SIAM J. CONTROL AND OPTIMIZATION
Vol. 15, No. 3, May 1977

MONOTONE MAPPINGS WITH APPLICATION IN DYNAMIC PROGRAMMING*

DIMITRI P. BERTSEKAS†

Abstract. The structure of many sequential optimization problems over a finite or infinite horizon can be summarized in the mapping that defines the related dynamic programming algorithm. In this paper we take as a starting point this mapping and obtain results that are applicable to a broad class of problems. This approach has also been taken earlier by Denardo under contraction assumptions. The analysis here is carried out without contraction assumptions and thus the results obtained can be applied, for example, to the positive and negative dynamic programming models of Blackwell and Strauch. Most of the existing results for these models are generalized and several new results are obtained relating mostly to the convergence of the dynamic programming algorithm and the existence of optimal stationary policies.

1. Introduction. It is well known that dynamic programming (D.P. for short) is the principal method for analysis of sequential optimization problems. It is also known that it is possible to describe each iteration of a D.P. algorithm by means of a certain mapping which maps the set of extended real-valued functions defined on the state space into itself. In problems with a finite, say N, number of stages, after N successive applications of this mapping (i.e. after N steps of the D.P. algorithm) one obtains the optimal value function of the problem. In problems with an infinite number of stages one hopes that the sequence of functions generated by successive application of the D.P. iteration converges in some sense to the optimal value function of the problem. Furthermore it is possible to define the optimization problem itself in terms of the underlying mapping.

To illustrate this viewpoint let us consider formally the deterministic optimal control problem of finding a control law, i.e. a finite sequence of control functions, $\pi = \{\mu_0, \mu_1, \cdots, \mu_{N-1}\}$ which minimizes

$$J_\pi(x_0) = \sum_{k=0}^{N-1} g[x_k, \mu_k(x_k)] \tag{1}$$

subject to the system equation constraint

$$x_{k+1} = f[x_k, \mu_k(x_k)], \qquad k = 0, 1, \cdots, N-1. \tag{2}$$

The states x_k belong to a state space S and the controls $\mu_k(x_k)$ are elements of a control space C. The initial state x_0 is known and f, g are given functions. The D.P. algorithm for this problem is given by

$$J_0(x) = 0, \tag{3}$$

$$J_{k+1}(x) = \inf_u \{g(x, u) + J_k[f(x, u)]\}, \qquad k = 0, \cdots, N-1, \tag{4}$$

and the optimal value of the problem $J^*(x_0)$ is obtained at the Nth step of the D.P. algorithm

$$J^*(x_0) = \inf_\pi J_\pi(x_0) = J_N(x_0).$$

---

* Received by the editors December 3, 1974, and in revised form July 1, 1976.

† Department of Electrical Engineering and the Coordinated Science Laboratory, University of Illinois, Urbana, Illinois 61801. This work was carried out at the Coordinated Science Laboratory and was supported by the National Science Foundation under Grant ENG 74-19332.

One may also obtain the value $J_\pi(x_0)$ corresponding to any $\pi = \{\mu_0, \mu_1, \cdots, \mu_{N-1}\}$ by means of the algorithm

$$J_{0,\pi}(x) = 0, \tag{5}$$

$$J_{k+1,\pi}(x) = g[x, \mu_k(x)] + J_{k,\pi}[f(x, \mu_k(x))], \quad k = 0, \cdots, N-1, \tag{6}$$

$$J_\pi(x_0) = J_{N,\pi}(x_0).$$

Now it is possible to formulate the problem above as well as to describe the D.P. algorithm (3), (4) by means of the mapping H given by

$$H(x, u, J) = g(x, u) + J[f(x, u)]. \tag{7}$$

Let us define the mapping T by

$$T(J)(x) = \inf_u H(x, u, J), \tag{8}$$

and for any function $\mu: S \to C$ the mapping T_μ by

$$T_\mu(J)(x) = H[x, \mu(x), J]. \tag{9}$$

Both T and T_μ map the set of real-valued (or perhaps extended real-valued) functions on S into itself. Then in view of (5), (6) we may write the cost functional $J_\pi(x_0)$ of (1) as

$$J_\pi(x_0) = (T_{\mu_0} T_{\mu_1} \cdots T_{\mu_{N-1}})(J_0)(x_0), \tag{10}$$

where J_0 is the zero function on $S(J_0(x) = 0, \forall x \in S)$, and $(T_{\mu_0} T_{\mu_1} \cdots T_{\mu_{N-1}})$ denotes the composition of the mappings $T_{\mu_0}, T_{\mu_1}, \cdots, T_{\mu_{N-1}}$. Similarly the D.P. algorithm (3), (4) may be described by

$$J_{k+1}(x) = T(J_k)(x), \qquad k = 0, 1, \cdots, N-1, \tag{11}$$

and we have

$$J^*(x_0) = \inf_\pi J_\pi(x_0) = T^N(J_0)(x_0), \tag{12}$$

where T^N is the composition of T with itself N times.

One may consider also an infinite horizon version of the deterministic problem above whereby we seek a sequence $\pi = \{\mu_0, \mu_1, \cdots\}$ that minimizes

$$J_\pi(x_0) = \lim_{N\to\infty} \sum_{k=0}^{N-1} g[x_k, \mu_k(x_k)] \tag{13}$$

subject to the system equation constraint (2). In this case one needs, of course, to make assumptions which ensure that the limit in (13) is well defined for each π and x_0. A primary question of interest is whether the optimal value function J^* satisfies Bellman's functional equation

$$J^*(x) = \inf_u \{g(x, u) + J^*[f(x, u)]\}$$

or equivalently whether

$$J^*(x) = T(J^*)(x) \quad \forall x \in S,$$

and J^* is a fixed point of the mapping T. This question has been answered in the affirmative for broad classes of problems [1], [3], [6], [11]. Other questions relate to the existence and characterization of optimal policies. It is also of both computational and analytical interest to know whether

$$J^*(x) = \lim_{N\to\infty} T^N(J_0)(x) \quad \forall x \in S. \tag{14}$$

When (14) holds, the D.P. algorithm yields in the limit the optimal value function of the problem. While (14) holds in discounted and positive dynamic programming models [1], [11], it has been proved only under restrictive finiteness assumptions for the negative model of Strauch (see [11, Thm. 9.1]). In fact for such models (14) may easily fail to hold as the following example shows:

Example. Let $S = [0, \infty)$, $C = (0, \infty)$ be the state and control spaces respectively. Let the system equation be

$$x_{k+1} = 2x_k + u_k, \qquad k = 0, 1, \cdots,$$

and let the cost per stage be defined by

$$g(x, u) = x.$$

Then it can be easily verified that

$$J^*(x) = \inf_\pi J_\pi(x) = +\infty \quad \forall x \in S$$

while

$$T^N(J_0)(0) = 0 \quad \forall N = 1, 2, \cdots.$$

The deterministic optimization problem described above is representative of a plethora of sequential optimization problems of practical interest which may be formulated in terms of mappings similar to the mapping H of (7). A class of such problems has been formulated and analyzed by Denardo [4]. His framework however is restricted by contraction and boundedness assumptions which preclude the use of his results in many types of problems including the positive and negative models of Blackwell [3] and Strauch [11]. The purpose of this paper is to provide a broader framework than the one of Denardo which includes in particular positive and negative models. Questions such as those described above for the deterministic problem are analyzed in this broader setting. Most of the existing results on positive and negative models are generalized. Some entirely new results are also obtained, most notably a necessary and sufficient condition for convergence of the D.P. algorithm (Proposition 11). This result yields in turn powerful a priori verifiable sufficient conditions for convergence of the D.P. algorithm as well as for existence of an optimal stationary policy (Proposition 12). Since under our assumptions we cannot rely on contraction properties, the line of analysis is entirely different from the one of Denardo and utilizes primarily the monotonicity of the mappings involved.

2. Notation and assumptions. The following notational conventions will be used throughout the paper:

1. S and C are two given nonempty sets referred to as the *state space* and *control space* respectively.

2. For each $x \in S$ there is given a nonempty subset $U(x)$ of C referred to as the *control constraint set at* x.

3. We denote by M the set of all functions $\mu: S \to C$ such that $\mu(x) \in U(x)$ for all $x \in S$. We denote by Π the set of all sequences $\pi = \{\mu_0, \mu_1, \cdots\}$ such that $\mu_k \in M$ for all k. Elements of Π are referred to as *policies*. Elements of Π of the form $\pi = \{\mu, \mu, \cdots\}$ where $\mu \in M$ are referred to as *stationary policies*.

4. We denote

F: The set of all extended real valued functions $J: S \to [-\infty, \infty]$.

B: The Banach space of all bounded real-valued functions $J: S \to (-\infty, \infty)$ with the sup norm $\|\cdot\|$ defined by

$$\|J\| = \sup_{x \in S} |J(x)| \quad \forall J \in B.$$

The unit function in F will be denoted e $[e(x) = 1, \forall x \in S]$.

5. For all $J, J' \in F$ we write

$$J = J' \quad \text{if } J(x) = J'(x) \quad \forall x \in S,$$

$$J \leqq J' \quad \text{if } J(x) \leqq J'(x) \quad \forall x \in S.$$

6. For any sequence $\{J_k\}$ with $J_k \in F$ for all k we denote by $\lim_{k\to\infty} J_k$ the pointwise limit of $\{J_k\}$ (assuming it is well defined as an extended real-valued function), and by $\liminf_{k\to\infty} J_k$ the pointwise limit inferior of $\{J_k\}$. Throughout the paper the convergence analysis is carried out within the set of extended real numbers, i.e. $+\infty$ or $-\infty$ are allowed as limits of sequences of extended real numbers. For any collection $\{J_a | a \in A\} \subset F$ parameterized by the elements of a set A we denote $\inf_{\alpha \in A} J_a$ the function taking value $\inf_{\alpha \in A} J_\alpha(x)$ at each $x \in S$.

7. We are given a mapping $H: S \times C \times F \to [-\infty, +\infty]$ and we define for each $\mu \in M$ the mapping $T_\mu: F \to F$ by

$$T_\mu(J)(x) = H[x, \mu(x), J] \quad \forall x \in S. \tag{15}$$

We define also the mapping $T: F \to F$ by

$$T(J)(x) = \inf_{u \in U(x)} H(x, u, J) \quad \forall x \in S. \tag{16}$$

We denote by T^k, $k = 1, 2, \cdots$, the composition of T with itself k times. For convenience we also define $T^0(J) = J$ for all $J \in F$. For any $\pi = \{\mu_0, \mu_1, \cdots\} \in \Pi$ we denote by $(T_{\mu_0} T_{\mu_1} \cdots T_{\mu_k})$ the composition of the mappings $T_{\mu_0}, \cdots, T_{\mu_k}$, $k = 0, 1, \cdots$.

The following assumption will be in effect throughout the paper.

Monotonicity assumption. There holds for every $x \in S$, $u \in U(x)$, $J, J' \in F$,

$$H(x, u, J) \leqq H(x, u, J') \quad \text{if } J \leqq J'.$$

Notice that the monotonicity assumption implies the following relations:

$$J \leqq J' \Rightarrow T(J) \leqq T(J') \qquad \forall J, J' \in F,$$

$$J \leqq J' \Rightarrow T_\mu(J) \leqq T_\mu(J') \quad \forall J, J' \in F, \quad \mu \in M.$$

We shall make frequent use of these relations.

3. Problem formulation. We are given a function $\bar{J} \in F$ satisfying

$$\bar{J}(x) > -\infty \quad \forall x \in S$$

and we consider for every $\pi = \{\mu_0, \mu_1, \cdots\} \in \Pi$ the function $J_\pi \in F$ defined by

(17) $$J_\pi(x) = \lim_{N \to \infty} (T_{\mu_0} T_{\mu_1} \cdots T_{\mu_{N-1}})(\bar{J})(x) \quad \forall x \in S.$$

We refer to J_π as the *value function of* π. Under the assumptions that we will introduce shortly J_π is well defined. Throughout the paper we will be concerned with the optimization problem

(18) $$\text{minimize } J_\pi(x) \quad \text{subject to } \pi \in \Pi.$$

The optimal value of this problem for a fixed $x \in S$ is denoted by $J^*(x)$,

(19) $$J^*(x) = \inf_{\pi \in \Pi} J_\pi(x).$$

We refer to the function $J^* \in F$ as the *optimal value function.* We say that a policy $\pi^* \in \Pi$ is *optimal at* $x \in S$ if $J_{\pi^*}(x) = J^*(x)$ and we say that a policy $\pi^* \in \Pi$ is *optimal* if $J_{\pi^*} = J^*$. For any stationary policy $\pi = (\mu, \mu, \cdots\} \in \Pi$ we write $J_\pi = J_\mu$. Thus a stationary policy $\pi^* = \{\mu^*, \mu^*, \cdots\}$ is optimal if $J^* = J_{\mu^*}$.

For every result to be shown *one of the following three assumptions will be in effect.*

Assumption C (Contraction assumption). The functions $\bar{J}$, $T(J)$, and $T_\mu(J)$ belong to B for all $\mu \in M$ and $J \in B$, and for every $\pi = \{\mu_0, \mu_1, \cdots\} \in \Pi$ the limit

$$\lim_{N \to \infty} (T_{\mu_0} T_{\mu_1} \cdots T_{\mu_{N-1}})(\bar{J})(x)$$

exists and is a real number for each $x \in S$. Furthermore there exist a positive integer m, and scalars ρ, α with $0 < \rho < 1$, $0 < \alpha$ such that for all $J, J' \in B$ there holds

(20) $$\|T_\mu(J) - T_\mu(J')\| \leqq \alpha \|J - J'\| \quad \forall \mu \in M,$$

(21) $$\|(T_{\mu_0}, T_{\mu_1} \cdots T_{\mu_{m-1}})(J) - (T_{\mu_0} T_{\mu_1} \cdots T_{\mu_{m-1}})(J')\| \leqq \rho \|J - J'\| \quad \forall \mu_0, \cdots, \mu_{m-1} \in M.$$

Assumption I (Uniform increase assumption). There holds

(22) $$\bar{J}(x) \leqq H(x, u, \bar{J}) \quad \forall x \in S, \quad u \in U(x).$$

Assumption D (Uniform decrease assumption). There holds

(23) $$\bar{J}(x) \geqq H(x, u, \bar{J}) \quad \forall x \in S, \quad u \in U(x).$$

It is easy to see that under each of these assumptions the limit in (17) is well defined as a real number or $\pm\infty$. Indeed in the case of Assumption I we have using (22)

$$(24) \qquad \bar{J} \leqq T_{\mu_0}(\bar{J}) \leqq (T_{\mu_0}T_{\mu_1})(\bar{J}) \leqq \cdots \leqq (T_{\mu_0}T_{\mu_1}\cdots T_{\mu_{N-1}})(\bar{J}) \leqq \cdots$$

while in the case of Assumption D we have using (23)

$$(25) \qquad \bar{J} \geqq T_{\mu_0}(\bar{J}) \geqq (T_{\mu_0}T_{\mu_1})(\bar{J}) \geqq \cdots \geqq (T_{\mu_0}T_{\mu_1}\cdots T_{\mu_{N-1}})(\bar{J}) \geqq \cdots$$

In both cases the limit in (17) clearly exists for each $x \in S$.

A large number of sequential optimization problems which are of interest in practice may be viewed as special cases of the abstract problem formulated above. We provide below some examples. Several other examples can be found in the author's textbook [1], and in the paper by Denardo [4] who considered a somewhat different problem under assumptions similar to Assumption C.

1. *Deterministic optimal control problems with additive cost functional.*

$$(26) \qquad \text{minimize} \lim_{N\to\infty} \sum_{k=0}^{N-1} \alpha^k g[x_k, \mu_k(x_k)]$$

subject to

$$x_{k+1} = f[x_k, \mu_k(x_k)], \quad \mu_k \in M, \quad k = 0, 1, \cdots.$$

If we define

$$(27) \qquad H(x, u, J) = g(x, u) + \alpha J[f(x, u)],$$

then problem (26) is equivalent to our abstract problem (18) for $\bar{J}(x) = 0$, $\forall x \in S$. Assumption C is satisfied if $0 < \alpha < 1$ and g is uniformly bounded, i.e., there exists a scalar $b > 0$ such that

$$(28) \qquad |g(x, u)| \leqq b \quad \forall x \in S, \quad u \in U(x).$$

This case corresponds to a discounted problem and is examined in [1, §§ 6.1–6.3]. Assumption I is satisfied if $0 < \alpha$ and

$$(29) \qquad g(x, u) \geqq 0 \quad \forall x \in S, \quad u \in U(x)$$

while Assumption D is satisfied if $0 < \alpha$ and

$$(30) \qquad g(x, u) \leqq 0 \quad \forall x \in S, \quad u \in U(x).$$

These cases are covered in [1, §§ 6.4, 7.1]. If g is extended real valued some care must be exercised in the definition of the mapping H in (27) so that the forbidden sum $(+\infty, -\infty)$ does not arise. This can be done by defining under Assumption I [c.f. (29)] $H(x, u, J) \equiv -\infty$ if $J(x) = -\infty$ for some $x \in S$, and by defining under Assumption D [c.f. (30)] $H(x, u, J) = +\infty$ if $J(x) = +\infty$ for some $x \in S$. We mention that state constraints of the form $x_k \in X$, $\forall k = 0, 1, \cdots$, can be incorporated under I in the cost functional by defining $g(x, u) = +\infty$ whenever $x \notin X$. Note that the deterministic versions of Blackwell's positive D.P. model [3] and Strauch's negative D.P. model [11] are covered under Assumption D and Assumption I respectively.

Deterministic optimal control problems with nonstationary cost per stage and system equation (including finite horizon problems) may be reformulated into the form of problem (26) (see [1, § 6.7]). A generalization of problem (26) is obtained if the scalar α is replaced by a function $\alpha(x, u)$ in (27) and the discount factor depends on the state x and the control u. Then Assumption C is satisfied if the assumption $0<\alpha<1$ is replaced by

$$0 \leqq \inf \{\alpha(x, u) | x \in S, u \in U(x)\} \leqq \sup \{\alpha(x, u) | x \in S, u \in U(x)\} < 1$$

and (28) holds. If $0 \leqq \alpha(x, u)$ for all $x \in S$, $u \in U(x)$, then Assumption I or D is satisfied if (29) or (30) holds respectively.

2. *Stochastic optimal control with additive cost functional.* This problem is obtained from problem (18) when $\bar{J} \equiv 0$ and

$$H(x, u, J) = \underset{w}{E} \{g(x, u, w) + \alpha J[f(x, u, w)] | x, u\},$$

where w is an uncertain parameter element of a *countable* set W with given probability distribution depending on x and u. Such problems are examined in [1, Chaps. 6 and 7] and include a large variety of Markovian decision problems with countable state space. Assumption C holds if $0<\alpha<1$ and $|g(x, u, w)| \leqq b$ for some $b>0$ and all $x \in S$, $u \in U(x)$, $w \in W$. Assumptions I and D hold if $a>0$ and $g(x, u, w) \geqq 0$ or $g(x, u, w) \leqq 0$ respectively for all x, u, w. A generalized version is obtained when α is replaced by a function $\alpha(x, u, w)$ satisfying similar assumptions as the corresponding functions in the previous example. This case covers certain discounted semi-Markov decision problems.

When the set W is not countable then matters are complicated by the need to impose a measurable space structure on S, C, and W and to require that the functions $\mu \in M$ be measurable (in the works of Blackwell, Strauch, and Hinderer [3], [11], [6], S, C, and W are Borel subsets of complete separable metric spaces and μ is required to be Borel measurable). Because of these restrictions the reformulation of the stochastic control problem into the form of the abstract problem (18) is not straightforward. Recent work of S. Shreve and the author [10] has demonstrated however that the framework of this paper is applicable in its entirety as well as convenient once the stochastic control problem is converted to a deterministic control problem (such as the one of the previous example) for which the state space is the set of all probability measures on S. For a detailed treatment we refer to the thesis of Shreve [12].

3. *Minimax control problem with additive cost functional.* This problem is obtained from problem (18) when $\bar{J} \equiv 0$ and

$$H(x, u, J) = \sup_{w \in W(x,u)} \{g(x, u, w) + \alpha J[f(x, u, w)]\}.$$

Here again w is an uncertain parameter belonging to a set W, and $W(x, u)$ is a given subset of W for each $x \in S$, $u \in U(x)$. Under assumptions analogous to those of the previous two examples, Assumptions C, I, or D can be shown to hold. The problem of reachability over an infinite horizon examined by the author in [2] can be shown to be a special case of this problem.

4. *Stochastic optimal control problems with exponential cost functional.* Under similar assumptions for w as in Example 2 consider

$$H(x, u, J) = \underset{w}{E}\{J[f(x, u, w)]\, e^{g(x,u,w)}|x, u\}.$$

This problem corresponds to minimization of the exponential cost functional

$$J_\pi(x) = \lim_{N\to\infty} \underset{w_k}{E}\left\{\exp\left(\sum_{k=0}^{N-1} g[x_k, \mu_k(x_k), w_k]\right)\right\}$$

subject to the system equation $x_{k+1} = f[x_k, \mu_k(x_k), w_k]$. An example of a finite horizon version of this problem has been considered in [7]. Here we take $\bar{J}(x) = 1$, $\forall x \in S$. If $g(x, u, w) \geqq 0$ for all (x, u, w) then Assumption I is satisfied, while if $g(x, u, w) \leqq 0$ for all (x, u, w) then Assumption D is satisfied.

4. Results under Assumption C. As mentioned earlier, a variation of our problem under Assumption C has been analyzed by Denardo. We shall restrict ourselves to providing some results which yield the connection between Denardo's framework and the one considered here.

PROPOSITION 1. *Let Assumption C hold. Then:*

(a) *For every $J \in B$, $\pi \in \Pi$ and $x \in S$ there holds*

$$J_\pi(x) = \lim_{N\to\infty} (T_{\mu_0} T_{\mu_1} \cdots T_{\mu_{N-1}})(\bar{J})(x) = \lim_{N\to\infty} (T_{\mu_0} T_{\mu_1} \cdots T_{\mu_{N-1}})(J)(x).$$

(b) *The function J^* belongs to B and is the unique fixed point of T within B, i.e., $J^* = T(J^*)$ and if $J' \in B$, $J' = T(J')$, then $J' = J^*$. Furthermore if $J' \in B$ is such that $T(J') \leqq J'$ then $J^* \leqq J'$ while if $J' \leqq T(J')$ then $J' \leqq J^*$.*

(c) *For every $\mu \in M$ the function J_μ belongs to B and is the unique fixed point of T_μ within B.*

(d) *There holds*

$$\lim_{N\to\infty} \|T^N(J) - J^*\| = 0 \quad \forall J \in B,$$

$$\lim_{N\to\infty} \|T_\mu^N(J) - J_\mu\| = 0 \quad \forall J \in B, \quad \mu \in M.$$

(e) *A stationary policy $\pi^* = \{\mu^*, \mu^*, \cdots\} \in \Pi$ is optimal if and only if*

$$T_{\mu^*}(J^*) = T(J^*).$$

Equivalently π^* is optimal if and only if

$$T_{\mu^*}(J_{\mu^*}) = T(J_{\mu^*}).$$

(f) *If there exists an optimal policy, there exists an optimal stationary policy.*

(g) *For any $\varepsilon > 0$ there exists a stationary policy $\pi_\varepsilon = \{\mu_\varepsilon, \mu_\varepsilon, \cdots\}$ such that*

$$\|J^* - J_{\mu_\varepsilon}\| \leqq \varepsilon.$$

Proof. Since the proof uses similar arguments as those in [4] (see also [1, Chap. 6, Prob. 4]) it will be abbreviated.

(a) For any integer $k \geqq 0$ write $k = nm + q$ where q, n are nonnegative integers and $0 \leqq q < m$. Then for any $J, J' \in B$ using (20), (21) we obtain

$$\|(T_{\mu_0} \cdots T_{\mu_{k-1}})(J) - (T_{\mu_0} \cdots T_{\mu_{k-1}})(J')\| \leqq \rho^n \alpha^q \|J - J'\|$$

from which

$$\lim_{k\to\infty} (T_{\mu_0} \cdots T_{\mu_{k-1}})(\bar{J}) = \lim_{k\to\infty} (T_{\mu_0} \cdots T_{\mu_{k-1}})(J) \quad \forall J \in B.$$

(b), (c), (d) Relation (20) can be used to show (compare with the proof of Proposition 3) that

$$T^N(\bar{J})(x) = \inf_{\pi\in\Pi} (T_{\mu_0} \cdots T_{\mu_{N-1}})(\bar{J})(x) \quad \forall x \in S, \quad N = 1, 2, \cdots,$$

and it follows from (21) that T and T_μ, $\mu \in M$ are m-stage contraction mappings, i.e., $\|T^m(J) - T^m(J')\| \leqq \bar{\rho}\|J - J'\|$ and $\|T^m_\mu(J) - T^m_\mu(J')\| \leqq \bar{\rho}_\mu \|J - J'\|$ for some $\bar{\rho} \in (0, 1)$, $\bar{\rho}_\mu \in (0, 1)$ and all $J, J' \in B$. Hence T and T_μ have unique fixed points in B. The fixed point of T_μ is clearly J_μ and hence part (c) is proved. Let $\tilde{J}^*$ be the unique fixed point of T. We have $\tilde{J}^* = T(\tilde{J}^*)$. For any $\bar{\varepsilon} > 0$ take $\bar{\mu} \in M$ such that

$$T_{\bar{\mu}}(\tilde{J}^*) \leqq \tilde{J}^* + \bar{\varepsilon} e.$$

Using (20) it follows that $T^2_{\bar{\mu}}(\tilde{J}^*) \leqq T_{\bar{\mu}}(\tilde{J}^*) + \alpha\bar{\varepsilon} e \leqq \tilde{J}^* + (1+\alpha)\bar{\varepsilon} e$. Continuing in the same manner we obtain

$$T^m_{\bar{\mu}}(\tilde{J}^*) \leqq \tilde{J}^* + (1 + \alpha + \cdots + \alpha^{m-1})\bar{\varepsilon} e.$$

Using (21) we obtain

$$\begin{aligned} T^{2m}_{\bar{\mu}}(\tilde{J}^*) &\leqq T^m_{\bar{\mu}}(\tilde{J}^*) + \rho(1 + \alpha + \cdots + \alpha^{m-1})\bar{\varepsilon} e \\ &\leqq \tilde{J}^* + (1+\rho)(1 + \alpha + \cdots + \alpha^{m-1})\bar{\varepsilon} e. \end{aligned}$$

Proceeding similarly we obtain for all $k \geqq 1$,

$$T^{km}_{\bar{\mu}}(\tilde{J}^*) \leqq \tilde{J}^* + (1 + \rho + \cdots + \rho^{k-1})(1 + \alpha + \cdots + \alpha^{m-1})\bar{\varepsilon} e.$$

Taking the limit as $k \to \infty$ and using the fact $J_{\bar{\mu}} = \lim_{k\to\infty} T^{km}_{\bar{\mu}}(\tilde{J}^*)$ we obtain

$$J_{\bar{\mu}} \leqq \tilde{J}^* + \frac{1}{1-\rho}(1 + \alpha + \cdots + \alpha^{m-1})\bar{\varepsilon} e.$$

Taking $\bar{\varepsilon} = (1-\rho)(1 + \alpha + \cdots + \alpha^{m-1})^{-1}\varepsilon$ we obtain

$$J_{\bar{\mu}} \leqq \tilde{J}^* + \varepsilon e. \tag{31}$$

Since $J^* \leqq J_{\bar{\mu}}$ and $\varepsilon > 0$ is arbitrary we obtain $J^* \leqq \tilde{J}^*$. We also have

$$J^* = \inf_{\pi\in\Pi} \lim_{N\to\infty} (T_{\mu_0} \cdots T_{\mu_{N-1}})(\tilde{J}^*) \geqq \lim_{N\to\infty} T^N(\tilde{J}^*) = \tilde{J}^*.$$

Hence $J^* = \tilde{J}^*$ and J^* is the unique fixed point of T. Thus part (b) is proved. Part (d) follows immediately from the contraction property of T and T_μ.

(e) If π^* is optimal then $J_{\mu^*} = J^*$ and the result follows from part (b) and (c). If $T_{\mu^*}(J^*) = T(J^*)$ then $T_{\mu^*}(J^*) = J^*$ and hence $J_{\mu^*} = J^*$ by part (c). If $T_{\mu^*}(J_{\mu^*}) = T(J_{\mu^*})$ then $J_{\mu^*} = T(J_{\mu^*})$ and $J_{\mu^*} = J^*$ by part (b).

(f) Let $\pi^*=\{\mu_0^*, \mu_1^*, \cdots\}$ be an optimal policy. Then using parts (a) and (b)

$$J^* = J_{\pi^*} = \lim_{k\to\infty} (T_{\mu_0^*}\cdots T_{\mu_k^*})(\bar{J})$$

$$= \lim_{k\to\infty} (T_{\mu_0^*}\cdots T_{\mu_k^*})(J^*) \geqq \lim_{k\to\infty} (T_{\mu_0^*}T^k)(J^*) = T_{\mu_0^*}(J^*) \geqq T(J^*) = J^*.$$

Hence $T_{\mu_0^*}(J^*) = T(J^*)$ and by part (e) the stationary policy $\{\mu_0^*, \mu_0^*, \cdots\}$ is optimal.

(g) This part was proved earlier in the proof of part (b), [cf. (31)]. Q.E.D.

For additional results and computational methods the reader is referred to Denardo's paper [4] and the author's textbook [1, Chap. 6]. Notice that part (a) shows that $\bar{J}$ may be replaced by any function $J \in B$. Thus it is often possible to select $\bar{J}$ in a way that Assumption I or D is satisfied and obtain alternative proofs of parts of Proposition 1 by using the results of the next section.

5. Results under Assumptions I or D. In our analysis under Assumptions I or D we will occasionally need to assume one or more of the following continuity properties for the mapping H. Assumptions I.1 and I.2 will be used in conjunction with Assumption I, while Assumptions D.1 and D.2 will be used in conjunction with Assumption D.

Assumption I.1. If $\{\bar{J}_k\} \subset F$ is a sequence satisfying $\bar{J} \leqq \bar{J}_k \leqq \bar{J}_{k+1}$ for all k, then

$$\lim_{k\to\infty} H(x, u, \bar{J}_k) = H(x, u, \lim_{k\to\infty} \bar{J}_k) \quad \forall x \in S, \quad u \in U(x). \tag{32}$$

Assumption I.2. There exists a scalar $\alpha > 0$ such that for all scalars $r > 0$ and functions $J \in F$ with $\bar{J} \leqq J$ there holds

$$H(x, u, J) \leqq H(x, u, J+re) \leqq H(x, u, J)+\alpha r \quad \forall x \in S, \quad u \in U(x), \tag{33}$$

where e denotes the unit function $[e(x)=1, \forall x \in S]$.

Assumption D.1. If $\{\bar{J}_k\} \subset F$ is a sequence satisfying $\bar{J}_{k+1} \leqq \bar{J}_k \leqq \bar{J}$ for all k, then

$$\lim_{k\to\infty} H(x, u, \bar{J}_k) = H(x, u, \lim_{k\to\infty} \bar{J}_k) \quad \forall x \in S, \quad u \in U(x). \tag{34}$$

Assumption D.2. There exists a scalar $\alpha > 0$ such that for all scalars $r > 0$ and functions $J \in F$ with $J \leqq \bar{J}$ there holds

$$H(x, u, J) - \alpha r \leqq H(x, u, J-re) \leqq H(x, u, J) \quad \forall x \in S, \quad u \in U(x), \tag{35}$$

where e denotes the unit function $[e(x)=1, \forall x \in S]$.

Notice that both the deterministic and the stochastic optimal control problems of § 3 satisfy I.1, I.2, D.1, D.2. The minimax control problem of § 3 satisfies I.1, I.2, D.2 while additional assumptions are needed in order that D.1 is satisfied as well. The mapping of Example 4 in § 3 satisfies I.1, D.1, and D.2 while if it is assumed that $|g(x, u, w)| \leqq b$ for some scalar b and all (x, u, w), then I.2 holds as well.

Dynamic programming and the finite horizon version of the problem. It is both interesting and helpful in the analysis that follows to consider the finite horizon version of our problem which involves finding for any positive integer N

$$J_N(x) = \inf_{\pi \in \Pi} (T_{\mu_0} \cdots T_{\mu_{N-1}})(\bar{J})(x) \tag{36}$$

as well as a policy attaining the infimum above (if one exists). We refer to this problem as the *N-stage problem*. We have the following results:

PROPOSITION 2. *Let* I *and* I.2 *hold. Then* $J_N = T^N(\bar{J})$ *for all* $N = 1, 2, \cdots$.

Proof. For any $\varepsilon > 0$ let $\bar{\mu}_k \in M$, $k = 0, 1, \cdots, N-1$, be such that

$$T_{\bar{\mu}_k}[T^{N-k-1}(\bar{J})] \leqq T^{N-k}(\bar{J}) + \varepsilon e, \qquad k = 0, 1, \cdots, N-1.$$

Such functions exist since $\bar{J}(x) > -\infty$ for all $x \in S$ and $T^{N-k}(\bar{J}) \geqq \bar{J}$ by I. We have using I.2,

$$\begin{aligned}
J_N = \inf_{\pi \in \Pi} (T_{\mu_0} \cdots T_{\mu_{N-1}})(\bar{J}) &\leqq (T_{\bar{\mu}_0} \cdots T_{\bar{\mu}_{N-1}})(\bar{J}) \\
&\leqq (T_{\bar{\mu}_0} \cdots T_{\bar{\mu}_{N-2}})[T(\bar{J}) + \varepsilon e] \\
&\leqq (T_{\bar{\mu}_0} \cdots T_{\bar{\mu}_{N-3}})[(T_{\bar{\mu}_{N-2}} T)(\bar{J}) + \alpha \varepsilon e] \\
&\cdots\cdot \\
&\leqq (T_{\bar{\mu}_0} \cdots T_{\bar{\mu}_{N-2}} T)(\bar{J}) + \alpha^{N-1} \varepsilon e \\
&\cdots\cdot \\
&\leqq (T_{\bar{\mu}_0} \cdots T_{\bar{\mu}_{N-3}} T^2)(\bar{J}) + (\alpha^{N-2} + \alpha^{N-1}) \varepsilon e \\
&\cdots\cdot \\
&\leqq T^N(\bar{J}) + \left(\sum_{k=0}^{N-1} \alpha^k \right) \varepsilon e.
\end{aligned}$$

Since ε is arbitrary we obtain $J_N \leqq T^N(\bar{J})$. On the other hand we have, by the definition of T and J_N, $T^N(\bar{J}) \leqq J_N$. Hence $J_N = T^N(\bar{J})$. Q.E.D.

Proposition 2 may fail to hold in the absence of I.2 even if I.1 holds as the following counterexample shows.

Counterexample 1. Take $S = \{0\}$, $C = U(0) = (0, 1]$, $\bar{J}(0) = 0$, $H(0, u, J) = 1$ if $J(0) > 0$, $H(0, u, J) = u$ if $J(0) \leqq 0$. Then $(T_{\mu_0} \cdots T_{\mu_{N-1}})(\bar{J})(0) = 1$ for every $\pi \in \Pi$ and $N \geqq 2$ and hence $J_N(0) = 1$ for $N \geqq 2$. On the other hand we have $T^N(\bar{J})(0) = 0$ for all N. Here I and I.1 are satisfied but I.2 is violated.

PROPOSITION 3. *Let* D *hold. Assume that either* D.1 *holds or else* D.2 *holds and* $T^N(\bar{J})(x) > -\infty$ *for all* $x \in S$. *Then* $J_N = T^N(\bar{J})$.

Proof. Let D.1 hold. For each $k = 0, 1, \cdots, N-1$ consider a sequence $\{\mu_k^i\} \subset M$ such that

$$\lim_{i \to \infty} T_{\mu_k^i}[T^{N-k-1}(\bar{J})] = T^{N-k}(\bar{J}), \qquad k = 0, \cdots, N-1.$$

$$T_{\mu_k^i}[T^{N-k-1}(\bar{J})] \geqq T_{\mu_k^{i+1}}[T^{N-k-1}(\bar{J})],$$
$$k = 0, \cdots, N-1, \quad i = 0, 1, \cdots.$$

We have by using D.1,

$$\begin{aligned} J_N &\leqq \lim_{\substack{i_0\to\infty\\ \cdots\\ i_{N-1}\to\infty}} (T_{\mu_0^{i_0}} \cdots T_{\mu_{N-1}^{i_{N-1}}})(\bar{J}) \\ &= \lim_{\substack{i_0\to\infty\\ \cdots\\ i_{N-2}\to\infty}} (T_{\mu_0^{i_0}} \cdots T_{\mu_{N-2}^{i_{N-2}}})[\lim_{i_{N-1}\to\infty} T_{\mu_{N-1}^{i_{N-1}}}(\bar{J})] \\ &= \lim_{\substack{i_0\to\infty\\ \cdots\\ i_{N-2}\to\infty}} (T_{\mu_0^{i_0}} \cdots T_{\mu_{N-2}^{i_{N-2}}})[T(\bar{J})] \\ &\cdots\cdots \\ &= T^N(\bar{J}). \end{aligned}$$

On the other hand we have clearly $T^N(\bar{J}) \leqq J_N$ and hence $J_N = T^N(\bar{J})$.

Let D.2 hold and assume $T^N(\bar{J})(x) > -\infty\ \forall x \in S$. For any $\varepsilon > 0$ let $\bar{\mu}_k \in M$, $k = 0, 1, \cdots, N-1$, be such that

$$\begin{aligned} &T_{\bar{\mu}_{N-1}}(\bar{J}) \leqq T(\bar{J}) + \varepsilon e, \\ &(T_{\bar{\mu}_{N-2}} T_{\bar{\mu}_{N-1}})(\bar{J}) \leqq T[T_{\bar{\mu}_{N-1}}(\bar{J})] + \varepsilon e, \\ &\cdots\cdots \\ &(T_{\bar{\mu}_0} \cdots T_{\bar{\mu}_{N-1}})(\bar{J}) \leqq T[(T_{\bar{\mu}_1} \cdots T_{\bar{\mu}_{N-1}})(\bar{J})] + \varepsilon e. \end{aligned}$$

The assumption $T^N(\bar{J})(x) > -\infty$, $\forall x \in S$ guarantees that such functions, $\bar{\mu}_k$ exist. We have using D.2,

$$\begin{aligned} T^N(\bar{J}) &\geqq T^{N-1}[T_{\bar{\mu}_{N-1}}(\bar{J}) - \varepsilon e] \geqq (T^{N-1} T_{\bar{\mu}_{N-1}})(\bar{J}) - \alpha^{N-1}\varepsilon e \\ &\geqq T^{N-2}[(T_{\bar{\mu}_{N-2}} T_{\bar{\mu}_{N-1}})(\bar{J}) - \varepsilon e] - \alpha^{N-1}\varepsilon e \\ &\geqq (T^{N-2} T_{\bar{\mu}_{N-2}} T_{\bar{\mu}_{N-1}})(\bar{J}) - (\alpha^{N-2} + \alpha^{N-1})\varepsilon e \\ &\cdots\cdots \\ &\geqq (T_{\bar{\mu}_0} \cdots T_{\bar{\mu}_{N-1}})(\bar{J}) - \left(\sum_{k=0}^{N-1} \alpha^k\right)\varepsilon e \\ &\geqq J_N - \left(\sum_{k=0}^{N-1} \alpha^k\right)\varepsilon e. \end{aligned}$$

Since ε is arbitrary it follows that $T^N(\bar{J}) \geqq J_N$. On the other hand we have clearly $T^N(\bar{J}) \leqq J_N$ and hence $J_N = T^N(\bar{J})$. Q.E.D.

Proposition 3 may fail to hold if its assumptions are slightly relaxed.

Counterexample 2. Take $S = \{0\}$, $C = U(0) = (-1, 0]$, $\bar{J}(0) = 0$, $H(0, u, J) = u$ if $-1 < J(0)$, $H(0, u, J) = J(0) + u$ if $J(0) \leqq -1$. Then $(T_{\mu_0} \cdots T_{\mu_{N-1}})(\bar{J})(0) = \mu_0(0)$ and $J_N(0) = -1$, while $T^N(\bar{J})(0) = -N$ for every N. Here D and the assumption $T^N(\bar{J})(0) > -\infty$ are satisfied, but D.1 and D.2 are both violated.

Counterexample 3. Take $S = \{0, 1\}$, $C = U(0) = U(1) = (-\infty, 0]$, $\bar{J}(0) = \bar{J}(1) = 0$, $H(0, u, J) = u$ if $J(1) = -\infty$, $H(0, u, J) = 0$ if $J(1) > -\infty$, and $H(1, u, J) =$

u. Then $(T_{\mu_0}\cdots T_{\mu_{N-1}})(\bar{J})(0)=0$, $(T_{\mu_0}\cdots T_{\mu_{N-1}})(\bar{J})(1)=\mu_0(1)$ for all $N\geqq 1$. Hence $J_N(0)=0$, $J_N(1)=-\infty$. On the other hand we have $T^N(\bar{J})(0)=T^N(\bar{J})(1)=-\infty$ for all $N\geqq 2$. Here D and D.2 are satisfied, but D.1 and the assumption $T^N(\bar{J})(x)>-\infty$, $\forall x\in S$ are both violated.

Characterization of the optimal value function. We now consider the question whether Bellman's equation, [i.e. $J^*=T(J^*)$] holds within our generalized setting. We first prove a preliminary result which is of independent interest.

PROPOSITION 4. *Let* I, I.1, *and* I.2 *hold. Then given any* $\varepsilon>0$ *there exists a policy* $\pi_\varepsilon\in\Pi$ *such that*

$$J^*\leqq J_{\pi_\varepsilon}\leqq J^*+\varepsilon e. \tag{37}$$

Furthermore if the scalar α *in* I.2 *satisfies* $\alpha<1$ *the policy* π_ε *can be taken stationary.*

Proof. Let $\{\varepsilon_k\}$ be a sequence such that $\varepsilon_k>0$ for all k, and

$$\sum_{k=0}^{\infty}\alpha^k\varepsilon_k=\varepsilon. \tag{38}$$

For each $x\in S$ consider a sequence of policies $\{\pi_k[x]\}\subset\Pi$ of the form

$$\pi_k[x]=\{\mu_0^k[x],\mu_1^k[x],\cdots\}$$

such that for $k=0,1,\cdots,$

$$J_{\pi_k[x]}(x)\leqq J^*(x)+\varepsilon_k\quad \forall x\in S. \tag{39}$$

Such a sequence exists since we have $J^*(x)>-\infty$ under our assumptions.

The (admittedly confusing) notation used above and later in the proof should be interpreted as follows. The policy $\pi_k[x]=\{\mu_0^k[x],\mu_1^k[x],\cdots\}$ is associated with x. Thus $\mu_i^k[x]$ denotes, for each $x\in S$ and k, a function in M, while $\mu_i^k[x](z)$ denotes the value of $\mu_i^k[x]$ at an element $z\in S$. In particular $\mu_i^k[x](x)$ denotes the value of $\mu_i^k[x]$ at x.

Consider the functions $\bar{\mu}_k\in M$ defined by

$$\bar{\mu}_k(x)=\mu_0^k[x](x)\quad \forall x\in S, \tag{40}$$

and the functions $\bar{J}_k$ defined by

$$\bar{J}_k(x)=H[x,\bar{\mu}_k(x),\lim_{i\to\infty}(T_{\mu_1^k[x]}\cdots T_{\mu_i^k[x]})(\bar{J})]\quad \forall x\in S,\quad k=0,1,\cdots. \tag{41}$$

By using (39), (40), I, and I.1 we obtain

$$\bar{J}_k(x)=\lim_{i\to\infty}(T_{\mu_0^k[x]}\cdots T_{\mu_i^k[x]})(\bar{J})(x)=J_{\pi_k[x]}(x)\leqq J^*(x)+\varepsilon_k \qquad \forall x\in S,\quad k=0,1,\cdots. \tag{42}$$

We have using (41), (42), and I.2 for all $k=1,2,\cdots$ and $x\in S$,

$$\begin{aligned}T_{\bar{\mu}_{k-1}}(\bar{J}_k)(x)&=H[x,\bar{\mu}_{k-1}(x),\bar{J}_k]\\&\leqq H[x,\bar{\mu}_{k-1}(x),(J^*+\varepsilon_k e)]\leqq H[x,\bar{\mu}_{k-1}(x),J^*]+\alpha\varepsilon_k\\&\leqq H[x,\bar{\mu}_{k-1}(x),\lim_{i\to\infty}(T_{\mu_1^{k-1}[x]}\cdots T_{\mu_i^{k-1}[x]})(\bar{J})]+\alpha\varepsilon_k\\&=\bar{J}_{k-1}(x)+\alpha\varepsilon_k,\end{aligned}$$

and finally

$$T_{\bar{\mu}_{k-1}}(\bar{J}_k) \leqq \bar{J}_{k-1} + \alpha\varepsilon_k e \quad \forall k = 1, 2, \cdots. \tag{43}$$

Using this inequality and I.2 we obtain

$$T_{\bar{\mu}_{k-2}}[T_{\bar{\mu}_{k-1}}(\bar{J}_k)] \leqq T_{\bar{\mu}_{k-2}}(\bar{J}_{k-1} + \alpha\varepsilon_k e)$$
$$\leqq T_{\bar{\mu}_{k-2}}(\bar{J}_{k-1}) + \alpha^2\varepsilon_k e \leqq \bar{J}_{k-2} + (\alpha\varepsilon_{k-1} + \alpha^2\varepsilon_k)e.$$

Continuing in the same manner we obtain for $k = 1, 2, \cdots$,

$$(T_{\bar{\mu}_0} \cdots T_{\bar{\mu}_{k-1}})(\bar{J}_k) \leqq \bar{J}_0 + (\alpha\varepsilon_1 + \cdots + \alpha^k\varepsilon_k)e \leqq J^* + \left(\sum_{i=0}^{k} \alpha^i\varepsilon_i\right)e.$$

Since $\bar{J} \leqq \bar{J}_k$ it follows that

$$(T_{\bar{\mu}_0} \cdots T_{\bar{\mu}_{k-1}})(\bar{J}) \leqq J^* + \left(\sum_{i=0}^{k} \alpha^i\varepsilon_i\right)e.$$

Denote $\pi_\varepsilon = \{\bar{\mu}_0, \bar{\mu}_1, \cdots\}$. Then by taking limit above and using (38) we obtain $J_{\pi_\varepsilon} \leqq J^* + \varepsilon e$. If $\alpha < 1$ take $\varepsilon_k = \varepsilon(1-a)$ and $\pi_k[x] = \{\mu_0[x], \mu_1[x], \cdots\}$ for all k. Then the policy $\pi_\varepsilon = \{\bar{\mu}_0, \bar{\mu}_1, \cdots\}$ is stationary. Q.E.D.

By using Proposition 4 we can prove the following.

PROPOSITION 5. *Let* I, I.1, *and* I.2 *hold. Then*

$$J^* = T(J^*).$$

Furthermore if $J' \in F$ is such that $J' \geqq \bar{J}$ and $J' \geqq T(J')$, then $J' \geqq J^*$.

Proof. For every $\pi = \{\mu_0, \mu_1, \cdots\} \in \Pi$ and $x \in S$ we have using I.1

$$J_\pi(x) = \lim_{k\to\infty} (T_{\mu_0}T_{\mu_1} \cdots T_{\mu_k})(\bar{J})(x)$$
$$= T_{\mu_0}[\lim_{k\to\infty} (T_{\mu_1} \cdots T_{\mu_k})(\bar{J})](x)$$
$$\geqq T_{\mu_0}(J^*)(x) \geqq T(J^*)(x).$$

By taking the infimum of the left hand side over $\pi \in \Pi$

$$J^* \geqq T(J^*).$$

To prove the reverse inequality let $\varepsilon_1, \varepsilon_2$ be any positive scalars and let $\bar{\pi} = \{\bar{\mu}_0, \bar{\mu}_1, \cdots\}$ be such that

$$T_{\bar{\mu}_0}(J^*) \leqq T(J^*) + \varepsilon_1 e,$$
$$J_{\bar{\pi}_1} \leqq J^* + \varepsilon_2 e,$$

where $\bar{\pi}_1 = \{\bar{\mu}_1, \bar{\mu}_2, \cdots\}$. Such a policy exists by Proposition 4. We have

$$J_{\bar{\pi}} = \lim_{k\to\infty} (T_{\bar{\mu}_0}T_{\bar{\mu}_1} \cdots T_{\bar{\mu}_k})(\bar{J})$$
$$= T_{\bar{\mu}_0}[\lim_{k\to\infty} (T_{\bar{\mu}_1} \cdots T_{\bar{\mu}_k})(\bar{J})] = T_{\bar{\mu}_0}(J_{\bar{\pi}_1})$$
$$\leqq T_{\bar{\mu}_0}(J^*) + \alpha\varepsilon_2 e \leqq T(J^*) + (\varepsilon_1 + \alpha\varepsilon_2)e.$$

Since $J^* \leqq J_{\bar{\pi}}$ and ε_1, ε_2 can be taken arbitrarily small it follows that

$$J^* \leqq T(J^*).$$

Hence $J^* = T(J^*)$.

Assume that $J' \in F$ satisfies $J' \geqq \bar{J}$ and $J' \geqq T(J')$. Let $\{\varepsilon_k\}$ be any sequence with $\varepsilon_k > 0$ and consider a policy $\bar{\pi} = \{\bar{\mu}_0, \bar{\mu}_1, \cdots\} \in \Pi$ such that

$$T_{\bar{\mu}_k}(J') \leqq T(J') + \varepsilon_k e, \qquad k = 0, 1, \cdots.$$

We have using I.2,

$$\begin{aligned}
J^* &= \inf_{\pi \in \Pi} \lim_{k \to \infty} (T_{\mu_0} \cdots T_{\mu_k})(\bar{J}) \\
&\leqq \inf_{\pi \in \Pi} \liminf_{k \to \infty} (T_{\mu_0} \cdots T_{\mu_k})(J') \\
&\leqq \liminf_{k \to \infty} (T_{\bar{\mu}_0} \cdots T_{\bar{\mu}_k})(J') \\
&\leqq \liminf_{k \to \infty} (T_{\bar{\mu}_0} \cdots T_{\bar{\mu}_{k-1}})[T(J') + \varepsilon_k e] \\
&\leqq \liminf_{k \to \infty} (T_{\bar{\mu}_0} \cdots T_{\bar{\mu}_{k-2}})[T_{\bar{\mu}_{k-1}}(J' + \varepsilon_k e)] \\
&\leqq \liminf_{k \to \infty} (T_{\bar{\mu}_0} \cdots T_{\bar{\mu}_{k-2}})[T_{\bar{\mu}_{k-1}}(J') + \alpha \varepsilon_k e] \\
&\leqq \liminf_{k \to \infty} [(T_{\bar{\mu}_0} \cdots T_{\bar{\mu}_{k-1}})(J') + \alpha^k \varepsilon_k e] \\
&\cdots \\
&\leqq \lim_{k \to \infty} \left[T(J') + \left(\sum_{i=0}^{k} \alpha^i \varepsilon_i \right) e \right] \leqq J' + \left(\sum_{i=0}^{\infty} \alpha^i \varepsilon_i \right) e.
\end{aligned}$$

Since we may choose $\sum_{i=0}^{\infty} \alpha^i \varepsilon_i$ as small as desired it follows that $J^* \leqq J'$. Q.E.D.

The following counterexamples show that I.1 and I.2 are essential in order for Proposition 5 to hold.

Counterexample 4. Take $S = \{0, 1\}$, $C = U(0) = U(1) = (-1, 0]$, $\bar{J}(0) = \bar{J}(1) = -1$, $H(0, u, J) = u$ if $J(1) \leqq -1$, $H(0, u, J) = 0$ if $J(1) > -1$, and $H(1, u, J) = u$. Then $(T_{\mu_0} \cdots T_{\mu_{N-1}})(\bar{J})(0) = 0$ and $(T_{\mu_0} \cdots T_{\mu_{N-1}})(\bar{J})(1) = \mu_0(1)$ for $N \geqq 1$. Thus $J^*(0) = 0$, $J^*(1) = -1$ while $T(J^*)(0) = -1$, $T(J^*)(1) = -1$ and hence $J^* \neq T(J^*)$. Notice also that $\bar{J}$ is a fixed point of T while $\bar{J} \leqq J^*$ and $\bar{J} \neq J^*$. Here I and I.1 are satisfied but I.2 is violated.

Counterexample 5. Take $S = \{0, 1\}$, $C = U(0) = U(1) = \{0\}$, $\bar{J}(0) = \bar{J}(1) = 0$, $H(0, 0, J) = 0$ if $J(1) < \infty$, $H(0, 0, J) = \infty$ if $J(1) = \infty$, $H(1, 0, J) = J(1) + 1$. Then $(T_{\mu_0} \cdots T_{\mu_{N-1}})(\bar{J})(0) = 0$ and $(T_{\mu_0} \cdots T_{\mu_{N-1}})(\bar{J})(1) = N$. Thus $J^*(0) = 0$, $J^*(1) = \infty$. On the other hand we have $T(J^*)(0) = T(J^*)(1) = \infty$ and $J^* \neq T(J^*)$. Here I and I.2 are satisfied but I.1 is violated.

As a corollary to Proposition 5 we obtain the following:

COROLLARY 5.1. *Let* I, I.1 *and* I.2 *hold. Then for every stationary policy* $\pi = \{\mu, \mu, \cdots\}$ *there holds*

$$J_\mu = T_\mu(J_\mu).$$

Furthermore if $J' \in F$ *is such that* $J' \geqq \bar{J}$, $J' \geqq T_\mu(J')$, *then* $J' \geqq J_\mu$.

Proof. Consider the variation of our problem where the control constraint set is $U_\mu(x) = \{\mu(x)\}\ \forall x \in X$ rather than $U(x)$. Application of Proposition 5 yields the result. Q.E.D.

We now provide the counterpart of Proposition 5 under Assumption D.

PROPOSITION 6. *Let* D *and* D.1 *hold. Then*

$$J^* = T(J^*).$$

Furthermore if $J' \in F$ *is such that* $J' \leqq \bar{J}$, $J' \leqq T(J')$, *then* $J' \leqq J^*$.

Proof. We first show the following lemma:

LEMMA 1. *Let* D *hold. Then*

$$J^* = \lim_{N\to\infty} J_N, \tag{44}$$

where J_N *is the optimal value function of the N-stage problem defined by* (36).

Proof. Clearly we have $J^* \leqq J_N$ for all N and hence $J^* \leqq \lim_{N\to\infty} J_N$. Also for all $\pi = \{\mu_0, \mu_1, \cdots\} \in \Pi$ we have

$$(T_{\mu_0} \cdots T_{\mu_{N-1}})(\bar{J}) \geqq J_N,$$

and by taking limit of both sides we obtain $J_\pi \geqq \lim_{N\to\infty} J_N$, and hence $J^* \geqq \lim_{N\to\infty} J_N$. Q.E.D.

We return to the proof of Proposition 6. An entirely similar argument as the one of the proof of Lemma 1 shows that under D we have for all $x \in S$,

$$\lim_{N\to\infty} \inf_{u\in U(x)} H(x, u, J_N) = \inf_{u\in U(x)} \lim_{N\to\infty} H(x, u, J_N). \tag{45}$$

Using D.1 the above equation yields

$$\lim_{N\to\infty} T(J_N) = T(\lim_{N\to\infty} J_N). \tag{46}$$

By Proposition 3 we have $J_N = T^N(\bar{J})$ and hence $T(J_N) = T^{N+1}(\bar{J})$. Combining this relation with (44) and (46) we obtain $J^* = T(J^*)$.

To complete the proof, let $J' \in F$ be such that $J' \leqq \bar{J}$, $J' \leqq T(J')$. Then we have

$$\begin{aligned} J^* &= \inf_{\pi\in\Pi} \lim_{M\to\infty} (T_{\mu_0} \cdots T_{\mu_{N-1}})(\bar{J}) \\ &\geqq \lim_{N\to\infty} \inf_{\pi\in\Pi} (T_{\mu_0} \cdots T_{\mu_{N-1}})(\bar{J}) \\ &\geqq \lim_{N\to\infty} \inf_{\pi\in\Pi} (T_{\mu_0} \cdots T_{\mu_{N-1}})(J') \\ &\geqq \lim_{N\to\infty} T^N(J') \geqq J'. \end{aligned}$$

Hence $J^* \geqq J'$. Q.E.D.

In Counterexamples 2 and 3, Assumption D.1 does not hold. In both cases we have $J^* \neq T(J^*)$ as the reader can easily verify.

A cursory examination of the proof of Proposition 6 reveals that the only point where we used D.1 was in establishing the relation $\lim_{N\to\infty} T(J_N) = T(\lim_{N\to\infty} J_N)$ [cf. (46)]. Hence if this relation can be established independently then the result of Proposition 6 follows. In this manner we obtain the following corollary.

COROLLARY 6.1. *Let* D *hold and assume that* D.2 *holds, S is a finite set, and $J^*(x) > -\infty$ for all $x \in S$. Then $J^* = T(J^*)$. Furthermore if $J' \in F$ is such that $J' \leqq \bar{J}$, $J' \leqq T(J')$, then $J' \leqq J^*$.*

Proof. We will show that

$$\lim_{N\to\infty} H(x, u, J_N) = H(x, u, \lim_{N\to\infty} J_N) \quad \forall x \in S, \quad u \in U(x).$$

Then using (45) we obtain (46) and the result follows as in the proof of Proposition 6. Assume the contrary, i.e., that for some $\tilde{x} \in S$, $\tilde{u} \in U(\tilde{x})$, and $\varepsilon > 0$ there holds

$$H(\tilde{x}, \tilde{u}, J_k) - \varepsilon > H(\tilde{x}, \tilde{u}, \lim_{N\to\infty} J_N) \quad \forall k = 1, 2, \cdots.$$

Using the finiteness of S and the fact $J^*(x) = \lim_{N\to\infty} J_N(x) > -\infty$ for all x we obtain that for some positive integer $\bar{k}$ we have

$$J_k - \frac{\varepsilon}{\alpha} e \leqq \lim_{N\to\infty} J_N \quad \forall k \geqq \bar{k}.$$

By using D.2 we obtain for all $k \geqq \bar{k}$,

$$H(\tilde{x}, \tilde{u}, J_k) - \varepsilon \leqq H\left(\tilde{x}, \tilde{u}, J_k - \frac{\varepsilon}{\alpha} e\right) \leqq H(\tilde{x}, \tilde{u}, \lim_{N\to\infty} J_N)$$

which contradicts the earlier inequality. Q.E.D.

Similarly as under I we have the following corollary:

COROLLARY 6.2. *Let* D *and* D.1 *hold. Then for every stationary policy $\pi = \{\mu, \mu, \cdots\}$ there holds*

$$J_\mu = T_\mu(J_\mu).$$

Furthermore if $J' \in F$ is such that $J' \leqq \bar{J}$, $J' \leqq T_\mu(J')$ then $J' \leqq J_\mu$.

It is worth noting that Propositions 5 and 6 may form the basis for computation of J^* when the state space S is a finite set with n elements denoted $x_1, x_2, \cdots, x_n$. It follows from Proposition 5 that, under I, I.1, and I.2, $J^*(x_1), \cdots, J^*(x_n)$ solve the problem

$$\text{minimize} \sum_{i=1}^{n} \lambda_i$$

subject to

$$\lambda_i \geqq \inf_{u \in U(x_i)} H(x_i, u, J_\lambda), \qquad i = 1, \cdots, n,$$

$$\lambda_i \geqq \bar{J}(x_i), \qquad i = 1, \cdots, n,$$

where J_λ is the function taking values $J_\lambda(x_i) = \lambda_i$, $i = 1, \cdots, n$. Under D and D.1, or D, D.2 and $J^*(x) > -\infty$ $\forall x \in S$ the corresponding problem is

$$\text{maximize } \sum_{i=1}^{n} \lambda_i$$

subject to

$$\lambda_i \leqq \inf_{u \in U(x_i)} H(x_i, u, J_\lambda), \qquad i = 1, \cdots, n,$$

$$\lambda_i \leqq \bar{J}(x_i), \qquad i = 1, \cdots, n.$$

When $U(x_i)$ is also a finite set for all i, then the problems above become finite-dimensional nonlinear programming problems.

Characterization of optimal stationary policies. We have the following necessary and sufficient conditions for optimality of a stationary policy.

PROPOSITION 7. *Let* I, I.1, *and* I.2 *hold. Then a stationary policy* $\pi^* = \{\mu^*, \mu^*, \cdots\}$ *is optimal if and only if*

$$T_{\mu^*}(J^*) = T(J^*). \tag{47}$$

Furthermore if there exists an optimal policy there exists an optimal stationary policy.

Proof. If π^* is optimal then $J_{\mu^*} = J^*$ and the result follows from Proposition 5 and Corollary 5.1. Conversely if $T_{\mu^*}(J^*) = T(J^*)$ then $J^* = T(J^*)$ (by Proposition 5) and it follows that $T_{\mu^*}(J^*) = J^*$. Hence by Corollary 5.1, $J_{\mu^*} \leqq J^*$ and it follows that π^* is optimal. If $\bar{\pi} = \{\bar{\mu}_0, \bar{\mu}_1, \cdots\}$ is optimal then we have by using I.1

$$J^* = J_{\bar{\pi}} = \lim_{k \to \infty} (T_{\bar{\mu}_0} T_{\bar{\mu}_1} \cdots T_{\bar{\mu}_k})(\bar{J})$$

$$= T_{\bar{\mu}_0}[\lim_{k \to \infty} (T_{\bar{\mu}_1} \cdots T_{\bar{\mu}_k})](\bar{J}) \geqq T_{\bar{\mu}_0}(J^*) \geqq T(J^*) = J^*.$$

It follows that $T_{\bar{\mu}_0}(J^*) = T(J^*)$ and, by the result just proved, the stationary policy $\{\bar{\mu}_0, \bar{\mu}_0, \cdots\}$ is optimal. Q.E.D.

PROPOSITION 8. *Let* D *and* D.1 *hold. Then a stationary policy* $\pi^* = \{\mu^*, \mu^*, \cdots)$ *is optimal if and only if*

$$T_{\mu^*}(J_{\mu^*}) = T(J_{\mu^*}). \tag{48}$$

Proof. If π^* is optimal then $J_{\mu^*} = J^*$ and, using Proposition 6, and Corollary 6.2, we have $T_{\mu^*}(J_{\mu^*}) = J_{\mu^*} = J^* = T(J^*) = T(J_{\mu^*})$. Conversely if $T_{\mu^*}(J_{\mu^*}) = T(J_{\mu^*})$ then we obtain from Corollary 6.2, $J_{\mu^*} = T(J_{\mu^*})$, and Proposition 6 yields $J_{\mu^*} \leqq J^*$. Hence π^* is optimal. Q.E.D.

Examples where π^* satisfies (47) or (48) but is not optimal under D or I respectively are given in [1, § 6.4]. It is also easy to modify the proof of Proposition 7 and show the stronger result that if there exists an optimal policy at each $x \in S$ then there exists an optimal stationary policy.

Convergence of the dynamic algorithm—existence of optimal stationary policies. The D.P. algorithm consists of successive generation of the function $T(\bar{J}), T^2(\bar{J}), \cdots$. Under either Assumption I or D the function $J_\infty \in F$ defined by

$$J_\infty(x) = \lim_{N \to \infty} T^N(\bar{J})(x) \quad \forall x \in S \tag{49}$$

is well defined. We would like to investigate the question whether $J_\infty = J^*$. When Assumption D holds, the following proposition shows that we have $J_\infty = J^*$ under mild assumptions.

PROPOSITION 9. *Let* D *hold and assume that either* D.1 *holds or else* $J_N = T^N(\bar{J})$ *for all* N *where* J_N *is the optimal value function of the* N*-stage problem defined by* (36). *Then*

$$J_\infty = J^*.$$

Proof. By Lemma 1 we have that D implies $J^* = \lim_{N\to\infty} J_N$. Proposition 3 shows also that under our assumptions $J_N = T^N(\bar{J})$. Hence $J^* = \lim_{N\to\infty} T^N(\bar{J}) = J_\infty$. Q.E.D.

Under Assumptions I, I.1, and I.2 the equality $J_\infty = J^*$ may easily fail to hold even in very simple deterministic optimal control problems as shown in the example of § 1. This fact has been known since Strauch's work (see [11, p. 880]). Reference [2, p. 608] provides an example where $J_\infty \neq J^*$ even though there exists an optimal stationary policy. The following preliminary result shows that in order to have $J_\infty = J^*$ it is necessary and sufficient to have $J_\infty = T(J_\infty)$.

PROPOSITION 10. *Let* I, I.1, *and* I.2 *hold. Then*

$$J_\infty \leqq T(J_\infty) \leqq T(J^*) = J^*. \tag{50}$$

Furthermore the relation

$$J_\infty = T(J_\infty) = T(J^*) = J^* \tag{51}$$

holds if and only if

$$J_\infty = T(J_\infty). \tag{52}$$

Proof. Clearly we have $J_\infty \leqq J_\pi$ for all $\pi \in \Pi$ and it follows that $J_\infty \leqq J^*$. Furthermore by Proposition 5 we have $T(J^*) = J^*$. Also we have for all $k \geqq 1$,

$$T(J_\infty) = \inf_{u\in U(x)} H(x, u, J_\infty) \geqq \inf_{u\in U(x)} H[x, u, T^k(\bar{J})] = T^{k+1}(\bar{J}).$$

Taking limit of the right side we obtain $T(J_\infty) \geqq J_\infty$ which combined with $J_\infty \leqq J^*$ and $T(J^*) = J^*$ proves (50). If (51) holds then (52) also holds. Conversely let (52) hold. Then, since we have $J_\infty \geqq \bar{J}$, we obtain by Proposition 5, $J_\infty \geqq J^*$ which combined with (50) proves (51). Q.E.D.

In what follows we provide a necessary and sufficient condition for $J_\infty = T(J_\infty)$ (and hence also (51)) to hold under Assumptions I, I.1, and I.2. We subsequently obtain a useful sufficient condition for $J_\infty = T(J_\infty)$ to hold which at the same time guarantees the existence of an optimal stationary policy.

For any $J \in F$ we denote by $E(J)$ the *epigraph* of J, i.e. the subset of $S \times (-\infty, \infty)$ given by

$$E(J) = \{(x, \lambda) | J(x) \leqq \lambda\}. \tag{53}$$

Under I we have $T^k(\bar{J}) \leqq T^{k+1}(\bar{J})$ for all k, $J_\infty = \lim_{k\to\infty} T^k(\bar{J})$, and it follows easily that

$$E(J_\infty) = \bigcap_{k=0}^{\infty} E[T^k(\bar{J})]. \tag{54}$$

Consider for each $k \geqq 1$ the subset C_k of $S \times C \times (-\infty, \infty)$ given by

(55) $$C_k = \{(x, u, \lambda) | H[x, u, T^{k-1}(\bar{J})] \leqq \lambda, x \in S, u \in U(x)\}.$$

Denote $P(C_k)$ the projection of C_k on $S \times (-\infty, \infty)$,

(56) $$P(C_k) = \{(x, \lambda) | \exists u \in U(x) \text{ s.t. } (x, u, \lambda) \in C_k\}.$$

In the above relation and later the symbol $\exists$ denotes "there exists" and the initials "s.t." stand for "such that". Consider also the following set:

(57) $$\overline{P(C_k)} = \{(x, \lambda) | \exists \{\lambda_n\} \text{ s.t. } \lambda_n \to \lambda, (x, \lambda_n) \in P(C_k), n = 0, 1, \cdots\}.$$

The set $\overline{P(C_k)}$ is obtained from $P(C_k)$ by adding for each x the point $[x, \bar{\lambda}(x)]$ where $\bar{\lambda}(x)$ is the perhaps missing end point of the half line $\{\lambda | (x, \lambda) \in P(C_k)\}$. We have the following lemma:

LEMMA 2. *Let* I *hold. Then for all* $k \geqq 1$,

(58) $$P(C_k) \subset \overline{P(C_k)} = E[T^k(\bar{J})].$$

Furthermore we have

(59) $$P(C_k) = \overline{P(C_k)} = E[T^k(\bar{J})]$$

if and only if the infimum is attained for each $x \in S$ *in the relation*

(60) $$T^k(\bar{J})(x) = \inf_{u \in U(x)} H[x, u, T^{k-1}(\bar{J})].$$

Proof. If $(x, \lambda) \in E[T^k(\bar{J})]$ we have

$$T^k(\bar{J})(x) = \inf_{u \in U(x)} H[x, u, T^{k-1}(\bar{J})] \leqq \lambda.$$

Let $\{\varepsilon_n\}$ be a sequence such that $\varepsilon_n > 0$, $\varepsilon_n \to 0$ and let $\{u_n\}$ be a sequence such that

$$H[x, u_n, T^{k-1}(\bar{J})] \leqq T^k(\bar{J})(x) + \varepsilon_n \leqq \lambda + \varepsilon_n.$$

Then $(x, u_n, \lambda + \varepsilon_n) \in C_k$ and $(x, \lambda + \varepsilon_n) \in P(C_k)$ for all n. Since $\{\lambda + \varepsilon_n\} \to \lambda$ by (57) we obtain $(x, \lambda) \in \overline{P(C_k)}$. Hence

(61) $$E[T^k(\bar{J})] \subset \overline{P(C_k)}.$$

Conversely let $(x, \lambda) \in \overline{P(C_k)}$. Then by (55)–(57) there exists a sequence $\{\lambda_n\}$ with $\lambda_n \to \lambda$ and a corresponding sequence $\{u_n\}$ such that

$$T^k(\bar{J})(x) \leqq H[x, u_n, T^{k-1}(\bar{J})] \leqq \lambda_n.$$

Taking limit as $n \to \infty$ we obtain $T^k(\bar{J})(x) \leqq \lambda$ and $(x, \lambda) \in E[T^k(\bar{J})]$. Hence

$$\overline{P(C_k)} \subset E[T^k(\bar{J})]$$

and using (61), we obtain (58).

To prove that (59) is equivalent to the attainment of the infimum in (60) assume first that the infimum is attained by $\mu_{k-1}^*(x)$ for each $x \in S$. Then for each

$(x, \lambda) \in E[T^k(\bar{J})]$

$$H[x, \mu_{k-1}^*(x), T^{k-1}(\bar{J})] \leqq \lambda$$

implying by (56) that $(x, \lambda) \in P(C_k)$. Hence $E[T^k(\bar{J})] \subset P(C_k)$ and in view of (58) we obtain (59). Conversely if (59) holds we have $[x, T^k(\bar{J})(x)] \in P(C_k)$ for every x for which $T^k(\bar{J})(x) < \infty$. Then by (55), (56) there exists a $\mu_{k-1}^*(x) \in U(x)$ such that

$$H[x, \mu_{k-1}^*(x), T^{k-1}(\bar{J})] \leqq T^k(\bar{J})(x) = \inf_{u \in U(x)} H[x, u, T^{k-1}(\bar{J})].$$

Hence the infimum in (56) is attained for all x such that $T^k(\bar{J})(x) < \infty$. It is also trivially attained by all $u \in U(x)$ whenever $T^k(\bar{J})(x) = \infty$ and the proof is complete. Q.E.D.

Consider now the set $\bigcap_{k=1}^{\infty} C_k$ and define similarly as in (56), (57) the sets

(62) $$P\left(\bigcap_{k=1}^{\infty} C_k\right) = \left\{(x, \lambda) | \exists u \in U(x) \text{ s.t. } (x, u, \lambda) \in \bigcap_{k=1}^{\infty} C_k\right\},$$

(63) $$\overline{P\left(\bigcap_{k=1}^{\infty} C_k\right)} = \left\{(x, \lambda) | \exists \{\lambda_n\} \text{ s.t. } \lambda_n \to \lambda, (x, \lambda_n) \in P\left(\bigcap_{k=1}^{\infty} C_k\right)\right\}.$$

Using (54) and Lemma 2 it is easy to see that we have

(64) $$P\left(\bigcap_{k=1}^{\infty} C_k\right) \subset \bigcap_{k=1}^{\infty} P(C_k) \subset \bigcap_{k=1}^{\infty} \overline{P(C_k)} = \bigcap_{k=1}^{\infty} E[T^k(\bar{J})] = E(J_\infty),$$

(65) $$\overline{P\left(\bigcap_{k=1}^{\infty} C_k\right)} \subset \bigcap_{k=1}^{\infty} \overline{P(C_k)} = \bigcap_{k=1}^{\infty} E[T^k(\bar{J})] = E(J_\infty).$$

We have the following proposition:

PROPOSITION 11. *Let* I, I.1, *and* I.2 *hold. Then*:

(a) *There holds* $J_\infty = T(J_\infty)$ (*equivalently* $J_\infty = J^*$) *if and only if*

(66) $$\overline{P\left(\bigcap_{k=1}^{\infty} C_k\right)} = \bigcap_{k=1}^{\infty} \overline{P(C_k)}.$$

(b) *There holds* $J_\infty = T(J_\infty)$ (*equivalently* $J_\infty = J^*$) *and in addition the infimum in*

(67) $$J_\infty(x) = \inf_{u \in U(x)} H(x, u, J_\infty)$$

is attained for each $x \in S$ (*equivalently there exists an optimal stationary policy*) *if and only if*

(68) $$P\left(\bigcap_{k=1}^{\infty} C_k\right) = \bigcap_{k=1}^{\infty} \overline{P(C_k)}.$$

Proof. (a) Assume $J_\infty = T(J_\infty)$ and let $(x, \lambda) \in E(J_\infty)$, i.e.

$$\inf_{u \in U(x)} H(x, u, J_\infty) = J_\infty(x) \leqq \lambda.$$

Let $\{\varepsilon_n\}$ be any sequence with $\varepsilon_n > 0$, $\varepsilon_n \to 0$. Then there exists a sequence $\{u_n\}$ such that

$$H(x, u_n, J_\infty) \leqq \lambda + \varepsilon_n \quad \forall n = 1, 2, \cdots,$$

and hence

$$H[x, u_n, T^{k-1}(\bar{J})] \leqq \lambda + \varepsilon_n \quad \forall k, n = 1, 2, \cdots.$$

Hence $(x, u_n, \lambda + \varepsilon_n) \in C_k$ for all k, n and $(x, u_n, \lambda + \varepsilon_n) \in \bigcap_{k=1}^\infty C_k$ for all n. Hence $(x, \lambda + \varepsilon_n) \in P(\bigcap_{k=1}^\infty C_k)$ for all n and since $\lambda + \varepsilon_n \to \lambda$ we obtain $(x, \lambda) \in \overline{P(\bigcap_{k=1}^\infty C_k)}$. Therefore

$$E(J_\infty) \subset \overline{P\left(\bigcap_{k=1}^{\infty} C_k\right)}$$

and by (65) we obtain (66).

Conversely let (66) hold. Then we have by (65) $\overline{P(\bigcap_{k=1}^\infty C_k)} = E(J_\infty)$. Let $x \in S$ be such that $J_\infty(x) < \infty$. Then $[x, J_\infty(x)] \in \overline{P(\bigcap_{k=1}^\infty C_k)}$ and there exists a sequence $\{\lambda_n\}$ with $\lambda_n \to J_\infty(x)$ and a sequence $\{u_n\}$ such that

$$H[x, u_n, T^{k-1}(J^-)] \leqq \lambda_n \quad \forall k, n = 1, 2, \cdots.$$

Taking limit with respect to k and using I.1 we obtain

$$H(x, u_n, J_\infty) \leqq \lambda_n \quad \forall n = 1, 2, \cdots,$$

and since $T(J_\infty)(x) \leqq H(x, u_n, J_\infty)$ it follows that

$$T(J_\infty)(x) \leqq \lambda_n.$$

Taking limit as $n \to \infty$ we obtain

$$T(J_\infty)(x) \leqq J_\infty(x)$$

for all $x \in S$ such that $J_\infty(x) < \infty$. Since the inequality above holds also for all $x \in S$ with $J_\infty(x) = \infty$ we have

$$T(J_\infty) \leqq J_\infty.$$

On the other hand by Proposition 10 we have $J_\infty \leqq T(J_\infty)$. Combining the two inequalities we have $J_\infty = T(J_\infty)$.

(b) Assume $J_\infty = T(J_\infty)$ and that the infimum in (67) is attained for each $x \in S$. Then there exists a function $\mu^* \in M$ such that for all $(x, \lambda) \in E(J_\infty)$

$$H[x, \mu^*(x), J_\infty] \leqq \lambda.$$

Hence $H[x, \mu^*(x), T^{k-1}(\bar{J})] \leqq \lambda$ for all k and $[x, \mu^*(x), \lambda] \in \bigcap_{k=1}^\infty C_k$. As a result $(x, \lambda) \in P(\bigcap_{k=1}^\infty C_k)$. Hence

$$E(J_\infty) \subset P\left(\bigcap_{k=1}^{\infty} C_k\right)$$

and, by (64), equation (68) follows.

Conversely let (68) hold. We have for all $x \in S$ with $J_\infty(x) < \infty$,

$$[x, J_\infty(x)] \in E(J_\infty) = P\left(\bigcap_{k=1}^{\infty} C_k\right).$$

Hence there exists a $\mu^*(x) \in U(x)$ such that

$$[x, \mu^*(x), J_\infty(x)] \in \bigcap_{k=1}^{\infty} C_k$$

from which

$$H[x, \mu^*(x), T^{k-1}(\bar{J})] \leqq J_\infty(x) \quad \forall k = 0, 1, \cdots.$$

Taking limit and using I.1, we have

$$T(J_\infty)(x) \leqq H[x, \mu^*(x), J_\infty] \leqq J_\infty(x).$$

It follows that $T(J_\infty) \leqq J_\infty$ and since by Proposition 10, $J_\infty \leqq T(J_\infty)$ we finally obtain $J_\infty = T(J_\infty)$. Furthermore the inequality above shows that $\mu^*(x)$ attains the infimum in (67) when $J_\infty(x) < \infty$. When $J_\infty(x) = \infty$ every $u \in U(x)$ attains the infimum and the proof is complete. Q.E.D.

The proposition above states that the equality $J_\infty = T(J_\infty)$, which in view of Proposition 10 is equivalent to the validity of interchanging infimum and limit as shown below:

$$J_\infty = \lim_{k\to\infty} \inf_{\pi \in \Pi} (T_{\mu_0} \cdots T_{\mu_k})(\bar{J}) = \inf_{\pi\in\Pi} \lim_{k\to\infty} (T_{\mu_0} \cdots T_{\mu_k})(\bar{J}) = J^*,$$

is in fact equivalent to the validity of interchanging projection and intersection in the manner of (66) or (68).

The following proposition provides a compactness assumption which guarantees that (68) holds. If C is a topological space (see e.g. [5]) we say that a subset U of C is compact if every collection of open sets that covers U has a finite subcollection that covers U. The empty set in particular is considered to be compact. Any sequence $\{u_n\}$ belonging to a compact set $U \subset C$ has at least one accumulation point $\bar{u} \in U$, i.e., a point $\bar{u} \in U$ every (open) neighborhood of which contains an infinite number of elements of $\{u_n\}$. Furthermore all accumulation points of $\{u_n\}$ belong to U. If $\{U_n\}$ is a sequence of nonempty compact sets of C and $U_n \supset U_{n+1}$ for all n, then the intersection $\bigcap_{n=1}^\infty U_n$ is nonempty and compact. This yields the following lemma which will be useful in what follows.

LEMMA 3. *Let C be a topological space, $f\colon C \to [-\infty, +\infty]$ be a function, and U be a subset of C. Assume that the set $U(\lambda)$ defined by*

$$U(\lambda) = \{u \in U \mid f(u) \leqq \lambda\}$$

is compact for each $\lambda \in (-\infty, \infty)$. Then f attains a minimum over U.

Proof. If $f(u) = +\infty$ for all $u \in U$ then every $u \in U$ attains the minimum. If $f^* = \inf\{f(u) \mid u \in U\} < +\infty$ let $\{\lambda_n\}$ be a scalar sequence such that $\lambda_n > \lambda_{n+1}$ for all n and $\lambda_n \to f^*$. Then the sets $U(\lambda_n)$ are nonempty, compact, and satisfy $U(\lambda_n) \supset U(\lambda_{n+1})$ for all n. Hence the intersection $\bigcap_{n=1}^\infty U(\lambda_n)$ is nonempty and compact. Let u^* be any point in the intersection. Then $u^* \in U$ and $f(u^*) \leqq \lambda_n$ for all n, and it follows that $f(u^*) \leqq f^*$. Hence u^* attains the minimum of f over U. Q.E.D.

PROPOSITION 12. *Let* I, I.1 *and* I.2 *hold and let the control space C be a topological space. Assume that there exists a nonnegative integer $\bar{k}$ such that for each $x \in S$, $\lambda \in (-\infty, \infty)$ and $k \geqq \bar{k}$ the set*

$$U_k(x, \lambda) = \{u \in U(x) \mid H[x, u, T^k(\bar{J})] \leqq \lambda\} \tag{69}$$

is compact. Then

$$P\Big(\bigcap_{k=1}^{\infty} C_k\Big)=\bigcap_{k=1}^{\infty} \overline{P(C_k)} \tag{70}$$

and (by Propositions 10 *and* 11) *there holds*

$$J_\infty = T(J_\infty) = T(J^*) = J^*.$$

Furthermore there exists an optimal stationary policy.

Proof. By (64) it will be sufficient to show that

$$P\Big(\bigcap_{k=1}^{\infty} C_k\Big)\supset\bigcap_{k=1}^{\infty} P(C_k), \qquad \bigcap_{k=1}^{\infty} P(C_k)=\bigcap_{k=1}^{\infty} \overline{P(C_k)}. \tag{71}$$

Let $(x,\lambda)\in\bigcap_{k=1}^{\infty}P(C_k)$. Then there exists a sequence $\{u_n\}$ such that

$$H[x,u_n,T^k(\bar{J})]\leqq H[x,u_n,T^n(\bar{J})]\leqq\lambda \quad \forall n\geqq k,$$

or equivalently

$$u_n\in U_k(x,\lambda) \quad \forall n\geqq k.$$

Since $U_k(x,\lambda)$ is compact for $k\geqq\bar{k}$ it follows that the sequence $\{u_n\}$ has an accumulation point $\bar{u}$ and

$$\bar{u}\in U_k(x,\lambda) \quad \forall k\geqq\bar{k}.$$

Hence

$$H[x,\bar{u},T^k(\bar{J})]\leqq\lambda$$

and $(x,\bar{u},\lambda)\in\bigcap_{k=1}^{\infty}C_k$. It follows that $(x,\lambda)\in P(\bigcap_{k=1}^{\infty}C_k)$ and

$$P\Big(\bigcap_{k=1}^{\infty} C_k\Big)\supset\bigcap_{k=1}^{\infty} P(C_k).$$

Also by the compactness of $U_k(x,\lambda)$ and the result of Lemma 3 it follows that the infimum in (60) is attained for every $x\in S$ and $k>\bar{k}$. Hence, by Lemma 2, $P(C_k)=\overline{P(C_k)}$ for $k>\bar{k}$ and

$$\bigcap_{k=1}^{\infty} P(C_k)=\bigcap_{k=1}^{\infty} \overline{P(C_k)}.$$

Thus (71) is proved. Q.E.D.

The following proposition shows also that a stationary optimal policy may be obtained in the limit by means of the D.P. algorithm.

PROPOSITION 13. *Let the assumptions of Proposition* 12 *hold. Then*:

(a) *There exists a policy* $\pi^*=\{\mu_0^*,\mu_1^*,\cdots\}\in\Pi$ *attaining the minimum in the D.P. algorithm for all* $k\geqq\bar{k}$, *i.e.*

$$(T_{\mu_k^*}T^k)(\bar{J})=T^{k+1}(\bar{J}) \quad \forall k\geqq\bar{k}. \tag{72}$$

(b) *For every policy* π^* *satisfying* (72) *the sequence* $\{\mu_k^*(x)\}$ *has at least one accumulation point for each* $x\in S$ *with* $J^*(x)<\infty$.

(c) *If* $\mu^*: S \to C$ *is such that* $\mu^*(x)$ *is an accumulation point of* $\{\mu_k^*(x)\}$ *for all* $x \in S$ *with* $J^*(x) < \infty$, *and* $\mu^*(x) \in U(x)$ *for all* $x \in S$ *with* $J^*(x) = \infty$, *then the stationary policy* $\{\mu^*, \mu^*, \cdots\}$ *is optimal.*

Proof. (a) For an $x \in S$ such that $T^{k+1}(\bar{J})(x) < \infty$ consider a sequence $\{\lambda_n\}$ with $\lambda_n > \lambda_{n+1}$, for all n and $\lambda_n \to T^{k+1}(\bar{J})(x)$. Then the sets $U_k(x, \lambda_n)$ are nonempty and compact and hence their intersection is also nonempty and compact. Any point $\mu_k^*(x)$ in the intersection satisfies $(T_{\mu_k^*}T^k)(\bar{J})(x) = T^{k+1}(\bar{J})(x)$. For an $x \in S$ such that $T^{k+1}(\bar{J})(x) = \infty$ any element of $U(x)$, call it $\mu_k^*(x)$, satisfies $(T_{\mu_k^*}T^k)(\bar{J})(x) = T^{k+1}(\bar{J})(x)$.

(b) For any $\pi^* = \{\mu_0^*, \mu_1^*, \cdots\}$ satisfying (72) and $x \in S$ such that $J^*(x) < \infty$ we have

$$H[x, \mu_n^*(x), T^k(\bar{J})] \leqq H[x, \mu_n^*(x), T^n(\bar{J})] \leqq J^*(x) \quad \forall k \geqq \bar{k}, \quad n \geqq k.$$

Hence we have

$$\mu_n^*(x) \in U_k[x, J^*(x)] \quad \forall k \geqq \bar{k}, \quad n \geqq k.$$

Since $U_k[x, J^*(x)]$ is compact, $\{\mu_n^*(x)\}$ has at least one accumulation point. Furthermore each accumulation point $\mu^*(x)$ of $\{\mu_n^*(x)\}$ belongs to $U(x)$ and satisfies

$$H[x, \mu^*(x), T^k(\bar{J})] \leqq J^*(x) \quad \forall k \geqq \bar{k}. \tag{73}$$

(c) By taking the limit in (73) and using I.1 we obtain

$$H[x, \mu^*(x), J_\infty] = H[x, \mu^*(x), J^*] \leqq J^*(x)$$

for all $x \in S$ with $J^*(x) < \infty$. The relation above holds also trivially for all $x \in S$ with $J^*(x) = \infty$. Hence $T_{\mu^*}(J^*) \leqq J^* = T(J^*)$ which implies $T_{\mu^*}(J^*) = T(J^*)$. It follows, by Proposition 7, that $\{\mu^*, \mu^*, \cdots\}$ is optimal. Q.E.D.

The compactness of the sets $U_k(x, \lambda)$ of (69) may be verified in a number of important special cases. One such case is when $U_k(x, \lambda)$ is a finite set for all k, x, λ. Simply consider the discrete topology on C, i.e. the topology consisting of all subsets of U. In this topology a set is compact if and only if it is finite. For this case the relation $J_\infty = J^*$ for the negative model of Strauch has been shown earlier [11]. There are many other important cases where the compactness of $U_k(x, \lambda)$ can be verified. Several examples have been given in [1, (Chap. 6 and 7)]. It is not our intention to provide an extensive list. Instead we state as an illustration two sets of assumptions which guarantee compactness of $U_k(x, \lambda)$ in the case of the mapping

$$H(x, u, J) = g(x, u) + \alpha(x, u)J[f(x, u)]$$

corresponding to a deterministic optimal control problem.

Assume that $g(x, u) \geqq 0$, $\alpha(x, u) \geqq 0$ for all $x \in S$, $u \in U(x)$ and take $\bar{J}(x) = 0$, $\forall x \in S$. Then compactness of $U_k(x, \lambda)$ is guaranteed if:

(a) $S = R^n$ (n-dimensional Euclidean space), $C = R^m$, $U(x) \equiv C$, f, g, α are continuous in (x, u) and g satisfies $\lim_{n\to\infty} g(x_n, u_n) = \infty$ for every bounded sequence $\{x_n\}$ and every sequence $\{u_n\}$ for which $|u_n| \to \infty$ ($|\cdot|$ is a norm on R^m).

(b) $S = R^n$, $C = R^m$, f, g, and α are continuous, $U(x)$ is compact and nonempty for each $x \in R^n$, and $U(\cdot)$ is a continuous point-to-set mapping from R^n to the set of all compact subsets of R^m.

Aside from the result of Strauch mentioned earlier, other general sufficient conditions which guarantee that an optimal stationary policy exists for special cases of our problem are those of Maitra for discounted problems (see [9] and [6, Thm. 5.11]), and Kushner for free end time problems [8]. In these cases Assumption C is satisfied. In both cases the sufficient conditions or existence of an optimal stationary policy can be shown to follow from Proposition 12.

We finally show that the compactness of the sets $U_k(x, \lambda)$ of (69) guarantees existence of an optimal stationary policy under Assumption C which can be obtained in the limit by means of the D.P. algorithm.

PROPOSITION 14. *The conclusions of Proposition* 13 *hold if Assumptions* I. I.1, *and* I.2 *are replaced by the Contraction Assumption* C.

Proof. (a) The proof of this part is identical to the corresponding proof in Proposition 13.

(b) Let $\pi^* = \{\mu_0^*, \mu_1^*, \cdots\}$ satisfy (72) and define

$$\varepsilon_k = \sup \{\|T^i(\bar{J}) - J^*\| \,|\, i \geqq k\}, \qquad k = 0, 1, \cdots.$$

We have from (20), (72) and the fact $T(J^*) = J^*$,

$$\|(T_{\mu_n^*}T^n)(\bar{J}) - J^*\| = \|T^{n+1}(\bar{J}) - T^{n+1}(J^*)\|$$
$$\leqq \alpha\|T^n(\bar{J}) - T^n(J^*)\| = \alpha\|T^n(\bar{J}) - J^*\| \quad \forall n \geqq \bar{k},$$
$$\|(T_{\mu_n^*}T^n)(\bar{J}) - (T_{\mu_n^*}T^k)(\bar{J})\| \leqq \alpha\|T^n(\bar{J}) - T^k(\bar{J})\|$$
$$\leqq \alpha\|T^n(\bar{J}) - J^*\| + \alpha\|T^k(\bar{J}) - J^*\|$$
$$\forall n \geqq \bar{k}, \quad k = 0, 1, \cdots.$$

From the above two relations we obtain

$$H[x, \mu_n^*(x), T^k(\bar{J})] \leqq H[x, \mu_n^*(x), T^n(\bar{J})] + 2\alpha\varepsilon_k$$
$$\leqq J^*(x) + 3\alpha\varepsilon_k \quad \forall n \geqq k, \quad k \geqq \bar{k}.$$

It follows that $\mu_n^*(x) \in U_k[x, J^*(x) + 3\alpha\varepsilon_k]$ for all $n \geqq k$ and $k \geqq \bar{k}$, and $\{\mu_n^*(x)\}$ has an accumulation point by the compactness of $U_k[x, J^*(x) + 3\alpha\varepsilon_k]$.

(c) If $\mu^*(x)$ is an accumulation point of $\{\mu_n^*(x)\}$ then $\mu^*(x) \in U_k[x, J^*(x) + 3\alpha\varepsilon_k]$ for all $k \geqq \bar{k}$ or equivalently

$$(T_{\mu^*}T^k)(\bar{J})(x) \leqq J^*(x) + 3\alpha\varepsilon_k \quad \forall x \in S, \quad k \geqq \bar{k}.$$

By using (20) we have for all k

$$\|(T_{\mu^*}T^k)(\bar{J}) - T_{\mu^*}(J^*)\| \leqq \alpha\|T^k(\bar{J}) - J^*\| \leqq \alpha\varepsilon_k.$$

Combining the two inequalities above we obtain

$$T_{\mu^*}(J^*)(x) \leqq J^*(x) + 4\alpha\varepsilon_k \quad \forall x \in S, \quad k \geqq \bar{k}.$$

Since $\varepsilon_k \to 0$ [cf. Prop. 1, part (d)] we obtain $T_{\mu^*}(J^*) \leqq J^*$. Using the fact $J^* = T(J^*) \leqq T_{\mu^*}(J^*)$, we obtain $T_{\mu^*}(J^*) = J^*$ which implies, by Proposition 1, that the stationary policy $\{\mu^*, \mu^*, \cdots\}$ is optimal. Q.E.D.

REFERENCES

[1] D. P. BERTSEKAS, *Dynamic Programming and Stochastic Control*, Academic Press, New York, 1976.

[2] ———, *Infinite time reachability of state space regions by using feedback control*, IEEE Trans. Automatic Control, AC-17 (1972), pp. 604–613.

[3] D. BLACKWELL, *Positive dynamic programming*, Proc. 5th Berkeley Symposium on Mathematics, Statistics, and Probability, vol. 1, 1965, pp. 415–418.

[4] E. V. DENARDO, *Contraction mappings in the theory underlying dynamic programming*, SIAM Rev., 9 (1967), pp. 165–177.

[5] J. DUGUNDJI, *Topology*, Allyn and Bacon, Boston, 1968.

[6] K. HINDERER, *Foundations of Non-Stationary Dynamic Programming with Discrete-Time Parameter*, Springer-Verlag, New York, 1970.

[7] D. H. JACOBSON, *Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games*, IEEE Trans. Automatic Control, AC-18 (1973), pp. 124–131.

[8] H. J. KUSHNER, *Introduction to Stochastic Control*, Holt, Rinehart, and Winston, New York, 1971.

[9] A. MAITRA, *Discounted dynamic programming on compact metric spaces*, Sankhyā, Ser. A, 30 (1968), pp. 211–216.

[10] S. SHREVE AND D. P. BERTSEKAS, *Equivalent deterministic and stochastic optimal control problems*, Proc. 1976 IEEE Conference on Decision and Control, Clearwater Beach, Fla., 1976.

[11] R. E. STRAUCH, *Negative dynamic programming*, Ann. Math. Statist., 37 (1966), pp. 871–890.

[12] S. SHREVE, *Dynamic programming in complete separable metric spaces*, Ph.D. thesis, Dept. of Mathematics, Univ. of Illinois, Urbana, Jan. 1977.