V^N - V^k \leqslant \beta_{1,N-k} \sup(V^k - V^{k-1}) \tag{4.13}$$

holds and is improving with increasing k and fixed $N > k$.

(ii) *If $\beta < 1$, then estimate* (4.13) *has the asymptotic error bound*

$$b_k := \frac{\beta \operatorname{sp}(V^k - V^{k-1})}{\min(\inf_s |\beta x + (1 - \beta) V^k(s)|, \inf_s |\beta y + (1 - \beta) V^k(s)|)}, \tag{4.14}$$

where $x := \inf(V^k - V^{k-1})$ and $y := \sup(V^k - V^{k-1})$, unless

$$\inf_s |\beta x + (1 - \beta) V^k(s)| = \inf_s |\beta y + (1 - \beta) V^k(s)| = 0.$$

(iii) *If $\beta \geqslant 1$, then the estimate* (4.13) *has the asymptotic error bound*

$$b_k := \operatorname{sp}(V^k - V^{k-1})/\min(|x|, |y|). \tag{4.15}$$

Proof. (i) Using (4.11) we get

$$V^N - V^k = \sum_{\nu=1}^{N-k} (V^{k+\nu} - V^{k+\nu-1})$$

$$= \sum_{1}^{N-k} (U^\nu V^k - U^\nu V^{k-1}) \leqslant \sum_{1}^{N-k} \beta^\nu \sup(V^k - V^{k-1}),$$

and the first inequality of (4.13) is proved in the same way. That (4.13) improves with increasing k follows, since e.g.

$$\begin{aligned} &V^{k+1} + \beta_{1,N-k-1} \sup(V^{k+1} - V^k) \\ &\quad \leqslant V^k + \sup(V^{k+1} - V^k) + \beta_{1,N-k-1} \sup(V^{k+1} - V^k) \\ &\quad = V^k + \beta_{N-k-1} \sup(V^{k+1} - V^k) \leqslant V^k + \beta_{1,N-k} \sup(V^k - V^{k-1}). \end{aligned}$$

(ii) and (iii). It follows from the definition, that the error bound $b(N, k)$ of (4.13) can be written as

$$b(N, k) = \frac{\beta \operatorname{sp}(V^k - V^{k-1})}{\inf_s \min(| \beta x + \delta_N V^k(s)| , | \beta y + \delta_N V^k(s)|)}, \tag{4.16}$$

where $\delta_N := (\beta_{N-k})^{-1} \to (1 - \beta)^+$ for $N \to \infty$. As is well known, we are allowed to interchange inf and min in (4.16). Moreover, if $N \to \infty$, then $| g_N(s)| := | \beta x + \delta_N V^k(s)|$ converges uniformly in s towards $| g(s)| := | \beta x + (1 - \beta)^+ V^k(s)|$, since

$$| \, | g_N(s)| - | g(s)| \, | \leqslant | g_N(s) - g(s)| \leqslant | \delta_N - (1 - \beta)^+ | \cdot \| V^k \| .$$

Now it is well known and easy to prove that uniform convergence implies the convergence of the sequence of infima, i.e. $\inf_s | \beta x + \delta_N V^k(s)|$ converges towards $\inf_s | \beta x + (1 - \beta)^+ V^k(s)|$ for $N \to \infty$. Therefore $b(N, k)$ tends to b_k, provided that the denominator in (4.14) and (4.15) does not vanish. So let us assume that $\beta < 1$ and that exactly one of the two terms in the denominator of (4.14), say $\inf_s | \beta x + (1 - \beta) V^k(s)|$ vanishes. Then there exists a sequence of points $s_n \in S$ such that $\beta x + (1 - \beta) V^k(s_n) \to 0$ for $n \to \infty$. It follows that

$$\begin{aligned} \beta | x - y | &= \lim_n | \beta(x - y) - \beta x - (1 - \beta) V^k(s_n)| \\ &= \lim_n | \beta y + (1 - \beta) V^k(s_n)| \geqslant \inf_s | \beta y + (1 - \beta) V^k(s)| > 0. \end{aligned}$$

But then $b(N, k)$ tends to infinity for $N \to \infty$, and this case is contained in (4.14). Next let us assume that $\beta \geqslant 1$, and that exactly one of the terms x and y vanishes. Then $b(N, k)$ tends to infinity for $N \to \infty$, and this case is contained in (4.15). Finally, if $\beta \geqslant 1$ and if $x = y = 0$, then $b(N, k)$ equals zero for all N and hence $b_k = 0$. But this case is again contained in (4.15) according to our convention $0/0 := 0$. ∎

As an immediate consequence of (4.13) one gets the estimate

$$\| V^N - V^k \| \leqslant \beta_{1,N-k} \| V^k - V^{k-1} \| , \qquad N > k \geqslant 1 \tag{4.17}$$

which also follows from

$$\| Uv - Uw \| \leqslant \beta \| v - w \|, \qquad v, w \in \mathfrak{S}_U . \tag{4.18}$$

D. In this subsection we generalize the method that led to estimate (4.13). By this way we get an infinite family of estimates, from which however, all but a finite number may be discarded. Then we find best estimates in some computational attractive classes of estimates. In particular we get a generalization of estimate (4.13).

Our general method for constructing estimates consists of two steps. (i) We select $\nu \in \mathbb{N}_0$ and split $V^N - V^\nu$ as

$$V^N - V^\nu = \sum_{\mu=0}^{m-1} (V^{i_{\mu+1}} - V^{i_\mu}),$$

where $m \in \mathbb{N}$ and $i_\mu \in \mathbb{N}_0$ are arbitrary, and $i_0 = \nu$ and $i_m = N$. Then we get immediately

$$\sum_0^{m-1} \inf(V^{i_{\mu+1}} - V^{i_\mu}) \leqslant V^N - V^\nu \leqslant \sum_0^{m-1} \sup(V^{i_{\mu+1}} - V^{i_\mu}). \tag{4.19}$$

(ii) We "reduce" the terms in the sums of (4.19) by means of (4.11), in order to arrive at an estimate, in which at most $V^0, V^1,..., V^k$ occur. Such an estimate will be called *admissible*. For later purposes we admit at first arbitrary reductions, not only those that lead to admissible estimates. Therefore, we have

LEMMA 4.6. *If $m \in \mathbb{N}$ is arbitrary and if $(i_\mu) \in \mathbb{N}_0^{m+1}$ and $(\rho_\mu) \in \mathbb{N}_0{}^m$ satisfy*

$$\rho_\mu \leqslant \min(i_\mu , i_{\mu+1}), \qquad 0 \leqslant \mu < m, \tag{4.20}$$

and $i_m = N$, then

$$\sum_0^{m-1} \beta^{\rho_\mu} \inf W_\mu \leqslant V^N - V^\nu \leqslant \sum_0^{m-1} \beta^{\rho_\mu} \sup W_\mu , \tag{4.21}$$

holds, where $\nu := i_0$, and $W_\mu := V^{i_{\mu+1}-\rho_\mu} - V^{i_\mu-\rho_\mu}$.

The set E_1 of those estimates (4.21) that are admissible (with respect to fixed N and k) is determined by the conditions $\nu \leqslant k$, $i_{\mu+1} - \rho_\mu \leqslant k$ and $i_\mu - \rho_\mu \leqslant k$. The last two conditions may be written as

$$\rho_\mu \geqslant \max(i_\mu , i_{\mu+1}) - k, \qquad 0 \leqslant \mu < m. \tag{4.22}$$

Denote by E_0 the set of all estimates of the form (4.1). If e and e' are two estimates in E_0 and if e' is an improvement of e (for the definition cf. the beginning of this section), then we write $e' \leqslant e$. The relation "$\leqslant$" is a partial ordering on E_0. If $E' \subset E''$ are nonempty subsets of E_0, then E' is said to be *complete* in E'', if for any $e \in E'' - E'$ there exists $e' \in E'$ such that $e' \leqslant e$; then obviously all estimates in $E'' - E'$ can be discarded. If $E' = \{e'\} \subset E''$ is complete, then e' is said to be a *best estimate* within E''.

In the next theorem we show that E_1 contains a finite complete subset E_2. We do not know whether all estimates in E_2 are best estimates within E_1.

THEOREM 4.7. *Let E_2 be the finite set of those estimates in E_1, for which*

(a$_1$) $\rho_\mu = (\max(i_\mu, i_{\mu+1}) - k)^+$, $0 \leqslant \mu < m$;

(a$_2$) $i_\mu \neq i_0$ *for* $1 \leqslant \mu \leqslant m$;

(a$_3$) *if* $m \geqslant 2$, *then* $i_2 > k$;

(a$_4$) (i_μ) *is strictly increasing for* $1 \leqslant \mu \leqslant m$.

Then E_2 is complete in E_1.

Proof. Define property (a$_2'$) as: $i_\mu \neq i_\nu$ for $\mu \neq \nu$, and property (a$_3'$) as: if $m \geqslant 2$, then $i_\mu > k$ for all $\mu \geqslant 2$. Denote by F_1, F_2, and F_3 the set of those estimates in E_1, that satisfy (a$_1$), (a$_1$) $\wedge$ (a$_2'$) and (a$_1$) $\wedge$ (a$_2'$) $\wedge$ (a$_3'$), respectively. Put $F_0 := E_1$, $F_4 := E_2$. Since (a$_1$) $\wedge$ (a$_2'$) $\wedge$ (a$_3'$) $\wedge$ (a$_4$) is equivalent to (a$_1$) $\wedge$ (a$_2$) $\wedge$ (a$_3$) $\wedge$ (a$_4$) and since completeness is a transitive relation on the system of nonempty subsets of E_0, it is sufficient to show that F_{j+1} is complete in F_j for $0 \leqslant j \leqslant 3$. For this purpose we shall only regard the upper bounds and use the abbreviation $M(i, j) := \sup(V^i - V^j)$. The considerations for the lower bounds are very similar. (a) When reducing $M(i_{\mu+1}, i_\mu)$ as described above, the best we can do is to reduce not at all, if $\max(i_\mu, i_{\mu+1}) \leqslant k$, and otherwise to reduce in such a way, that the largest of the terms i_μ, $i_{\mu+1}$ is reduced to k. Therefore, F_1 is complete in F_0. (b) Any estimate e in F_1 is determined by $m \in \mathbb{N}$ and $x := (i_\mu) \in \mathbb{N}_0^{m+1}$, such that $i_0 \leqslant k$ and $i_m = N$; we shall write then $e = e(x)$. Now we select $e(x) \in F_1$, for which $i_\kappa = i_\lambda$ for some κ, λ, $0 \leqslant \kappa < \lambda \leqslant m$. Then, using the abbreviation

$$l(i_{\mu+1}, i_\mu) := \beta^{\rho_\mu} M(i_{\mu+1} - \rho_\mu, i_\mu - \rho_\mu),$$

we get

$$l(i_\lambda, i_\kappa) = 0 = M(i_\lambda, i_\kappa) \leqslant \sum_{t=0}^{\lambda-\kappa-1} M(i_{\lambda-t}, i_{\lambda-t-1}) \leqslant \sum_{t=0}^{\lambda-\kappa-1} l(i_{\lambda-t}, i_{\lambda-t-1}).$$

Therefore, if $y := (i_0, i_1, \ldots, i_\kappa, i_{\lambda+1}, \ldots, i_m)$, then $e(y) \in F_1$, and $e(y) \leqslant e(x)$. Iterating the transition from $e(x)$ to $e(y)$, if necessary, we can find $e(z) \in F_2$ such that $e(z) \leqslant e(x)$. Therefore, F_2 is complete in F_1. (c) Choose $e(x) \in F_2$

for which $m \geqslant 2$ and $i_\mu \leqslant k$ for some $\mu \geqslant 2$. Then we get for λ, the largest of the indices for which $i_\lambda \leqslant k$

$$l(i_\kappa, i_0) = M(i_\lambda, i_0) \leqslant \sum_0^{\lambda-1} M(i_{\lambda-\mu}, i_{\lambda-\mu-1}) \leqslant \sum_0^{\lambda-1} l(i_{\lambda-\mu}, i_{\lambda-\mu-1}).$$

Therefore, if $y := (i_0, i_\lambda, i_{\lambda+1}, \ldots, i_m)$, then $e(y) \in F_3$ and $e(y) \leqslant e(x)$. Therefore F_3 is complete in F_2. (d) Choose $e(x) \in F_3$ for which $x =: (i_\mu)$ is not strictly increasing for $1 \leqslant \mu \leqslant m$. (Note that condition ($a_2$) does not exclude the possibility that $i_0 > i_1$.) Then there exists λ, $0 < \lambda < m$, such that $i_\lambda > \max(i_{\lambda-1}, i_{\lambda+1})$. Put $\rho(a, b) := (\max(a, b) - k)^+$, $i := i_{\lambda-1}$, $g := i_\lambda$, $j := i_{\lambda+1}$. It follows that $\rho(i, g) = \rho(g, j) = (g - k)^+ =: \rho \geqslant \rho(i, j)$, and that $|i - j| \leqslant k$. Therefore

$$\begin{aligned} &l(j, g) + l(g, i) \\ &\quad = \beta^\rho[M(j - \rho, g - \rho) + M(g - \rho, i - \rho)] \geqslant \beta^\rho M(j - \rho, i - \rho) \\ &\quad \geqslant \beta^{\rho(i,j)} M(j - \rho(i, j), i - \rho(i, j)) = l(j, i). \end{aligned}$$

It follows that, if $y := (i_0, i_1, \ldots, i_{\lambda-1}, i_{\lambda+1}, \ldots, i_m)$, then $e(y)$ belongs to F_3 and $e(y) \leqslant e(x)$. Iterating the transition from $e(x)$ to $e(y)$, if necessary, we can find $e(z) \in F_4$ such that $e(z) \leqslant e(x)$. Therefore F_4 is complete in F_3. ∎

E. In real world problems, it will in general be impossible to scan all estimates in E_2. Therefore, one will look for a set $E_3 \subset E_1$ of easily computable estimates, and then one will again search for best estimates within E_3 or at least for a small complete subset.

Particularly attractive from the computational point of view are those estimates in E_1 for which both $i_{\mu+1} - \rho_\mu$ and $i_\mu - \rho_\mu$ are independent of μ, $0 \leqslant \mu < m$, since then (4.21) simplifies considerably. Now this independence of μ holds iff $i_{\mu+1} - i_\mu = c$ and $i_{\mu+1} - \rho_\mu = d$ for some integer c and some $d \in \mathbb{N}_0$, and then $i_\mu = \nu + \mu c$ for $0 \leqslant \mu \leqslant m$. If the estimate belongs to E_1, then $\nu < N$. For later purposes we admit also estimates in $E_0 - E_1$, but for convenience and without real loss of generality we assume $\nu < N$, hence $c > 0$. Now let us fix an arbitrary tupel $(\nu, c, d) \in \mathbb{N}_0 \times \mathbb{N}^2$ and define (i_μ) and (ρ_μ) by $i_\mu := \nu + \mu c$, $0 \leqslant \mu \leqslant m$; $\rho_\mu := i_{\mu+1} - d$. Then $\rho_\mu \geqslant 0$ for all μ iff $d \leqslant \nu + c$, and condition (4.20) is satisfied iff $c \leqslant d$; and then $d \geqslant 1$. Therefore we have

LEMMA 4.8. *If* $(\nu, c, d) \in \mathbb{N}_0 \times \mathbb{N}^2$ *satisfies*

(i) $\nu < N$,

(ii) $m := (N - \nu)/c \in \mathbb{N}$,

(iii) $c \leqslant d \leqslant \nu + c$, then

$$\beta^{\nu+c-d}(\beta^c)_{m-1}\inf(V^d - V^{d-c}) \leqslant V^N - V^\nu \leqslant \beta^{\nu+c-d}(\beta^c)_{m-1}\sup(V^d - V^{d-c}). \tag{4.23}$$

It is obvious that estimate (4.23) is admissible iff $\max(\nu, d) \leqslant k$. Now we fix some c, $1 \leqslant c \leqslant k$. Then the set $E_3(c)$ of those estimates (4.23) which are admissible, are determined by those pairs $(\nu, d) \in \mathbb{N}_0 \times \mathbb{N}$ for which $\max(\nu, d) \leqslant k$ and conditions (ii) and (iii) of Lemma 4.8 are satisfied.

THEOREM 4.9. *Fix $c \in \mathbb{N}$, $1 \leqslant c \leqslant k$. Let t be the positive*[3] *remainder of $N - k$ under division by c, and put $n := (N - k - t)/c$. Then*

$$\beta^t(\beta^c)_n \inf(V^k - V^{k-c}) \leqslant V^N - V^{k+t-c} \leqslant \beta^t(\beta^c)_n \sup(V^k - V^{k-c}) \tag{4.24}$$

belongs to E_2 and is a best estimate within $E_3(c)$. In particular, estimate (4.13) *belongs to E_2 and is a best estimate within $E_3(1)$. Moreover, estimate* (4.24) *is improving for fixed c and increasing $k < N$.*

Proof. (a) At first we characterize the estimates in $E_2 \cap E_3(c)$. It follows from the definition of E_2 in Theorem 4.7, that an estimate $e(\nu, d) \in E_3(c)$ belongs to E_2 iff

$$\rho_\mu = \nu + (\mu + 1)\, c - d = (\nu + (\mu + 1)\, c - k) \tag{4.25}$$

for $0 \leqslant \mu < m$ and $\nu + 2c > k$ if $m := (N - \nu)/c \geqslant 2$. Considering (4.25) for $\mu = m - 1$ and $\mu = 0$, respectively, we realize that for any estimate in $E_3(c)$ the relation (4.25) is true for $0 \leqslant \mu < m$ iff $d = k$ and $\nu + c > k$, in which case also the condition $\nu + 2c > k$ is fulfilled, irrespective of the value of m. Hence $e(\nu, d) \in E_3(c)$ belongs to E_2 iff $d = k \leqslant \nu + c$. Therefore, the estimates in $E_2 \cap E_3(c)$ are characterized by those numbers $\nu \in \mathbb{N}_0$, for which $\nu \leqslant k \leqslant \nu + c$ and $(N - \nu)/c \in \mathbb{N}$. It follows, if $m_0 := \min\{\mu \in \mathbb{N} : N - \mu c \leqslant k\}$, that

$$E_2 \cap E_3(c) = \begin{cases} \{e(N - m_0 c, k)\}, & \text{if} \quad (N - k)/c \notin \mathbb{N}, \\ \{e(N - m_0 c, k), e(N - m_0 c - c, k)\}, & \text{otherwise.} \end{cases}$$

In particular, (4.24) is the estimate $e(N - m_0 c, k)$, and thus belongs always to E_2. (b) Now we regard the set of those pairs $(\nu, d) \in \mathbb{N}_0 \times \mathbb{N}$, which determine the estimates $e(\nu, d)$ in $E_3(c)$; cf. Fig. 1, where $k = 10$, $c = 3$, and

[3] t is defined as the unique integer such that $1 \leqslant t \leqslant c$ and $(N - k - t)/c \in \mathbb{N}_0$.

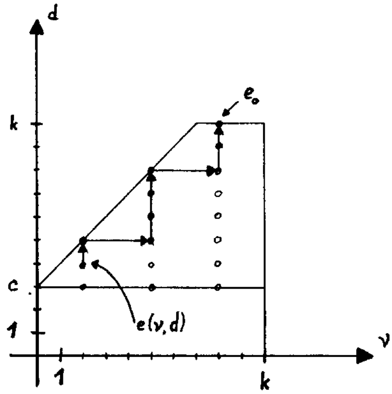

FIGURE 1

$N \equiv 2 \pmod 3$ is assumed. In order to prove that $e_0 := e(N - m_0 c, k)$ is a best estimate within $E_3(c)$, we proceed in two steps: (b_1) For any $e(\nu, d) \in E_3(c)$ such that $d < \min(k, \nu + c)$, $e(\nu, d + 1)$ belongs again to $E_3(c)$, and (4.11) implies $e(\nu, d + 1) \leqslant e(\nu, d)$. ($b_2$) Now we select $e(\nu, d) \in E_3(c)$ such that $d = \min(k, \nu + c)$ and $e(\nu, d) \neq e_0$. Then $\nu \leqslant k - c$, hence $d = \nu + c$, and $e(\nu + c, \nu + c)$ belongs to $E_3(c)$. Now we are going to show that

$$e(\nu + c, \nu + c) \leqslant e(\nu, \nu + c). \tag{4.26}$$

In fact, from (4.23) we see that the upper bound of $e(\nu + c, \nu + c)$ is $V^{\nu+c} + \beta^c(\beta^c)_{m-1} \sup(V^{\nu+c} - V^\nu)$, where $m := (N - \nu - c)/c$. Now we have

$$\begin{aligned} &V^{\nu+c} + \beta^c(\beta^c)_{m-1} \sup(V^{\nu-c} - V^\nu) \\ &\leqslant V^\nu + \sup(V^{\nu+c} - V^\nu) + \beta^c(\beta^c)_{m-1} \sup(V^{\nu+c} - V^\nu) \\ &= V^\nu + (\beta^c)_m \sup(V^{\nu+c} - V^\nu), \end{aligned}$$

and the last term is the upper bound of $e(\nu, \nu + c)$. Exactly the same reasoning holds for the lower bounds, and therefore (4.26) is true. Now, looking at Fig. 1, we realize that (b_1) and (4.26) imply that e_0 is a best estimate within $E_3(c)$. (c) Finally, if c is fixed, then the definition of the set $E_3(c) = E_3(c, k)$ shows that $E_3(c, k') \subset E_3(c, k)$ for $k' < k$. Now the optimality property of estimate (4.24) implies that it is improving with increasing k. ∎

In general, estimate (4.24) has no asymptotic error bound, since the number t, as a function of N, is oscillating for $N \to \infty$. However, we can get an asymptotic error bound by letting N run through the sequence $(cn + t + k, n \in \mathbb{N})$, where t is kept fixed. Thus we get

THEOREM 4.10. *Fix* $(c, t) \in \mathbb{N}^2$ *such that* $1 \leqslant t \leqslant c \leqslant k$. *Then estimate*

(4.24) *holds for* $N := cn + t + k$ *and arbitrary* $n \in \mathbb{N}$, *and it has the asymptotic error bound (for* $n \to \infty$)

$$b := \frac{\beta^t \operatorname{sp}(V^k - V^{k-c})}{\min(\inf_s |\beta^t x + (1 - \beta^c)^+ V^\nu(s)|, \inf_s |\beta^t y + (1 - \beta^c)^+ V^\nu(s)|)} \qquad (4.27)$$

where $\nu := k + t - c$, $x := \inf(V^k - V^{k-c})$, $y := \sup(V^k - V^{k-c})$, *unless* $\beta < 1$ *and*

$$\inf_s |\beta^t x + (1 - \beta^c) V^\nu(s)| = \inf_s |\beta^t y + (1 - \beta^c) V^\nu(s)| = 0.$$

The proof is very similar to that given for Theorem 4.5, and is therefore omitted.

5. Estimates for V^N Which use Auxiliary Functions

In the sequel we understand by a fixed point of the operator U a map $v \in \mathfrak{S}_U$ that satisfies the so-called optimality equation (or Bellman equation)

$$v = Uv. \qquad (5.1)$$

A. At first we assume $\beta < 1$. Then, as a well known consequence (cf. Blackwell [2] and Strauch [14]) of the fixpoint theorem for contractions, there exists $V := \lim V^n$, V is the unique fixed point of U, and estimate (3.2) holds.

Now let us look what we can derive from the results of Section 4, assuming that V is known. (If S and A are finite, one possible way for computing V is the method of policy iteration.) At first we obtain easily from Corollary 4.3 the following.

Corollary 5.1. *For* $\beta < 1$ *holds:*

(i) *If* $\sup V^c \leqslant \beta_{m-1} \inf r' + \beta^m \inf V^c$ *for some* $c \leqslant N$, *and some* $m \in \mathbb{N}$, *then* $V^N \leqslant V$.

(ii) *If* $\inf V^c \geqslant \beta_{m-1} \sup r' + \beta^m \sup V^c$ *for some* $c \leqslant N$ *and some* $m \in \mathbb{N}$, *then* $V \leqslant V^N$.

Next we consider Theorem 4.8 for fixed ν, c, d and $N_m := \nu + mc$, $m \in \mathbb{N}$. Letting m tend to infinity and replacing afterwards ν by N, yields

THEOREM 5.2. *If* $\beta < 1$, *and* $1 \leqslant c \leqslant d \leqslant N + c$, *then*

$$\beta^{N+c-d}(1-\beta^c)^{-1}\inf(V^d - V^{d-c}) \leqslant V - V^N \leqslant \beta^{N+c-d}(1-\beta^c)^{-1}\sup(V^d - V^{d-c}). \tag{5.2}$$

Obviously, estimate (5.2) is admissible iff $d \leqslant k$, and for fixed $c \in \mathbb{N}$ the case $d = k$, i.e., the estimate

$$\beta^{N+c-k}(1-\beta^c)^{-1}\inf(V^k - V^{k-c}) \leqslant V - V^N \leqslant \beta^{N+c-k}(1-\beta^c)^{-1}\sup(V^k - V^{k-c}) \tag{5.3}$$

is according to (4.11) a best estimate among the set of those estimates (5.2) for which $c \leqslant d \leqslant k$. Moreover, (5.3) is improving with increasing k and fixed c. One realizes immediately that (5.3) improves and generalizes estimate (3.2). However, in subsection B we shall derive a further improvement on (5.3).

By the way, (5.2) provides us also with a set $E^4(c)$ of *estimates for* V, given $V^0, V^1, \ldots, V^k$, provided that N, d, c are restricted by the conditions $1 \leqslant c \leqslant d \leqslant \min(k, N + c)$ and $0 \leqslant N \leqslant k$; note that N in (5.2) corresponds to ν in (4.23). It follows from Theorem 4.9 that for fixed $c \leqslant k$ the set $E_4'(c)$ of those estimates

$$\beta^j(1-\beta^c)^{-1}\inf(V^k - V^{k-c}) \leqslant V - V^{k+j-c} \leqslant \beta^j(1-\beta^c)^{-1}\sup(V^k - V^{k-c}) \tag{5.4}$$

for which $1 \leqslant j \leqslant c$, is complete in $E_4(c)$. In particular, if $c = 1$, then

$$\beta(1-\beta)^{-1}\inf(V^{k-1} - V^k) \leqslant V^k - V \leqslant \beta(1-\beta)^{-1}\sup(V^{k-1} - V^k), \tag{5.5}$$

an estimate derived by Porteus [10, Theorem 1], is a best estimate within the set of estimates

$$\beta^{N+1-d}(1-\beta)^{-1}\inf(V^{d-1} - V^d) \leqslant V^N - V \leqslant \beta^{N+1-d}(1-\beta)^{-1}\sup(V^{d-1} - V^d), \tag{5.6}$$

for which $1 \leqslant d \leqslant \min(k, N + 1)$, $0 \leqslant N \leqslant k$. It follows that MacQueen's estimate (3.1), which is contained in (5.6) as the special case $d := k$, $N := k - 1$, is improved by Porteus' estimate (5.5).

B. Now we shall drop the assumption $\beta < 1$. Then, in general, $\lim_n V^n$ will not exist, and the method of subsection A is no longer applicable.

However, Hinderer/Hübner [5, Section 4] showed by an example that the operator U may have a fixed point also for values of β, for which $\lim_n V^n$ does not exist. Therefore, when working with fixed points of U instead of $\lim_n V^n$, we can still exploit the properties of U, and are thus able to extend some of our results from the case $\beta < 1$ to the case $\beta > 1$. However, we shall go one step further: We do not compare V^N only with fixed points of U, but more generally with "auxiliary" functions of the form $w^\nu := U^\nu w$, where w is an arbitrarily chosen element of $\mathfrak{S}_U$, and we shall derive bounds for V^N in terms of $V^0, V^1, ..., V^k$ and $w^0, w^1, ..., w^{k'}$, where k' is a given positive integer. Such an estimate will again be called *admissible*. The particular choice of w will lead to several types of estimates.

Our general approach is the following one: We select $\nu \in \mathbb{N}_0$ and $n \in \mathbb{N}_0$ and obtain from

$$V^N - w^\nu = (V^N - w^n) + (w^n - w^\nu)$$

and (4.11) for any $\rho \geqslant 0$ the upper estimate

$$V^N - w^\nu \leqslant \beta^\rho \sup(V^{N-\rho} - w^{n-\rho}) + \sup(w^n - w^\nu), \tag{5.7}$$

provided that $\rho \leqslant \min(N, n)$. And then we can use Lemma 4.8 with V^0 replaced by w and N replaced by n, in order to get an upper bound for $\sup(w^n - w^\nu)$. The same applies to lower bounds. Using in addition Theorem 4.9, one derives easily

THEOREM 5.3. (i) *If* $w \in \mathfrak{S}_U$, $w^\mu := U^\mu w$ *for* $\mu \in \mathbb{N}_0$, *and if* $(\nu, \rho, n, c, d) \in \mathbb{N}_0^3 \times \mathbb{N}^2$ *satisfies*

(a) $c \leqslant d \leqslant \nu + c$,

(b) $\nu < n$,

(c) $m := (n - \nu)/c \in \mathbb{N}$,

(d) $\rho \leqslant \min(n, N)$,

then

$$\beta^\rho \inf(V^{N-\rho} - w^{n-\rho}) + \beta^{\nu+c-d}(\beta^c)_{m-1} \inf(w^d - w^{d-c}) \leqslant V^N - w^\nu \leqslant \beta^\rho \sup(V^{N-\rho} - w^{n-\rho}) + \beta^{\nu+c-d}(\beta^c)_{m-1} \sup(w^d - w^{d-c}). \tag{5.8}$$

(ii) *Estimate* (5.8) *is admissible iff*

$$\rho \geqslant \max(N - k, n - k') \quad \textit{and} \quad \max(\nu, d) \leqslant k'.$$

(iii) *Fix* c *and* $n > k'$, *let* t *be the positive remainder of* $n - k'$ *under the* *division by c, and put $n' := (n - k' - t)/c$ and $\rho' := \max(N - k, n - k')$. Then*

$$\beta^{\rho'} \inf(V^{N-\rho'} - w^{n-\rho'}) + \beta^t(\beta^c)_{n'} \cdot \inf(w^{k'} - w^{k'-c}) \\ \leqslant V^N - w^{k'+t-c} \leqslant \beta^\rho \sup(V^{N-\rho'} - w^{n-\rho'}) + \beta^t(\beta^c)_{n'} \cdot \sup(w^{k'} - w^{k'-c}) \tag{5.9}$$

is a best estimate among those estimates (5.8) *which are admissible.*

Note that the condition

$$\max(N - k, n - k') \leqslant \rho \leqslant \min(n, N),$$

which is necessary for (5.8) to be admissible, can only be fulfilled if $N - k \leqslant n \leqslant N + k'$.

Our first special case of Theorem 5.3 is:

THEOREM 5.4. (i) *If w is a fixed point of U and $0 \leqslant k < N$, then*

$$\beta^{N-k} \inf(V^k - w) \leqslant V^N - w \leqslant \beta^{N-k} \sup(V^k - w). \tag{5.10}$$

(ii) *If $\beta = 1$, and $w \in \mathfrak{S}_U$ is a solution of $w + g = Uw$ for some constant $g \in \mathbb{R}$, then*

$$\inf(V^{N-\rho} - w) + (\nu - n + \rho + cm) g \\ \leqslant V^N - w \leqslant \sup(V^{N-\rho} - w) + (\nu - n + \rho + cm) g \tag{5.11}$$

holds for all $(\rho, \nu, n, c) \in \mathbb{N}_0^3 \times \mathbb{N}$ such that

(a) $\nu < n$,

(b) $(n - \nu)/c \in \mathbb{N}$,

(c) $N - k \leqslant \rho \leqslant \min(n, N)$.

(iii) *The estimate*

$$\inf(V^k - w) + (N - k) g \leqslant V^N - w \leqslant \sup(V^k - w) + (N - k) g \tag{5.12}$$

is a best estimate within the set of estimates (5.11).

(iv) *The estimates* (5.10) *and* (5.12) *are improving with increasing k.*

Proof. A direct proof of part (i) can be given by means of Lemma 4.4, but we can deal jointly with parts (i) and (ii), since in both cases $w + g = Uw$ holds for some constant g. It follows, that $U^\lambda w = \beta_{\lambda-1} g + w$ for all $\lambda \in \mathbb{N}$.

Inserting this equation into (5.8) shows that for arbitrary (ν, ρ, n, c, d) satisfying conditions (a)–(d) of Theorem 5.3, we have e.g. the upper bound

$$V^N - w \leqslant \beta^\rho \sup(V^{N-\rho} - w) + g(\beta_{\nu-1} - \beta_{\rho,n-1} + \beta_{\nu,\nu+c-1}(\beta^c)_{m-1}). \quad (5.13)$$

Now part (i) follows with $g = 0$, $c = d = 1$, $\nu = 0$, $\rho = n = N - k$. If $\beta = 1$, then (5.13) simplifies to

$$V^N - w \leqslant \sup(V^{N-\rho} - w) + g(\nu - n + \rho + cm), \quad (5.14)$$

which is independent of d. If we put $d = 1$ and $k' = n - 1$, then the conditions (a)–(d) of Theorem 5.3 together with the admissibility condition in part (ii) of Theorem 5.3 are equivalent to the conditions (a)–(c) in part (ii) of the present theorem. In order to prove part (iii) we use part (iii) of Theorem 5.3 with $k' = n - 1$, hence $\rho' = N - k$, $t = 1$, $n' = 0$. It follows that for any fixed c and n the estimate (5.12) is a best estimate among those estimates (5.11) for which $\max(N - k, n - k') \leqslant \rho$ and $\nu \leqslant k'$. Since the latter conditions are no real restrictions, part (iii) is proved. Finally, part (iv) follows quickly by means of Lemma 4.4. ∎

Remark 1. Assume $\beta < 1$. Then estimate (5.10) with $w = V$ is an improvement of estimate (5.3): In fact, for any $n \in \mathbb{N}$ we have

$$\begin{aligned}\sup(V^k - V) &\leqslant \sum_{\nu=0}^{n-1} \sup(V^{k+\nu c} - V^{k+(\nu+1)c}) + \sup(V^{k+nc} - V) \\ &\leqslant \sup(V^{k-c} - V^k)\,(\beta^c)_{1,n} + \| V^{k+nc} - V \| \\ &\to \beta^c(1 - \beta^c)^{-1} \sup(V^{k-c} - V^k), \qquad (n \to \infty),\end{aligned}$$

from which the desired result follows.

Remark 2. Estimate (4.9), which was derived from Rieder's estimate (3.12) as an improvement of Martin's estimate (3.6), is improved by (5.10). In fact, if $\beta < 1$, then

$$(1 - \beta)^{-1} \inf r' \leqslant V \leqslant (1 - \beta)^{-1} \sup r' \quad (5.15)$$

follows from (4.6), and then (5.10) implies the estimate

$$\begin{aligned}&\beta^{N-k}(\inf V^k - (1 - \beta)^{-1} \sup r') \\ &\quad \leqslant V^N - V \leqslant \beta^{N-k}(\sup V^k - (1 - \beta)^{-1} \inf r'),\end{aligned} \quad (5.16)$$

which, for $k = 0$, improves on (4.9).

THEOREM 5.5. (i) *Estimate* (5.10) *has the asymptotic error bound*

$$b_k := \begin{cases} 0, & \text{if } \beta < 1 \text{ and if } w(s) \neq 0 \text{ for all } s \in S, \\ \dfrac{\operatorname{sp}(V^k - w)}{\min(|\sup(V^k - w)|, |\inf(V^k - w'|)}, & \text{if } \beta > 1. \end{cases} \tag{5.17}$$

(ii) *Estimate* (5.12) *has the asymptotic error bound zero, provided that* $g \neq 0$.

The proof is similar to that given for Theorem 4.5, and is therefore omitted. Now we are going to compare V^N with $r' := \sup_{a \in D(\cdot)} r(\cdot, a)$.

THEOREM 5.6. (i) *If* $0 \leqslant k < N$, *then*

$$\begin{aligned} &\beta_{1,N-k-1} \inf r' + \beta^{N-k} \inf V^k \\ &\qquad \leqslant V^N - r' \leqslant \beta_{1,N-k-1} \sup r' + \beta^{N-k} \sup V^k. \end{aligned} \tag{5.18}$$

(ii) *Estimate* (5.8) *has the asymptotic error bound*

$$b_k := \frac{\operatorname{sp} r' + (\beta - 1)^+ \operatorname{sp} V^k}{\min(\inf_s |x(s)|, \inf_s |y(s)|)}, \tag{5.19}$$

where

$$\begin{aligned} y(s) &:= \sup r' + (\beta - 1)^+ \sup V^k + \beta^{-1}(1 - \beta)^+ r'(s), \\ x(s) &:= \inf r' + (\beta - 1)^+ \inf V^k + \beta^{-1}(1 - \beta)^+ r'(s), \end{aligned}$$

unless $\beta < 1$ *and*

$$\inf_s |\beta \sup r' + (1 - \beta) r'(s)| = \inf_s |\beta \inf r' + (1 - \beta) r'(s)| = 0.$$

(iii) *Estimate* (5.18) *is improving with increasing* k.

Proof. (i) If $k \leqslant N - 2$, then the assertion follows from part (i) of Theorem 5.3 with $w \equiv 0$, $\nu = c = d = 1$ and $\rho = n = N - k$. If $k = N - 1$, then (5.18) follows easily from

$$\beta \inf v + r' \leqslant Uv \leqslant \beta \sup v + r', \qquad v \in \mathfrak{S}_U. \tag{5.20}$$

(ii) The proof is similar to that of Theorem 4.5 and is therefore omitted. Finally, part (iii) is an easy consequence of (5.20). ∎

We remark that (5.18) is an improvement of (4.6), since, e.g.

$$r' + \beta_{1,N-k-1} \sup r' \leqslant \beta_{N-k-1} \sup r'.$$

6. Estimates for "Good" Policies

At first we must introduce a certain subset of $\mathfrak{S}_U$. Let $\mathfrak{D}$ be the system of those measurable subsets of D which contain the graph of a measurable map from S into A. For any $K \in \mathfrak{D}$ we define an operator $U_K: \mathfrak{S}_m \to \mathbb{R}^S$ by means of

$$(U_K v)(s) := \sup_{a \in K(s)} \left[r(s, a) + \beta \int q(s, a, dt)\, v(t) \right], \qquad s \in S. \tag{6.1}$$

Denote by $\mathfrak{S}_D$ the set of those functions v in $\mathfrak{S}_m$ which have the following property: For every sequence of sets $K_n \in \mathfrak{D}$, $n \in \mathbb{N}_0$, we have $U_{K_0} v \in \mathfrak{S}_m$, $U_{K_1} U_{K_0} v \in \mathfrak{S}_m$, $U_{K_2} U_{K_1} U_{K_0} v \in \mathfrak{S}_m$,.... The set $\mathfrak{S}_D$ has the following properties:

(i) $\mathfrak{S}_D$ is the largest of those subsets of $\mathfrak{S}_m$ which are mapped into itself by each of the operators U_K, $K \in \mathfrak{D}$. In particular, $\mathfrak{S}_D$ is a subset of $\mathfrak{S}_U$.

(ii) $\mathfrak{S}_D$ is under the sup-metric a closed subset of the Banach space $\mathfrak{S}_m$, hence a complete metric space.

(iii) $\mathfrak{S}_D$ contains the set $\mathfrak{S}_b$ of all bounded and $\mathfrak{S}$-measurable functions from S into $\mathbb{R}$. This can be seen by an application of Theorems 13.2 and 14.4 in Hinderer [3], since using $D_n \in \mathfrak{D}$ instead of D at stage n, $n \in \mathbb{N}_0$, yields a dynamic program that is nonstationary in the "restriction sets D_n".

Particularly important is the special case where K consists of the graph of a decision function $f \in F$. Then we denote U_K by U_f,[4] and we have

$$U_f v(s) = r(s, f(s)) + \beta \int q(s, f(s), dt)\, v(t), \qquad s \in S. \tag{6.2}$$

It follows from property (iii) above that for any $v \in \mathfrak{S}_b$, $m, n \in \mathbb{N}_0$ and $f_\mu \in F$, $0 \leqslant \mu \leqslant m$, the function $U_{f_0} U_{f_1} \cdots U_{f_m} U^n v$ is defined and belongs to $\mathfrak{S}_D$. Moreover, it is easy to see and well known that the expected conditional total reward $V_\pi{}^n$ under the n-stage policy $\pi = (f_\nu)$ has the representation

$$V_\pi{}^n = U_{f_0} U_{f_1} \cdots U_{f_{n-1}} V^0. \tag{6.3}$$

Let us call a decision function $f \in F$ a *maximizer* of the function $v \in \mathfrak{S}_D$, if for all $s \in S$ the map

$$a \to r(s, a) + \beta \int q(s, a, dt)\, v(t)$$

attains its supremum at $f(s)$. In other words, f is a maximizer of v iff $U_f v = Uv$. An essential part of the method of backward induction in dynamic

[4] In former papers, e.g. in Hinderer [3], we denoted U_f by L_f.

programming says (cf. Hinderer [3, Theorem 17.6]), that a sufficient (but not necessary) condition for the optimality of the N-stage policy $(\varphi_N, \varphi_{N-1}, ..., \varphi_1)$ is the following one: φ_{n+1} is a maximizer of V^n, $0 \leqslant n < N$.

Now assume that we have carried through value iteration for k steps, $k \geqslant 1$, and that we have found maximizers φ_{n+1} of V^n, $0 \leqslant n < k$. What are natural candidates for a good N-stage policy π? One candidate always at hand is obtained by *extrapolation*. We just continue to use the maximizer φ_k until stage N, i.e. we suspect that $\pi := (\varphi_k, \varphi_k, ..., \varphi_k, \varphi_{k-1}, ..., \varphi_1) \in \Delta_N$ is "good". (Note that in case $k = 1$ the policy π is myopic). If U has a fixed point w, which has a maximizer f, then also $\pi := (f, f, ..., f, \varphi_k, \varphi_{k-1}, ..., \varphi_1) \in \Delta_N$ may be a good policy. The conjecture about the goodness of the policy π is under certain assumptions supported by the turnpike-theorem for the horizon, given by Shapiro [13] and improved by Hinderer/Hübner [5].

Our general method for deriving upper bounds for $V^N - V_\pi^N$, where π is a "good" policy, is the following one. Let $\varphi_{\nu+1}$ be a maximizer of V^ν for $0 \leqslant \nu \leqslant k$, and let f be an appropriately chosen decision function. Then we consider the N-stage policy $\pi := (f, f, ..., f, \varphi_k, \varphi_{k-1}, ..., \varphi_1)$. At first we get, if v is an "auxiliary" function

$$V^N - V_\pi^N = (V^N - v) - (V_\pi^N - v). \tag{6.4}$$

Now, if v is appropriately chosen, we have available bounds for $V^N - v$, derived in Sections 4 and 5. Moreover, since our basic inequality (4.11) holds obviously also for the operator U_f, it will turn out, that for appropriately chosen v and f the lower bounds for $V^N - v$ carry over to corresponding lower bounds for $V_\pi^N - v$. Combining both bounds yields an upper bound for $V^N - V_\pi^N$ according to (6.4). We shall carry through this program only for the most important of our estimates, namely (4.13), (4.24), (5.10), (5.12) and (5.18).

THEOREM 6.1. (i) *Assume $1 \leqslant c \leqslant k < N$, and assume that $\varphi_{\nu+1}$ is a maximizer of V^ν for $0 \leqslant \nu < k$. Let t be the positive remainder of $N - k$ under division by c, and put $n := (N - k - t)/c$. Put*

$$\tau := (\varphi_{k+t-c}, \varphi_{k+t-c-1}, ..., \varphi_{k-c+1}) \qquad \text{and} \qquad \sigma := (\varphi_k, \varphi_{k-1}, ..., \varphi_{k-c+1}).$$

Then the N-stage policy $\pi := (\tau, \sigma, \sigma, ..., \sigma, \varphi_k, \varphi_{k-1}, ..., \varphi_1)$ satisfies

$$0 \leqslant V^N - V_\pi^N \leqslant \beta^t(\beta^c)_n \operatorname{sp}(V^k - V^{k-c}). \tag{6.5}$$

In particular, we have for $c = 1$

$$0 \leqslant V^N - V_\pi^N \leqslant \beta_{1,N-k} \operatorname{sp}(V^k - V^{k-1}). \tag{6.6}$$

(ii) *For any myopic N-stage policy* π *we have*

$$0 \leqslant V^N - V_\pi^N \leqslant \beta_{1,N-1}\,\mathrm{sp}(V^1 - V^0). \tag{6.7}$$

(iii) *The instimates* (6.5) *and* (6.6), *considered as estimates for* $V^N - V_\pi^N$, *are improving with increasing* k *and fixed* c.

Proof. (i) For any policy $\rho := (f_0, f_1, ..., f_{n-1})$ put

$$M_\rho v := U_{f_0} U_{f_1} \cdots U_{f_{n-1}} v, \qquad v \in \mathfrak{S}_D .$$

It follows from the optimality criterion, stated above, that $M_\psi V^0 = V^\nu$ for $1 \leqslant \nu \leqslant k$ and $\psi = (\varphi_\nu, \varphi_{\nu-1}, ..., \varphi_1)$. Now we get, using (4.11) for the operators U_{φ_ν},

$$V_\pi^N - V^{k+t-c} = \sum_{\nu=0}^{n} (M_\tau M_\sigma^\nu V^k - M_\tau M_\sigma^\nu V^{k-c}) \geqslant \beta^t (\beta^c)_n \inf(V^k - V^{k-c}).$$

Now (6.5) follows from (6.4) with $v := V^{k+t-c}$ and (4.24), whereas (6.6) is the special case $c = 1$. (ii) Estimate (6.7) is obtained from (6.6) with $k = 1$. (iii) The assertion follows from (4.12), distinguishing for (6.5) the two cases $t > 1$ and $t = 1$ for the comparison of the estimate for k and the estimate for $k + 1$. ∎

THEOREM 6.2. *Assume that* $k \in \mathbb{N}_0$, *that* $\varphi_{\nu+1}$ *is a maximizer of* V^ν *for* $0 \leqslant \nu < k$, *and that there exists a constant* $g \in \mathbb{R}$ *and* $w \in \mathfrak{S}_D$ *such that* $w + g = Uw$. *Then, if* f *is a maximizer of* w, *the N-stage policy* $\pi := (f, f, ..., f, \varphi_k, \varphi_{k-1}, ..., \varphi_1)$ *satisfies*

$$0 \leqslant V^N - V_\pi^N \leqslant \beta^{N-k}\,\mathrm{sp}(V^k - w), \tag{6.8}$$

and this estimate for $V^N - V_\pi^N$ *is improving with increasing* k.

Proof. The properties of f and w imply $U_f w = Uw = w + g$, hence $U_f^\nu w = \beta_{\nu-1} g + w$, $\nu \in \mathbb{N}$. We obtain, using (4.11)

$$V_\pi^N - w = U_f^{N-k} V^k - U_f^{N-k} w + \beta_{N-k-1} g \geqslant \beta^{N-k} \inf(V^k - w) + \beta_{N-k-1} g.$$

Now (6.8) follows from (5.13) with $c = d = 1$, $\nu = 0$, $\rho = n = N - k$ and from (6.4) with $v := w$. That (6.8) is improving with increasing k is a consequence of (4.12), observing that $\mathrm{sp}(v + g) = \mathrm{sp}\, v$ for any constant g. ∎

THEOREM 6.3. *Assume that* $\varphi_{\nu+1}$ *is a maximizer of* V^ν *for* $0 \leqslant \nu < k$, *where* $k \in \mathbb{N}_0$, *and assume that there exists a decision function* f *such that*

$a \to r(s, a)$ attains its supremum at $f(s)$, $s \in S$. Then the N-stage policy $\pi := (f, f, ..., f, \varphi_k, \varphi_{k-1}, ..., \varphi_1)$ satisfies

$$0 \leqslant V^N - V_\pi^N \leqslant \beta_{1,N-k-1} \operatorname{sp} r' + \beta^{N-k} \operatorname{sp} V^k. \tag{6.9}$$

Proof. It is easy to see that $U_f v \geqslant r' + \beta \inf v$ for $v \in \mathfrak{S}_D$.
Then induction on n yields

$$U_f^n v \geqslant r' + \beta_{1,n-1} \inf r' + \beta^n \inf v, \qquad n \in \mathbb{N}. \tag{6.10}$$

Now we get

$$V_\pi^N = U_f^{N-k} V^k \geqslant r' + \beta_{1,N-k-1} \inf r' + \beta^{N-k} \inf V^k,$$

and the assertion follows from (5.18). ∎

If $V^0 \equiv 0$ and $k = 0$, then the policy π in Theorem 6.3 is myopic, and the estimates (6.9) and (6.7) coincide.

In general, in order to determine the policies π in Theorem 6.1 and 6.2, one will need to compute also $V^0, V^1, ..., V^k$. If only $V^0, V^1, ..., V^l$ for some $l < k$ is available and hence π is not known, one may still be interested to get upper bounds for $V^N - V_\pi^N$, because these bounds will indicate whether it pays to compute also $V^{l+1}, V^{l+2}, ..., V^k$, in order to find the policy π. Such estimates may be obtained with the aid of (4.12). In fact, it is easily seen, that the estimates (6.5), (6.6) and (6.8) can be generalized—under the assumptions stated there—to

$$0 \leqslant V^N - V_\pi^N \leqslant \beta^{t+b}(\beta^c)_n \operatorname{sp}(V^{k-b} - V^{k-b-c}), \qquad 0 \leqslant b \leqslant k - c, \tag{6.5'}$$

$$0 \leqslant V^N - V_\pi^N \leqslant \beta_{1+b,N-k+b} \operatorname{sp}(V^{k-b} - V^{k-b-1}), \qquad 0 \leqslant b \leqslant k - 1, \tag{6.6'}$$

$$0 \leqslant V^N - V_\pi^N \leqslant \beta^{N-k+b} \operatorname{sp}(V^{k-b} - w), \qquad 0 \leqslant b \leqslant k. \tag{6.8'}$$

In some cases the estimates (6.5'), (6.6'), and (6.8') for $b > 0$ can be improved considerably by means of the following result of Hübner [7, Theorem 1.1]:

LEMMA 6.4. *Assume that S and A are finite. Put*

$$\alpha := \max_{\substack{(s,a) \in D \\ (s',a') \in D}} \left[1 - \sum_{t \in S} \min\{q(s, a, t), q(s', a', t)\} \right]. \tag{6.11}$$

Then for $k \in \mathbb{N}$ holds

$$\operatorname{sp}(U^k v - U^k w) \leqslant (\alpha\beta)^k \operatorname{sp}(v - w), \qquad v, w \in \mathbb{R}^S. \tag{6.12}$$

Note that (6.12) improves on (4.14), since $\alpha \leqslant 1$.

Now it follows from Lemma 6.4 that, in case of finite state space and finite action space, the upper bounds in (6.5′), (6.6′) and (6.8′) may be multiplied by α^b.

7. Numerical Example

In this section we are going to test the estimates derived in Sections 4, 5 and 6. We choose a very simple example, namely a slight variant of Howard's well known toymaker example, since here an explicit solution—which is needed for purposes of comparison—is easily obtainable. However, if $V^0, V^1, \ldots, V^k$ are already available, then the estimates may be easily computed also in much more complicated examples.

The example is defined by $S := A := D(s) := \{1, 2\}$, q and r as in Table 1, $\beta \in (0, \infty)$, and $V^0(1) := 105$, $V^0(2) := 100$.

TABLE 1

(s, a)	$q(s, a, 1)$	$q(s, a, 2)$	$r(s, a)$
(1, 1)	0.5	0.5	6
(1, 2)	0.8	0.2	4
(2, 1)	0.4	0.6	−3
(2, 2)	0.7	0.3	−5

A. At first we shall derive an explicit solution of the problem for arbitrary horizon N and for β in a neighborhood of 1. Using an idea in Hinderer/Hübner [5, Sect. 4], we prove

Lemma 7.1. *If $\beta \geqslant \beta' := (-27 + (849)^{1/2})/3 \approx 0.712$, then the decision function $g \equiv 2$ is a maximizer of V^n for all $n \geqslant 1$; and if $\beta \leqslant \frac{4}{3}$, then $f \equiv 1$ is a maximizer of V^0.*

Proof. Denote by M_{n+1} the set of maximizers of V^n, $n \in \mathbb{N}_0$, and put for $v \in \mathbb{R}^s$

$$(Lv)(s, a) := r(s, a) + \beta \sum_{t \in S} q(s, a, t)\, v(t), \qquad s \in S. \tag{7.1}$$

and $z(v) := v(1) - v(2)$. From

$$Lv(s, 1) - Lv(s, 2) = 2 - 0.3\beta(v(1) - v(2)), \qquad s \in S,$$

we conclude that $g \in M_{n+1}$ iff $z(V^n) \geqslant 20/3\beta$, and that $f \in M_{n+1}$ iff $z(V^n) \leqslant 20/3\beta$; in particular $\{f, g\} \cap M_{n+1} \neq \varnothing$ for all $n \in \mathbb{N}_0$. Therefore, since

$$Lv(1, a) - Lv(2, a) = 9 + 0.1\beta z(v), \qquad a \in A,$$

we get $z(V^n) = 9 + 0.1\beta z(V^{n-1})$, $n \in \mathbb{N}$; hence

$$z(V^n) = 9 \sum_{m=0}^{n-1} (0.1\beta)^m + 5(0.1\beta)^n, \qquad n \in \mathbb{N}.$$

Since $z(V^n)$ is increasing in n, we know that $g \in M_{n+1}$ for all $n \geqslant 1$, iff $z(V^1) = 9 + 0.5\beta \geqslant 20/3\beta$, i.e. if $\beta \geqslant \beta'$. Moreover, $f \in M_1$ iff $z(V^0) = 5 \leqslant 20/3\beta$, i.e. if $\beta \leqslant \frac{4}{3}$. ▮

Now let us fix $\beta \in \langle \beta', \frac{4}{3} \rangle$. Then the value iteration (2.3) simplifies by means of Lemma 7.1 to the linear system of difference equations

$$V^{n+1} = c + \beta B V^n, \qquad n \in \mathbb{N}, \tag{7.2}$$

where

$$c := \begin{pmatrix} 4 \\ -5 \end{pmatrix} \quad \text{and} \quad B := \begin{pmatrix} 0.8 & 0.2 \\ 0.7 & 0.3 \end{pmatrix},$$

and the initial values

$$V^1 = \begin{pmatrix} 6 + 102.5\beta \\ -3 + 102\beta \end{pmatrix}. \tag{7.3}$$

From (7.2) we get

$$V^{n+1} = \sum_{\nu=0}^{n-1} \beta^\nu B^\nu c + \beta^n B^N V^1, \qquad n \in \mathbb{N}. \tag{7.4}$$

Now one easily computes, e.g. by induction on n, that

$$B^n = 9^{-1} \left[\begin{pmatrix} 7 & 2 \\ 7 & 2 \end{pmatrix} + (0.1)^n \begin{pmatrix} 2 & -2 \\ -7 & 7 \end{pmatrix} \right], \qquad n \in \mathbb{N}_0. \tag{7.5}$$

Straightforward calculation yields then

LEMMA 7.2. *For arbitrary* $\beta \in \langle \beta', \frac{4}{3} \rangle$ *and* $N \geqslant 2$ *the toymaker example has the following solution:*

(i) *If* $\beta \neq 1$, *then*

$$V^N = 2(1-\beta)^{-1} + (1-0.1\beta)^{-1}\begin{pmatrix} 2 \\ -7 \end{pmatrix}$$
$$+ \beta^{N-1}(4 + 9^{-1} \times 921.5\beta - 2(1-\beta)^{-1})$$
$$+ (0.1\beta)^{N-1}\left[9^{-1}\begin{pmatrix} 18+\beta \\ -63-3.5\beta \end{pmatrix} - (1-0.1\beta)^{-1}\begin{pmatrix} 2 \\ -7 \end{pmatrix}\right].$$

(ii) *If* $\beta = 1$, *then*

$$V^N = 2N + 9^{-1}\begin{pmatrix} 959.5 \\ 869.5 \end{pmatrix} + 9^{-1}(0.1)^{N-1}\begin{pmatrix} -1 \\ 3.5 \end{pmatrix}.$$

Table 2 contains values of V^N for $N = 1, 4, 5, 9, 10, 100, 1000$ and $\beta = 0.98, 1, 1.1$, obtained by means of Lemma 7.2.

TABLE 2

	$\beta = 0.98$		$\beta = 1$	
N	$V^N(1)$	$V^N(2)$	$V^N(1)$	$V^N(2)$
1	106.450 000	96.960 000	108.500 000	99.000 000
4	106.302 012	96.325 644	114.611 000	104.611 500
5	106.221 388	96.243 606	116.611 100	106.611 150
9	105.910 552	95.932 725	124.611 111	114.611 111
10	105.836 687	95.858 860	126.611 111	116.611 111
100	102.804 760	92.826 933	306.611 111	296.611 111
1000	102.217 295	92.239 468	2106.611 111	2096.611 111

	$\beta = 1.1$	
N	$V^N(1)$	$V^N(2)$
1	118.750 000	109.200 000
4	164.098 597	153.986 986
5	182.283 902	172.171 625
9	275.120 967	265.008 607
10	304.408 344	294.295 984
100	$1.711\,632_{6}$	$1.711\,622_{6}$
1000	$3.067\,831_{43}$	$3.067\,831_{43}$

In order to apply the estimates in Sections 5 and 6 we need solutions w of $w + g = Uw$ for some constant g. Now, if $\beta \notin \{1, 10\}$, then

$$2(1-\beta)^{-1}(10-\beta)^{-1}\binom{20-11\beta}{34\beta-25}$$

is a fixed point of U, as shown in Hinderer/Hübner [5, Sect. 4]. Moreover, if $\beta = 1$, then $w := \binom{10}{0}$ satisfies $w + 2 = Uw$; cf. Howard [6, p. 41].

As a measure for the quality of our estimates of the form $v_1 \leqslant V^N \leqslant v_2$ and $0 \leqslant V^N - V_\pi{}^N \leqslant v_3$ we shall use the larger of the maximal absolute relative errors of v_1 and v_2, i.e.

$$\rho := \max[\|(V^N - v_1)/V^N\|, \|(V^N - v_2)/V^N\|], \tag{7.6}$$

and the numbers

$$\rho' := \| v_3/V^N \|. \tag{7.7}$$

Due to the unavoidable round-off errors, only (close) upper bounds for ρ and ρ' are given in Tables 3 and 4. We used only the most important of our estimates, namely (4.13), (4.24), (5.10), (5.12), (5.18), (6.6), (6.8) and (6.9).

Concerning our simple example we can draw the following *conclusions*:

TABLE 3

Upper Bounds for the Maximal Absolute Relative Error ρ of Estimates for V^N

	$N = 100$			$N = 1000$		
Estimate	0.98	1	1.1	0.98	1	1.1
(4.13) $k = 1$	1.4	1	0.26	1.5	1.5	0.26
(4.13) $k = 5$	1.5_{-4}	1.2_{-4}	2.9_{-5}	1.7_{-4}	1.7_{-4}	2.9_{-5}
(4.13) $k = 10$	7_{-8}	3.4_{-9}	6_{-7}	8_{-8}	5_{-10}	6.6_{-7}
(4.24) $k = 10, c = 5$	3_{-6}	3_{-6}	6_{-7}	3.7_{-6}	4_{-6}	3.3_{-7}
(5.10) $k = 0$	4.8_{-3}	—	3_{-2}	1_{-6}	—	3_{-2}
(5.10) $k = 5$	7_{-8}	—	1_{-6}	1_{-8}	—	3.3_{-7}
(5.12) $k = 0$	—	1.2_{-2}	—	—	1.7_{-3}	—
(5.12) $k = 5$	—	1.4_{-7}	—	—	2_{-8}	—
(5.18) $k = 5$	2.3	1.6	0.3	2.7	2.4	0.29

TABLE 4

Upper Bounds for the Maximal Absolute Relative Error ρ' of Upper Estimates for $V^N - V_\pi{}^N$

	$N = 100$			$N = 1000$		
Estimate	0.98	1	1.1	0.98	1	1.1
(6.6) $k = 1$	2.05	1.51	0.40	2.4	2.15	0.37
(6.6) $k = 5$	1.9_{-4}	1.5_{-4}	3.7_{-5}	2.2_{-4}	2.2_{-4}	3.7_{-5}
(6.6) $k = 10$	4.5_{-7}	3_{-7}	3.5_{-8}	5.4_{-7}	4.8_{-7}	3.5_{-8}
(6.8) $k = 5$	7.2_{-8}	1.7_{-7}	4.2_{-7}	9.1_{-16}	2.4_{-8}	4.2_{-7}
(6.9) $k = 5$	4.1	2.9	0.5	4.8	4.3	0.5

(a) The estimates (4.13), (5.10), (5.12), (6.6) and (6.8) are excellent for all values of N and β considered and for $k = 10$, they are still very good for $k = 5$, and (5.10) and (5.12) are satisfactory even for $k = 0$.

(b) The estimates (5.18) and (6.9) are very poor.

(c) For fixed k, estimates (5.10) and (5.12) are considerably better than (4.13); however, the latter estimate does not need computation of the auxiliary function w.

The following example, due to G. Hübner, shows that the generalization (4.24) of (4.13) may be useful: We take $S = A = D(s) = \{1, 2\}$, q the deterministic law that prescribes under any action the transition from $s = 1$ to $s = 2$ and from $s = 2$ to $s = 1$, $r(1, 1) = 1$ and $r(s, a) = 0$ otherwise, $\beta = 1$ and $V^0 \equiv 0$. Then (4.24) with $k = c = 2$, $N = 2n + 3$ and $N = 2n + 4$ yields the exact values of V^N for $N \geqslant 3$, viz.

$$V^N = \begin{cases} \binom{n+2}{n+1}, & \text{if} \quad N = 2n + 3, \\ \binom{n+2}{n+2}, & \text{if} \quad N = 2n + 4. \end{cases}$$

On the other hand, (4.13) for $k = 2$ yields only the poor estimates

$$\binom{1}{1} \leqslant V^N \leqslant \binom{N-1}{N-1}, \qquad N \geqslant 3.$$

Acknowledgments

The author is grateful to G. Hübner for contributing the preceding example, to W. Wierich for a revision of the manuscript and to W. Nicolai for checking the numerical calculations.

References

1. M. J. Beckmann, "Dynamic Programming of Economic Decisions," Springer, Berlin, 1968.
2. D. Blackwell, Discounted dynamic programming, *Ann. Math. Stat.* **36** (1965), 226–235.

2(a). D. Blackwell, D. Freedman, and M. Orkin, The optimal reward operator in dynamic programming, *Ann. Prob.* **2** (1974), 926–941.

3. K. Hinderer, "Foundations of Nonstationary Dynamic Programming with Discrete Time Parameter," Lecture Notes in Operations Research and Mathematical Systems, Vol. 33, Springer, Berlin, 1970.
4. K. Hinderer, On stochastic dynamic programs with absorbing sets, to be published.
5. K. Hinderer and G. Hübner, An improvement of J. F. Shapiro's turnpike-theorem for the horizon of finite stage, discrete dynamic programs, to appear in "Trans. 7th Prague Conference on Information Theory, Statistical Decision Functions, and Random Processes, Prague 1974."
6. R. Howard, "Dynamic Programming and Markov Processes," Technology Press and Wiley, New York, 1960.
7. G. Hübner, Improved procedures for eliminating suboptimal actions in Markov programming by the use of contraction properties, to be published in "Trans. 7th Prague Conference on Information Theory, Statistical Decision Functions, and Random Processes, Prague 1974."
8. J. MacQueen, A modified dynamic programming method for Markovian decision problems, *J. Math. Anal. Appl.* **14** (1966), 38–43.
9. J. J. Martin, "Bayesian Decision Problems and Markov Chains," Wiley, New York, 1967.
10. E. L. Porteus, Some bounds for discounted sequential decision processes, *Management Sc.* **18** (1971), 7–11.
11. U. Rieder, Bayessche dynamische Entscheidungs- und Stoppmodelle, Dissertation, Hamburg University, 1972.
12. H. Schellhaas, Zur Extrapolation in Markoffschen Entscheidungsmodellen mit Diskontierung, *Zeitschr. Oper. Res.* **18** (1974), 91–104.
13. J. F. Shapiro, Turnpike planning horizons for a Markovian decision model, *Management Sc.* **14** (1968), 292–300.
14. R. E. Strauch, Negative dynamic programming, *Ann. Math. Stat.* **37** (1966), 871–890.