The Annals of Mathematical Statistics
1972, Vol. 43, No. 6, 1894-1901

MARKOV DECISION PROCESSES WITH A NEW OPTIMALITY CRITERION: SMALL INTEREST RATES

BY STRATTON C. JAQUETTE

Science Applications, Incorporated

Finite state and action discrete time Markov decision processes with discounting are considered under the criterion of moment optimality. The case of small interest rates is studied, in particular the behavior of optimal policies as the interest rate approaches zero. Laurent expansions in the interest rate are developed for all moments of return for stationary policies, and a proof is given that there is a stationary policy which is moment optimal for all sufficiently small interest rates.

1. Introduction. In this paper the behavior of optimal policies for Markov decision processes is considered for the case of small interest rates. The problem was first discussed by Blackwell [1] for finite state discrete time Markov decision processes under the usual optimality criterion of maximizing the expected total discounted return. The result obtained in that paper was that a single stationary policy was optimal for all interest rates sufficiently close to zero. This work was extended by Miller and Veinott [5] who developed methods for constructing such an optimal policy, and by Veinott [6] who presented a unified treatment of the material with several generalizations. In this paper we consider these same problems for a new optimality criterion called moment optimality developed and discussed in Jaquette [2], [3], and [4].

Most of the notation used in this paper is standard for finite state and action discrete time Markov decision processes with discounting. Some notation peculiar to the criterion of moment optimality is explained in detail in Jaquette [3] or [4], and we give only a brief discussion of this notation here. As usual the finite state space is given by $S = \{1, 2, \cdots, s\}$, the finite action set associated with the ith state is A_i, and the set of all possible action vectors is $F \equiv \times_{i=1}^{s} A_i$. A policy is a sequence of elements of F, e.g., $\pi = (\mathbf{f}_1, \mathbf{f}_2, \cdots)$. Using the policy π means that if the process is in state i at time t, then the action taken is $f_t(i)$. We denote $\mathbf{f}_t$ for the policy π by $\pi(t)$. Let $f^\infty \equiv (\mathbf{f}, \mathbf{f}, \mathbf{f}, \cdots)$, then f^∞ is said to be stationary. Using any action vector $\mathbf{f}$, $\mathbf{r}(f)$ is the associated column vector of one period returns and $P(f)$ is the associated matrix of transition probabilities. Let $X_\pi(t)$ be the random state of the process at time t using policy π, and let $\mathbf{X}_\pi(t)$ be the random vector whose ith component is 1 if $X_\pi(t) = i$ and 0 otherwise. If $\mathbf{x}$ and $\mathbf{y}$ are vectors with components $x(i)$ and $y(i)$ respectively, $\mathbf{x} \circ \mathbf{y}$ is defined as the vector whose components are given by $[\mathbf{x} \circ \mathbf{y}]_i = x(i)y(i)$. Then the total discounted return random vector is given by

$$\mathbf{R}(\pi) \equiv \sum_{t=0}^{\infty} \beta^t \mathbf{r}(\pi(t)) \circ \mathbf{X}_\pi(t) .$$

Received January 31, 1972; revised April 11, 1972.

$[\mathbf{R}(\pi)]_i$ is the return random variable given that the process starts in state i, that is that $X_\pi(0) = i$, and thus if $\mathbf{p}$ is the probability vector representing the initial state distribution, the total return random variable using policy π is $\mathbf{R}(\pi) \cdot \mathbf{p}$. The nth moment of return we denote by $\mathbf{M}_n(\pi)$ defined by

$$\mathbf{M}_n(\pi) \equiv E[(\mathbf{R}(\pi))^n] \qquad (n = 1, 2, \cdots)$$
$$\mathbf{M}_0(\pi) \equiv \mathbf{1}$$

where $(\mathbf{R}(\pi))^n = \mathbf{R}(\pi) \circ \mathbf{R}(\pi) \circ \cdots \circ \mathbf{R}(\pi)$. We define the vector $\mathbf{N}^m(\pi)$ by

$$\mathbf{N}^m(\pi) \equiv (-\mathbf{M}_0(\pi), \mathbf{M}_1(\pi), -\mathbf{M}_2(\pi), \cdots, (-1)^{n+1}\mathbf{M}_m(\pi))$$

and the vector $\mathbf{N}(\pi)$ we define as $\mathbf{N}^\infty(\pi)$.

A policy π^* is said to be moment optimal if $\mathbf{N}(\pi^*) \succeq \mathbf{N}(\pi)$ for all policies π, i.e., π^* lexicographically maximizes the sequence of the moments of return vectors with alternating signs. In addition a policy π^* is said to be (m) moment optimal if $\mathbf{N}^m(\pi^*) \succeq \mathbf{N}^m(\pi)$ for all policies π. It is clear that a policy is moment optimal if and only if it is (m) moment optimal for all m. We also define sets F^m to be the set of all stationary (m) moment optimal policies:

$$F^m \equiv \{\mathbf{f} : \mathbf{f} \in F \text{ and } f^\infty \text{ is } (m) \text{ moment optimal}\} .$$

In this paper we develop two results. We first establish a relationship between the moments of return for stationary policies and the interest rate. This relationship is a Laurent expansion. The second result is that there exists a stationary policy which is moment optimal for all sufficiently small interest rates. The proofs of these results follow the treatment given in Veinott [6] for the Laurent expansion of the expected return and the treatment of moment optimality given in Jaquette [3].

2. Laurent expansion for the moments of return. We shall only treat Markov processes with a discrete time parameter in this paper. For an arbitrary stationary policy we let P denote the transition probability matrix and $\mathbf{r}$ the one period return vector. We assume discounting with a discount factor β, which is related to the interest rate ρ by $\beta = (1 + \rho)^{-1}$.

It is quite simple to verify the following recursion for the nth moment of return:

$$\mathbf{M}_n = \sum_{i=0}^{n} \beta^i \binom{n}{i} (\mathbf{r})^{n-i} \circ P\mathbf{M}_i$$

where we take $(\mathbf{r})^0 \equiv \mathbf{1}$ and $P\mathbf{M}_0 \equiv \mathbf{1}$. This can be seen by expanding the total return $\mathbf{R}$ into $\mathbf{r} + \beta\mathbf{R} \circ \mathbf{X}$, expanding $(\mathbf{r} + \beta\mathbf{R} \circ \mathbf{X}(1))^n$, and taking the expectation. We can rewrite the above expression in closed form as long as $\beta < 1$:

$$\mathbf{M}_n = (I - \beta^n P)^{-1} \sum_{i=0}^{n-1} \beta^i \binom{n}{i} (\mathbf{r})^{n-i} \circ P\mathbf{M}_i . \tag{2.1}$$

We can also write (2.1) in the same form that it takes for $n = 1$:

$$\mathbf{M}_n = (I - \beta_n P)^{-1} \mathbf{r}_n , \tag{2.2}$$

where $\beta_n \equiv \beta^n$ and $\mathbf{r}_n \equiv \sum_{i=0}^{n-1} \beta^i \binom{n}{i} (\mathbf{r})^{n-i} \circ P\mathbf{M}_i$. Equation (2.2) is the same form as the expression for the expected return using a stationary policy; cf. Veinott [6]. We shall treat the nth moment just like the first moment.

The resolvent of a matrix Q, denoted by $R_\lambda(Q)$, is defined as

$$R_\lambda(Q) \equiv (\lambda I - Q)^{-1} .$$

By defining $Q \equiv P - I$, we can rewrite (2.2) in terms of the resolvent and the pseudo-return $\mathbf{r}_n$:

$$\mathbf{M}_n = (1 + \rho_n)R_{\rho_n}(Q)\mathbf{r}_n , \tag{2.3}$$

where $\rho_n = \beta_n^{-1} - 1 = (1 + \rho)^n - 1$.

We now make some additional definitions which will enable us to apply some well-known results to the present problem. Let

$$P^* \equiv \lim_{n\to\infty} (n + 1)^{-1} \textstyle\sum_{i=0}^{n} P^i ,$$
$$H_\lambda \equiv R_\lambda(Q)(I - P^*) , \quad \text{and}$$
$$H \equiv R_0(Q - P^*)(I - P^*) .$$

It is well known that P^* exists for a stochastic matrix P. We now state Lemma 2.1, which can be found in Veinott's paper [6] in a more general setting, without proof, and we refer the reader to that paper for the details.

LEMMA 2.1. *If $\rho \neq 0$ and $\rho < ||H||^{-1}$, then*

$$R_\rho(Q) = \rho^{-1}P^* + \textstyle\sum_{i=0}^{\infty} (-\rho)^i H^{i+1} .$$

A direct application of Lemma 2.1 to (2.2) yields

$$\mathbf{M}_n = (1 + \rho)^n[\textstyle\sum_{i=-1}^{\infty} (\rho_n)^i u_i]\mathbf{r}_n \tag{2.4}$$

where we have rewritten (2.3) with $(1 + \rho_n)$ expressed as $(1 + \rho)^n$ and expressed $R_{\rho_n}(Q)$ in the form

$$R_{\rho_n}(Q) = \textstyle\sum_{i=-1}^{\infty} (\rho_n)^i u_i , \tag{2.5}$$

where

$$u_{-1} = P^* \quad \text{and} \quad u_i = -(-H)^{i+1} \qquad (i = 0, \cdots) .$$

The condition for convergence in Lemma 2.1 applied to (2.4) and (2.5) can be written as $(1 + \rho)^n - 1 < ||H||^{-1}$, which is satisfied for small enough ρ.

We shall now show that (2.5) can be expressed as a convergent Laurent series in ρ instead of ρ_n as written. Consider first the term for $i = -1$. We have $\rho_n^{-1}P^*$. Since $\rho_n = (1 + \rho)^n - 1$, we can write this as

$$\rho_n^{-1}P^* = \rho^{-1}[\textstyle\sum_{i=0}^{n-1} \binom{n}{i+1}\rho^i]^{-1}P^* . \tag{2.6}$$

The expression $[\sum_{i=0}^{n-1} \binom{n}{i+1}\rho^i]^{-1}$ is positive for $\rho \geqq 0$, and it clearly has no poles for $\rho \geqq 0$. We can in fact, write a power series for $[\sum_{i=0}^{n-1} \binom{n}{i+1}\rho^i]^{-1}$ which converges for $\rho \in [0, 1)$:

$$[\textstyle\sum_{i=0}^{n-1} \binom{n}{i+1}\rho^i]^{-1} = \sum_{i=0}^{\infty} a_i^n \rho^i , \tag{2.7}$$

where $a_0^n = n^{-1}$, $a_i^n = -n^{-1} \sum_{j=1}^{(n-1)\wedge i} \binom{n}{j+1} a_{i-j}^n$, and $(n - 1) \wedge i \equiv$ minimum of

$\{n-1, i\}$. By using (2.7) in (2.6) we obtain a Laurent expansion which converges in (0, 1):

$$\rho_n^{-1}P^* = (\textstyle\sum_{i=-1}^{\infty} a_{i+1}^n \rho^i)P^* . \tag{2.8}$$

For the remainder of (2.5) we have terms $(\rho_n)^i$ for $i \geqq 0$. We can write these as

$$(\textstyle\sum_{j=0}^{n-1} \binom{n}{j+1}\rho^j)^i ,$$

which is still a polynomial in ρ. We can therefore simply collect terms with the same power of ρ to obtain a new power series

$$\textstyle\sum_{i=0}^{\infty} (\rho_n)^i u_i = \sum_{i=0}^{\infty} \rho^i \bar{u}_i ,$$

where the $\bar{u}_i$ are polynomials in the u's. Thus we can combine this result with (2.8) to obtain the formal result

$$R_{\rho_n}(Q) = \textstyle\sum_{i=-1}^{\infty} \rho^i C_i^n , \tag{2.9}$$

where C_i^n are matrixes; $C_{-1}^n = a_0^n P^*$ and $C_i^n = a_{i+1}^n P^* + \bar{u}_i$ $(i \geqq 0)$. The Laurent series in (2.9) converges in the region specified by Lemma 2.1: ρ small enough that $(1+\rho)^n - 1 < ||H||^{-1}$.

We now show inductively that $\mathbf{r}_n$, as defined in (2.2), and $\mathbf{M}_n$ in (2.4) have convergent Laurent expansions in ρ of the forms

$$\mathbf{r}_n = \textstyle\sum_{i=-n+1}^{\infty} \rho^i \mathbf{b}_i^n \qquad (n = 1, 2, \cdots) , \tag{2.10}$$

and

$$\mathbf{M}_n = (1+\rho)^n \textstyle\sum_{i=-n}^{\infty} \rho^i \mathbf{y}_i^n \qquad (n = 1, 2, \cdots) , \tag{2.11}$$

where the vectors $\mathbf{b}_i^n$ and $\mathbf{y}_i^n$ are fixed vectors expressible in terms of P^*, H, $\mathbf{r}$, and binomial coefficients.

For the case $n = 1$, we have the case of the expected return, and using the results of Miller and Veinott [5] we have

$$\mathbf{r}_1 = \mathbf{r} \qquad \text{and} \qquad \mathbf{M}_1 = (1+\rho) \textstyle\sum_{i=-1}^{\infty} \rho^i u_i \mathbf{r} ,$$

so we can take $\mathbf{b}_0^1 = \mathbf{r}$, $\mathbf{b}_i^1 = \mathbf{0}$ $(i \neq 0)$ and $\mathbf{y}_i^1 = u_i \mathbf{r}$. The expression for $\mathbf{M}_1$ converges for $\rho < ||H||^{-1}$, and the expression for $\mathbf{r}$ converges for all ρ.

Suppose now that $\mathbf{r}_n$ has the form (2.10), that $\mathbf{M}_k$ has the form (2.11) for $k \leqq n$, and that these Laurent expansions converge in some common interval $(0, \bar{\rho}_n)$. Now consider $\mathbf{r}_{n+1}$. We combine the expressions for $\mathbf{M}_i$ from (2.11) in the expression for $\mathbf{r}_{n+1}$ given in (2.2) as follows to obtain $\mathbf{r}_{n+1}$.

$$\begin{aligned}
\mathbf{r}_{n+1} &= \textstyle\sum_{i=0}^{n} \beta^i \binom{n+1}{i} (\mathbf{r})^{n+1-i} \circ P[(1+\rho)^i \sum_{j=-i}^{\infty} \rho^j \mathbf{y}_j^i] \\
&= \textstyle\sum_{i=0}^{n} \binom{n+1}{i} (\mathbf{r})^{n+1-i} \circ P \sum_{j=-i}^{\infty} \rho^j \mathbf{y}_j^i \\
&= \textstyle\sum_{j=-n}^{\infty} \sum_{i=0}^{n \wedge (n+j)} \binom{n+1}{i+1} \rho^j (\mathbf{r})^{i+1} \circ P\mathbf{y}_j^{n-i} \\
&= \textstyle\sum_{j=-n}^{\infty} \rho^j (\sum_{i=0}^{n \wedge (n+j)} \binom{n+1}{i+1} (\mathbf{r})^{i+1} \circ P\mathbf{y}_j^{n-i}) \\
&= \textstyle\sum_{i=-n}^{\infty} \rho^i \mathbf{b}_i^{n+1} ,
\end{aligned}$$

where $\mathbf{b}_i^{n+1} \equiv \sum_{j=0}^{n \wedge (n+i)} \binom{n+1}{j+1} (\mathbf{r})^{j+1} \circ P\mathbf{y}_i^{n-j}$.

We can see that $\mathbf{r}_{n+1}$ has the form of (2.10). We can then use this expression in (2.4) to obtain

$$\begin{aligned}\mathbf{M}_{n+1} &= (1+\rho)^{n+1} \sum_{i=-1}^{\infty} \rho^i C_i^{n+1} \left(\sum_{j=-n}^{\infty} \rho^j \mathbf{b}_j^{n+1}\right) \\ &= (1+\rho)^{n+1} \sum_{i=-1}^{\infty} \sum_{j=-n}^{\infty} \rho^{i+j} C_i^{n+1} \mathbf{b}_j^{n+1} \\ &= (1+\rho)^{n+1} \sum_{i=-n-1}^{\infty} \rho^i \mathbf{y}_i^{n+1},\end{aligned}$$

where $\mathbf{y}_i^{n+1} = \sum_{j=-1}^{n+i} C_j^{n+1} \mathbf{b}_{i-j}^{n+1}$.

This has the form of (2.11). From the manner in which we calculated the expressions for $\mathbf{r}_{n+1}$ and $\mathbf{M}_{n+1}$ and the convergence of the expansions used in the induction hypothesis, it follows that there is some region of convergence, $(0, \bar{\rho}_{n+1})$, where we can take $\bar{\rho}_{n+1} \leqq \bar{\rho}_n$. This completes the induction.

We introduce the notation $\mathbf{M}_n(f, \rho)$, where $\mathbf{M}_n(f, \rho) \equiv \mathbf{M}_n(f^\infty)$. We append the ρ merely to make explicit the dependence of the moments of return on the interest rate. We now state the following result, which we have already proved.

THEOREM 1. *Choose any stationary policy f^∞. The* nth *moment of return using f^∞ has the Laurent expansion*

$$\mathbf{M}_n(f, \rho) = (1+\rho)^n \sum_{i=-n}^{\infty} \rho^i \mathbf{y}_i^n(f), \qquad n \geqq 1, \tag{2.12}$$

which converges in some interval $(0, \bar{\rho}_n)$, where $\mathbf{y}_i^n(f)$ are vectors depending only on i, n, and $\mathbf{f}$.

The calculation of the $\mathbf{y}_i^n(f)$ is complicated. We have given only an indication of the steps needed to perform the calculation of the $\mathbf{y}_i^n(f)$ in the proof of Theorem 1, but as we shall not need the $\mathbf{y}_i^n(f)$ in explicit form, we shall not pursue the subject further.

3. Moment optimal policies for small interest rates. We use Theorem 1 to show that there is a single stationary policy which is moment optimal for all sufficiently small interest rates. We begin by stating the following lemma.

LEMMA 3.1. *Let $h(x) = \sum_{i=-n}^{\infty} a_i x^i$ and suppose $h(x)$ converges in the region $(0, x_1)$. Then there exists an $x_0 > 0$ such that either* (a), (b) *or* (c) *holds.*

(a) $h(x) > 0$, *for all* $x \in (0, x_0)$

(b) $h(x) < 0$, *for all* $x \in (0, x_0)$

(c) $h(x) \equiv 0$, *for all* x.

PROOF. Suppose the first nonzero coefficient among the a_i is a_{-k} for some $k > 0$. Then for sufficiently small values of x, the term $a_{-k} x^{-k}$ dominates all other terms in the Laurent expansion. Since $h(x)$ is continuous on $(0, x_1)$, and $a_{-k} \neq 0$, the function $h(x)$ must have the same sign as a_{-k} in some region of x near zero. It then follows immediately that either (a) or (b) holds.

Suppose instead that $a_{-n} = \cdots = a_{-2} = a_{-1} = 0$. Then $h(x)$ has no poles at zero and is an ordinary power series. The conclusion that either (a), (b), or (c) holds in this case is easily reached under the assumptions and is given in detail in Jaquette ([2] Lemma 9.1).

We now define some notation which will simplify the discussion. For a fixed interest rate we have previously defined F^m as the set of action vectors generating stationary (m) moment optimal policies. The F^m clearly depend on the interest rate, and we now make this dependence explicit by writing $F^m(\rho)$. We then define new sets of action vectors G^m and associated interest rates ρ_m as follows:

$$(3.1)\qquad \begin{aligned} &G^m \equiv \{\mathbf{f} \in F\colon \text{ there exists a } \rho_m > 0 \text{ such that } \mathbf{f} \in F^m(\rho) \text{ all } \rho \in (0, \rho_m)\}, \\ &G_0 \equiv F \qquad\qquad \text{and} \\ &\rho_0 \equiv +\infty . \end{aligned}$$

We shall always associate the largest such ρ_m with G^m. The sets G^m are the sets of action vectors generating stationary policies which are (m) moment optimal for all interest rates in some region $(0, \rho_m)$. It follows, by definition, that if $\mathbf{f}$ and $\mathbf{g}$ are in G^m, they have the same moments of return for all kth moments, $k \leqq m$:

$$(3.2)\qquad \mathbf{M}_k(f, \rho) = \mathbf{M}_k(g, \rho) \quad \text{all } \mathbf{f}, \mathbf{g} \in G^m, \text{ all } k \leqq m, \text{ and all } \rho \in (0, \rho_m).$$

We now show that the sets G^m are in fact nonempty and that the definition (3.1) yields sets of some interest.

LEMMA 3.2. *G^m is nonempty.*

PROOF. We shall prove this lemma inductively by showing that $G^m = F^m(\rho)$ for all $\rho \in (0, \rho_m)$. For $m = 0$ the lemma holds trivially by definition (3.1) since $F^0(\rho) = F$.

Now assume that $G^k \neq \emptyset$ and $G^k = F^k(\rho)$ for all $\rho \in (0, \rho_k)$. For any $\mathbf{f}$ and $\mathbf{g}$ in G^k and for any ρ' in $(0, \rho_k]$ we have

$$\mathbf{M}_i(f, \rho) = \mathbf{M}_i(g, \rho) \qquad \text{for all } i \leqq k \text{ and } \rho \in (0, \rho') .$$

The relationship $G^{k+1} \subseteqq G^k$ must hold since $F^{k+1}(\rho) \subseteqq F^k(\rho)$ for all ρ, and $F^k(\rho) = G^k$ on $(0, \rho_k)$. Thus to find an element of G^{k+1} we seek a $\rho_{k+1} \in (0, \rho_k]$ and an $\mathbf{f} \in G^k$ such that $\mathbf{M}_{k+1}(f, \rho) \geqq (-1)^k \mathbf{M}_{k+1}(g, \rho)$ for all $\mathbf{g}$ in G^k and $\rho \in (0, \rho_{k+1})$. We do this by considering the difference $\mathbf{M}_{k+1}(f, \rho) - \mathbf{M}_{k+1}(g, \rho)$, which by Theorem 1, can be written as

$$(3.3)\qquad \mathbf{M}_{k+1}(f, \rho) - \mathbf{M}_{k+1}(g, \rho) = (1 + \rho)^{k+1} \textstyle\sum_{i=-k-1}^{\infty} \rho^i \mathbf{a}_i ,$$

where $\mathbf{a}_i = \mathbf{y}_i^{k+1}(f) - \mathbf{y}_i^{k+1}(g)$. Appealing to Lemma 3.1 we conclude that for each component of the function in (3.3) there is an open interval of ρ bounded below by zero on which the function does not change sign. Thus since the state space is finite, there exists a common open interval on which the vector function (3.3) does not change sign for the given pair $\mathbf{f}$ and $\mathbf{g}$ in G^k. Since there are only a finite number of pairs of $\mathbf{f}$ and $\mathbf{g}$, there is again a common interval $(0, \rho')$ with $\rho' > 0$ and $\rho' < \rho_k$ such that for any $\mathbf{f}$ and $\mathbf{g}$ in G^k and all $\rho \in (0, \rho')$, either $\mathbf{M}_{k+1}(f, \rho) \geqq \mathbf{M}_{k+1}(g, \rho)$ or $\mathbf{M}_{k+1}(f, \rho) \leqq \mathbf{M}_{k+1}(g, \rho)$.

Choose any $\mathbf{f}$ from the nonempty set $F^{k+1}(\rho')$. For this $\mathbf{f}$ we have $\mathbf{M}_i(f, \rho) = \mathbf{M}_i(g, \rho)$ for all $\mathbf{g} \in G^k$ and $\rho \in (0, \rho')$ and

$$\mathbf{M}_{k+1}(f, \rho') \geqq (-1)^k \mathbf{M}_{k+1}(g, \rho') \qquad \text{for all } \mathbf{g} \in G^k .$$

If $\mathbf{M}_{k+1}(f, \rho') \geqq (-1)^k \mathbf{M}_{k+1}(g, \rho')$, then we must have $\mathbf{M}_{k+1}(f, \rho) \geqq (-1)^k \mathbf{M}_{k+1}(g, \rho)$ for all $\rho \in (0, \rho')$. With this fact and knowing that $F^k(\rho) = G^k$ is constant on $(0, \rho')$, we conclude that $\mathbf{f} \in F^{k+1}(\rho)$ for all $\rho \in (0, \rho')$. We also conclude that $F^{k+1}(\rho)$ is constant on $(0, \rho')$. Taking $\rho_{k+1} = \rho'$, we see that $G^{k+1} = F^{k+1}(\rho)$ for all $\rho \in (0, \rho_{k+1})$. Since $F^{k+1}(\rho)$ is nonempty, so is G^{k+1}. This completes the induction step and completes the proof of the lemma.

From Lemma 3.2 we see that there are stationary policies which are (m) moment optimal for all interest rates in the interval $(0, \rho_m)$ for some $\rho_m > 0$. There may, of course, be nonstationary (m) moment optimal policies, but we shall not consider them in our attempts to find policies which are moment optimal for small interest rates.

We now give the following lemma which characterizes the behavior of the sets G^m for large m. The main result of this section, Theorem 2, follows directly from this lemma.

LEMMA 3.3. *There exists a finite number n such that* $G^n = G^{n+1} = \cdots = G^\infty \neq \emptyset$ *and we can choose the associated interest rate limits such that* $\rho_n = \rho_{n+1} = \cdots = \rho_\infty > 0$.

PROOF. This result that G^m is constant for large enough m is not difficult to establish, and a detailed discussion of the arguments needed to prove this lemma is given in Jaquette ([2] Lemma 5.3). We only summarize the arguments here. Taking $G^0 = F$, we have a finite set. As m increases, elements are eliminated from G^m and G^m is reduced. Since G^0 is finite, there must be a last number m such that $G^{m+1} \subset G^m$ $(G^{m+1} \neq G^m)$. Thus taking n to be this index plus one, we must have $G^n = G^{n+1} = \cdots = G^\infty$. Since $G^m \neq \emptyset$ for all m, we also must conclude that $G^\infty \neq \emptyset$.

We now consider the choice of the interest rate limits ρ_k for $k \geqq n$. Choose any $k \geqq n$. Since $G^k = G^{k+1}$, we have for all $\mathbf{f}$ and $\mathbf{g}$ in G^k and all ρ in $(0, \rho_{k+1})$

$$\mathbf{M}_{k+1}(f, \rho) = \mathbf{M}_{k+1}(g, \rho) .$$

This implies that for all $\rho > 0$ and all $\mathbf{f}$ and $\mathbf{g}$ in G^k $\mathbf{M}_{k+1}(f, \rho) = \mathbf{M}_{k+1}(g, \rho)$. Since $F^k(\rho) = G^k$ for in $(0, \rho_k)$, since all elements of $F^k(\rho)$ have identical $(k+1)$st moments for all ρ in $(0, \rho_k)$, and since $F^{k+1}(\rho) \subseteqq F^k(\rho)$ for any ρ, we conclude that $F^{k+1}(\rho)$ is constant on $(0, \rho_k)$. This means that we could have chosen $\rho_{k+1} = \rho_k$ and still preserved $F^{k+1}(\rho) = G^{k+1}$ for all ρ in $(0, \rho_{k+1})$. Since this holds true for all $k \geqq n$, we can choose the ρ_k associated with G^k such that $\rho_n = \rho_{n+1} = \cdots = \rho_\infty$. Since $\rho_n > 0$, we have $\rho_\infty > 0$. This complete the proof of the lemma.

The main result of this section, Theorem 2, is a direct corollary of Lemma 3.3. It follows since $G^\infty \neq \emptyset$ and $\rho^\infty > 0$:

THEOREM 2. *There is a stationary policy which is moment optimal for all interest rates in some interval* $(0, \rho)$ *where* $\rho > 0$.

We remark that G^∞ gives all the stationary policies which are moment optimal for small interest rates. The method of proof of Lemma 3.2 also indicates an algorithm to construct the sets G^m based on the vectors $\mathbf{y}_i^k(f)$. An alternative algorithm can be devised following the algorithms developed in Jaquette [3] for the discrete time case and in Jaquette [4] for the continuous time case; the only changes needed to those algorithms is the explicit accounting of the dependence on the interest rate, but it is no problem to account for this and to compute ρ_m and G^m directly. Since we cannot give a simple closed expression for the $\mathbf{y}_i^k(f)$, the latter algorithm is more practical.

REFERENCES

[1] BLACKWELL, D. (1962). Discrete dynamic programming. *Ann. Math. Statist.* **33** 719–726.
[2] JAQUETTE, S. C. (1971). Markov decision processes with a new optimality criterion. Technical Report No. 15, Department of Operations Research, Stanford Univ.
[3] JAQUETTE, S. C. (1973a). Markov decision processes with a new optimality criterion: discrete time. *Ann. Statist.* **1** 000–000.
[4] JAQUETTE, S. C. (1973b). Markov decision processes with a new optimality criterion: continuous time. *Ann. Statist.* **1** 000–000.
[5] MILLER, B. L. and VEINOTT, A. F. (1969). Discrete dynamic programming with a small interest rate. *Ann. Math. Statist.* **40** 366–370.
[6] VEINOTT, A. F. (1969). Discrete dynamic programming with sensitive discount optimality criteria. *Ann. Math. Statist.* **40** 1635–1660.

SCIENCE APPLICATIONS, INC.
1020 CORPORATION WAY
PALO ALTO, CALIFORNIA 94303