

## ALTERNATIVE THEORETICAL FRAMEWORKS FOR FINITE HORIZON DISCRETE-TIME STOCHASTIC OPTIMAL CONTROL\*

STEVEN E. SHREVE† AND DIMITRI P. BERTSEKAS‡

**Abstract.** Stochastic optimal control problems are usually analyzed under one of three types of assumptions: a) Countability assumptions on the underlying probability space—this eliminates all difficulties of measure theoretic nature; b) Semicontinuity assumptions under which the existence of optimal Borel measurable policies can be guaranteed; and c) Borel measurability assumptions under which the existence of  $p$ -optimal or  $p$ - $\epsilon$ -optimal Borel measurable policies can be guaranteed (Blackwell [3], Strauch [31]). In this paper we introduce a general theoretical framework based on outer integration which contains these three models as special cases. Within this framework all known results for finite horizon problems together with some new ones are proved and subsequently specialized. An important new feature of our specialization to the Borel measurable model is the introduction of universally measurable policies. We show that everywhere optimal or nearly optimal policies exist within this class and this enables us to dispense with the notion of  $p$ -optimality.

**1. Introduction.** Consider a stochastic optimal control problem with cost function

$$(1) \quad J = \sum_{k=0}^{N-1} g(x_k, u_k, w_k), \quad N: \text{positive integer or } +\infty,$$

subject to the system equation

$$(2) \quad x_{k+1} = f(x_k, u_k, w_k), \quad k = 0, 1, \dots,$$

where  $x_k, u_k$  are the state and control of the system and  $w_k$  is a random object with probability distribution parameterized by  $x_k$  and  $u_k$ . We wish to choose a policy, that is, a sequence of functions  $\{\mu_k\}$  from the state space  $S$  to the control space  $C$  so that when  $u_k = \mu_k(x_k)$  the expected value of  $J$  is minimized. (A precise definition of the problem will be postponed for later.)

The equation

$$(3) \quad J_{k+1}(x) = \inf_u E\{g(x, u, w) + J_k[f(x, u, w)] | x, u\}$$

with  $J_0(x) \equiv 0$ , and its limiting form

$$(4) \quad J^*(x) = \inf_u E\{g(x, u, w) + J^*[f(x, u, w)] | x, u\}$$

are the Dynamic Programming (DP for short) equations related to the problem above. In the case where  $w$  takes a single value and the problem is deterministic (more generally, where  $w$  can take a countable number of values), the functions in these equations exist in a well-defined mathematical sense and the theory of DP is well developed (see e.g. [1]). When  $w$  can take uncountably many values, acute difficulties arise from the need to impose a proper measure theoretic structure on the problem so that the expected value of the cost  $J$  of (1) and the expected values in (3) and (4) are well defined. A related difficulty stems from the need to balance the measurability restrictions on policies (necessary so that the expected cost corresponding to a policy is

\* Received by the editors April 14, 1977, and in revised form December 23, 1977. This work was supported by the National Science Foundation under Grant No. ENG 74-19332.

† Department of Statistics, University of California, Berkeley, California 94720.

‡ Department of Electrical Engineering and Coordinated Science Lab., University of Illinois, Urbana, Illinois 61801.

well defined) against a desire to admit enough policies to consideration so as to be able to find one which selects at or near the infimum in (3).

The measurability questions described above have been dealt with by a number of authors under the assumption that all spaces underlying the problem are Borel spaces (Borel subsets of complete separable metric spaces). There have been two main approaches. In the first approach, semicontinuity and compactness assumptions are imposed on the data of the problem (see § 5). Under these assumptions, the functions  $J_k$  in (3) can be shown to be semicontinuous and there exists a Borel measurable function  $\mu_k$  such that  $u_k = \mu_k(x)$  selects at or near the infimum in (3) for every  $x$ . For results in this direction see Maitra [15], Schäl [24]–[25], and Freedman [9]. Much of the work in stochastic programming (see Olsen [18]–[20], Rockafellar and Wets [22], [23] and the references quoted therein) also utilizes assumptions of this type. Some of this work employs additional convexity assumptions and is geared toward convex programming type results, i.e., duality and Kuhn–Tucker conditions for optimality, rather than resolution of the measure theoretic questions.

The second approach was introduced by Blackwell [3] and further refined by Strauch [31], Dynkin and Juskevic [8], Hinderer [12] and others. No assumptions other than Borel measurability of the data of the problem are made, and admissible policies are required to be Borel measurable. Under these conditions it is possible to prove the universal measurability of the optimal cost function and the existence for every  $\varepsilon > 0$  and probability measure  $p$  on  $S$  of a  $p$ - $\varepsilon$ -optimal policy (Strauch [31, Thms 7.1 and 8.1]). A  $p$ - $\varepsilon$ -optimal policy is one which leads to a cost which differs from the optimal cost by less than  $\varepsilon$  for  $p$ -almost every initial state. Even over a finite horizon the optimal cost function need not be Borel measurable, and there need not exist an everywhere  $\varepsilon$ -optimal policy (Blackwell [3, Example 2]). The difficulty arises from the inability to choose a Borel measurable function  $\mu_k: S \rightarrow C$  which nearly achieves the infimum in (3) uniformly in  $x$ . The nonexistence of such a function interferes with the construction of optimal policies via the DP algorithm (3), since one must first determine, at each stage  $k$ , a measure  $p_k$  with respect to which it is satisfactory to nearly achieve the infimum in (3) for  $p_k$ -almost every  $x$ . The difficulties in constructing nearly optimal policies over an infinite horizon are more acute. Furthermore, from an applications point of view, a  $p$ - $\varepsilon$ -optimal policy, even if it can be constructed, is a much less appealing object than an everywhere  $\varepsilon$ -optimal policy, since in many situations the distribution  $p$  is unknown or may change when the system is operated repetitively, in which case a new  $p$ - $\varepsilon$ -optimal policy must be computed.

In view of the undesirable features of  $p$ -optimality, Blackwell, Freedman and Orkin [4] have considered analytically measurable policies—a class that properly contains Borel measurable policies (see § 6). Their work deals with a special type of problem, that of minimization when the cost per stage is nonpositive. They show that a history remembering policy which is everywhere  $\varepsilon$ -optimal exists, and if the optimal cost functions  $J_k^*$ ,  $k = 1, \dots, N$ , are everywhere finite, this policy can be taken to be Markov. We relax the assumption of a nonpositive cost per stage and show the existence in Corollary 5.1 of an analytically measurable  $\varepsilon$ -optimal Markov policy under the assumption that the functions  $J_k^*$ ,  $k = 1, \dots, N$ , are everywhere finite. However we have been unable to show the strongest possible existence results for finite horizon problems within an analytically measurable policies framework (compare Corollaries 5.1 and 5.2). For this reason we have extended the class of admissible policies to include all universally measurable policies (a class properly containing the analytically measurable policies). A key fact here is that the composition of two universally measurable functions is universally measurable, while the

composition of two analytically measurable functions need not be analytically measurable [2]. In this paper we admit only Markov nonrandomized policies and, within this framework, cannot prove the result of Blackwell, Freedman and Orkin mentioned earlier for the case where  $J_k^*(x)$  can be  $-\infty$  for some  $x$  and  $k$ . It is shown elsewhere [2], [29], however, that under the assumption of Corollary 5.2(b), there exists for every  $\varepsilon > 0$  a nonrandomized semi-Markov and a randomized Markov universally measurable  $\varepsilon$ -optimal policy. We do not know whether such a policy can be taken to be analytically measurable rather than universally measurable. The fact mentioned earlier relating to composition of two analytically measurable functions interferes with the constructions involved in the proofs of [29].

The present paper has two main objectives. The first is to provide a general framework for finite horizon stochastic optimal control that includes as special cases the formulations described earlier. The second is to demonstrate that when universally measurable policies are admitted in the Borel space framework of Blackwell, then all basic results for stochastic problems can be shown to hold in a form that is as strong as for problems where measurability questions are of no essential concern. In particular, the existence of *everywhere*  $\varepsilon$ -optimal policies is assured as opposed to policies which are  $\varepsilon$ -optimal  $p$ -almost everywhere. Thus the notion of  $p$ -optimality can be dispensed with.

The paper is organized as follows. Section 2 formulates a general stochastic optimal control problem without any topological assumptions. The formulation is based on a notion of outer integration developed in Appendix A. The main results regarding the validity of the DP algorithm and the existence of optimal and nearly optimal policies are provided in § 3. These include all results known for special cases together with a new result [Proposition 1(b)] relating to the existence of a sequence of policies exhibiting what is referred to as  $\{\varepsilon_n\}$  dominated convergence to optimality. The results of § 3 are applied to special cases in § 4 (model without topological assumptions) and in § 5 (Borel space models with semicontinuity assumptions). Slight extensions of results by Freedman [9] are given in Corollaries 4.1 and 4.2. Section 6 is devoted to general Borel space models. We consider both analytically and universally measurable policies and prove an extended version of a measurable selection theorem by Brown and Purves [5]. Using this theorem we show that all the results of § 3 carry over to the Borel space model when universally measurable policies are allowed.

We note that some of the ideas and analysis in this paper (particularly the employment of universally measurable policies) have infinite horizon and imperfect state information counterparts described elsewhere [2], [28], [29]. Also, this paper considers exclusively nonrandomized Markov policies. Existence results relating to randomized and semi-Markov policies may be found in [2], [27], [29].

**2. Problem formulation.** Our notation will be as follows. For a set  $X$  we denote by  $F_X$  the set of all functions  $J: X \rightarrow [-\infty, +\infty]$ . For  $J_1, J_2 \in F_X$  we write  $J_1 = J_2$  if  $J_1(x) = J_2(x) \forall x \in X$ , and  $J_1 \leq J_2$  if  $J_1(x) \leq J_2(x) \forall x \in X$ . If  $J(x) \geq \varepsilon (J(x) \leq \varepsilon) \forall x \in X$ , where  $\varepsilon$  is a scalar, we write  $J \geq \varepsilon (J \leq \varepsilon)$ . If a sequence  $\{J_k\} \subset F_X$  increases (decreases) monotonically to  $J \in F_X$ , we write  $J_k \uparrow J (J_k \downarrow J)$ . If  $\{J_k\}$  converges pointwise to  $J$  we write  $J_k \rightarrow J$ . If  $J_1, J_2 \in F_X$ ,  $\varepsilon$  is a scalar, and  $J_1(x) \leq J_2(x) + \varepsilon \forall x \in X$ , we write  $J_1 \leq J_2 + \varepsilon$ . We adopt the usual conventions regarding ordering and arithmetic in the set of extended real numbers  $[-\infty, +\infty]$ , except that we take

$$-\infty + \infty = +\infty - \infty = +\infty.$$

The Cartesian product of sets  $A_1, A_2, \dots, A_n$  is denoted by  $A_1 A_2 \cdots A_n$ . If  $X$  and  $Y$

are sets, then  $\text{proj}_X$  is the projection mapping from  $XY$  to  $X$ . If  $E$  is a subset of some space  $X$ , we denote by  $\chi_E$  the indicator function of  $E$  [ $\chi_E(x) = 1$  if  $x \in E$ ,  $\chi_E(x) = 0$  if  $x \notin E$ ]. For any function  $f: X \rightarrow [-\infty, +\infty]$ , where  $X$  is some space, we use the notation  $f^+(x) = \max\{0, f(x)\}$ ,  $f^-(x) = \max\{0, -f(x)\}$ . The infimum over the empty set is taken by convention to be  $+\infty$  ( $\inf \emptyset = +\infty$ ).

The stochastic optimal control model we consider consists of the eight elements listed below:

$S$ —*State space*. A nonempty set.

$C$ —*Control space*. A nonempty set.

$(W, \mathcal{F})$ —*Disturbance space*. A measurable space.

$p(dw|x, u)$ —*Disturbance kernel*. For fixed  $(x, u) \in SC$ ,  $p(\cdot|x, u)$  is a probability measure on  $(W, \mathcal{F})$ .

$f(x, u, w)$ —*System function*. A mapping from  $SCW$  to  $S$ .

$g(x, u, w)$ —*One-stage cost function*. A mapping from  $SCW$  to  $[-\infty, +\infty]$ .

$M$ —*Control function space*. A nonempty set of mappings from  $S$  to  $C$ .

$N$ —*Horizon*. A positive integer.

The model is stationary in that the data does not change from one stage to the next. There is no essential loss of generality in this assumption, since a nonstationary model can be reduced to a stationary one by state augmentation ([24, § 8], [1, § 6.7]). We impose no assumptions for the time being on the set of control functions  $M$ . However, specific results will assume explicitly or implicitly various conditions on  $M$ , and in fact our line of analysis is geared toward demonstrating the type of properties of  $M$  that are essential for specific results to hold. In particular special cases the set  $M$  could be as large as the set of all functions  $\mu: S \rightarrow C$  or as restricted as the set of all linear functions  $\mu: S \rightarrow C$  ( $S, C$  assumed to be linear spaces). We shall use the letter  $x$  to represent an element of  $S$ , and the letter  $u$  to represent an element of  $C$ . Denote by  $\Pi_N$  the Cartesian product of  $N$  copies of  $M$  and define

$$(5) \quad \Gamma = \{(x, u): x \in S, u = \mu(x) \text{ for some } \mu \in M\}.$$

We denote by  $\Gamma_x$  the cross-section  $\{u: (x, u) \in \Gamma\}$ . We refer to an element of  $\Pi_N$  as a *policy*.

We have in mind a system operating as follows. A policy  $\pi = (\mu_0, \dots, \mu_{N-1}) \in \Pi_N$  is chosen. The system begins in some initial state  $x_0$  and subsequent states are specified by the system equation

$$(6) \quad x_{k+1} = f(x_k, u_k, w_k), \quad k = 0, \dots, N-2,$$

where

$$(7) \quad u_k = \mu_k(x_k), \quad k = 0, \dots, N-1,$$

and  $w_k$  is random with distribution  $p(dw_k|x_k, u_k)$ . The cost incurred at each stage of the operation is  $g(x_k, u_k, w_k)$ , so the total cost is

$$\sum_{k=0}^{N-1} g(x_k, u_k, w_k).$$

The expected total cost corresponding to the policy  $\pi$  is obtained by taking the expectation of the total cost with respect to the appropriate probability measure. If the integrals can be defined, this can be represented by

$$(8) \quad \int_W \cdots \int_W \sum_{k=0}^{N-1} g(x_k, u_k, w_k) p(dw_{N-1}|x_{N-1}, u_{N-1}) \cdots p(dw_0|x_0, u_0),$$

where (6) and (7) hold. But we have not yet imposed sufficient structure for the integrals to be defined, so we specify the cost corresponding to each policy and initial state by means of outer integration.

Given a probability space  $(X, \mathcal{B}, p)$  and a function  $f \in F_X$  with  $f \geq 0$ , the outer integral of  $f$  with respect to  $p$  is defined in Appendix A as

$$\int^* f dp = \inf \left\{ \int g dp : f \leq g, g \text{ is } \mathcal{B}\text{-measurable} \right\}.$$

Given an arbitrary  $f \in F_X$ , we define its *outer integral with respect to  $p$*  by

$$(9) \quad \int^* f dp = \int^* f^+ dp - \int^* f^- dp.$$

Since we take  $\infty - \infty$  to be  $+\infty$ , the outer integral of  $f$  is defined for every  $f \in F_X$ . If  $f$  is measurable with respect to the  $\sigma$ -algebra  $\mathcal{B}$ , we write  $\int f dp$  in place of  $\int^* f dp$ . Note that for such an  $f$  and  $\alpha \in (-\infty, +\infty]$ , we have

$$(10) \quad \alpha + \int f dp = \int (\alpha + f) dp.$$

For each  $\mu \in M$ , we define the mapping  $T_\mu : F_S \rightarrow F_S$  by

$$(11) \quad T_\mu(J)(x) = \int^* \{g[x, \mu(x), w] + J[f(x, \mu(x), w)]\} p(dw | x, \mu(x)) \quad \forall J \in F_S, \quad x \in S.$$

We also define the mapping  $T : F_S \rightarrow F_S$  by

$$(12) \quad T(J)(x) = \inf_{u \in \Gamma_x} \int^* \{g(x, u, w) + J[f(x, u, w)]\} p(dw | x, u) \quad \forall J \in F_S, \quad x \in S.$$

Note that since we have by definition  $\Gamma_x = \{u \in C : u = \mu(x) \text{ for some } \mu \in M\}$ , it follows that (12) can be replaced by

$$(13) \quad T(J)(x) = \inf_{\mu \in M} T_\mu(J)(x) \quad \forall x \in S.$$

For any  $\mu_1, \dots, \mu_k \in M$ , we denote by  $(T_{\mu_1} \dots T_{\mu_k})$  the composition of the mappings  $T_{\mu_1}, \dots, T_{\mu_k}$ . Similarly, we denote by  $T^k$  the composition of  $T$  with itself  $k$  times. For convenience we also use  $T^0$  to denote the identity mapping on  $F_S$ .

The *cost function corresponding to the policy  $\pi$*  is defined by

$$(14) \quad J_{N,\pi} = (T_{\mu_0} \dots T_{\mu_{N-1}})(J_0),$$

where  $J_0(x) = 0$  for every  $x \in S$ . The *optimal cost function* is given by

$$(15) \quad J_N^*(x) = \inf_{\pi \in \Pi_N} J_{N,\pi}(x) \quad \forall x \in S.$$

If measurability assumptions are made so that reference to the outer integral is unnecessary and finiteness assumptions are imposed to allow the interchange of summation and integration, then (14) reduces to the more traditional definition of expected cost corresponding to a policy given by (8). One type of measurability assumption is to assume  $W$  is countable and  $\mathcal{F}$  is the power set of  $W$ , so that integration reduces to summation. No measure structure need be imposed on  $S$  and  $C$ . A less trivial set of assumptions is obtained by letting  $S$  have a  $\sigma$ -algebra  $\mathcal{S}$ ,  $C$  have a  $\sigma$ -algebra  $\mathcal{C}$ , and assuming  $f$  is  $(\mathcal{S}\mathcal{C}\mathcal{F}, \mathcal{S})$  measurable and  $g$  is  $(\mathcal{S}\mathcal{C}\mathcal{F}, \mathcal{B})$  measurable,

where  $\mathcal{S}\mathcal{C}\mathcal{F}$  is the product  $\sigma$ -algebra and  $\mathcal{B}$  is the Borel  $\sigma$ -algebra in  $[-\infty, +\infty]$ . One must also assume  $p(B|x, u)$  is  $\mathcal{S}\mathcal{C}$  measurable for fixed  $B \in \mathcal{F}$ . If all mappings in  $M$  are  $(\mathcal{S}, \mathcal{C})$  measurable, then whenever  $\mu \in M$  and  $J$  is  $(\mathcal{S}, \mathcal{B})$  measurable,  $T_\mu(J)$  is also. Despite these assumptions, it may still occur that (8) and (14) do not agree, but if the possibility of  $+\infty - \infty$  occurring is limited, agreement can be guaranteed. This can be accomplished by requiring that  $g(x, u, w) > -\infty \forall x \in S, u \in C, w \in W$  (see (10)), or by requiring that for each  $\pi = (\mu_0, \dots, \mu_{N-1}) \in \Pi_N$  and each  $x_0 \in S$ ,

$$(16) \quad \int_W \dots \int_W \sum_{k=0}^{N-1} g^+(x_k, u_k, w_k) p(dw_{N-1}|x_{N-1}, u_{N-1}) \dots p(dw_0|x_0, u_0) < +\infty,$$

where (6) and (7) hold. If (16) holds for each  $\pi \in \Pi_N$  and  $x_0 \in S$ , one can in fact show by Fubini's theorem that

$$J_{N,\pi}(x) = E_{(\pi,x)} \left\{ \sum_{k=0}^{N-1} g(x_k, u_k, w_k) \right\},$$

where (6) and (7) hold and the expectation is with respect to the product measure on  $W \dots W$  generated by  $\pi$  from  $x_0 = x$ . This is also the case if for each  $\pi \in \Pi_N$  and  $x_0 \in S$ , (16) holds with  $g^+$  replaced by  $g^-$ .

We now introduce various notions of optimality. Let  $x \in S$  and  $\epsilon > 0$  be given. A policy  $\pi \in \Pi_N$  is  $\epsilon$ -optimal at  $x$  if

$$J_{N,\pi}(x) \leq \begin{cases} J_N^*(x) + \epsilon & \text{if } J_N^*(x) > -\infty, \\ -1/\epsilon & \text{if } J_N^*(x) = -\infty. \end{cases}$$

If  $\pi \in \Pi_N$  is  $\epsilon$ -optimal at every  $x \in S$ , we say  $\pi$  is  $\epsilon$ -optimal. A policy  $\pi \in \Pi_N$  is optimal at  $x$  if

$$J_{N,\pi}(x) = J_N^*(x).$$

If  $\pi \in \Pi_N$  is optimal at every  $x \in S$ , we say  $\pi$  is optimal.

Let  $\{\epsilon_n\}$  be a sequence of positive numbers with  $\epsilon_n \downarrow 0$ . A sequence of policies  $\{\pi_n\} \subset \Pi_N$  is said to exhibit  $\{\epsilon_n\}$  dominated convergence to optimality if

$$J_{N,\pi_n} \rightarrow J_N^*, \\ J_{N,\pi_n}(x) \leq J_N^*(x) + \epsilon_n \quad \text{if } J_N^*(x) > -\infty,$$

and

$$J_{N,\pi_n}(x) \leq J_{N,\pi_{n-1}}(x) + \epsilon_n \quad \text{if } J_N^*(x) = -\infty.$$

If  $\{\pi_n\}$  exhibits  $\{\epsilon_n\}$  dominated convergence to optimality and  $J_N^*(x) > -\infty$  for every  $x \in S$ , then by definition  $\pi_n$  is  $\epsilon_n$ -optimal.

**3. Main results.** For our results we shall need some regularity assumptions on the model. We list them here for convenience and shall refer to them explicitly when we wish to include them in the hypotheses of a proposition.

*Assumption A:* There is a subset  $F$  of  $F_S$  such that  $J_0 \in F$  and whenever  $J \in F$ , then  $T(J) \in F$ .

*Assumption B:* If  $J \in F$  as given in Assumption A and  $\epsilon > 0$ , then there exists  $\mu_\epsilon \in M$  such that

$$T_{\mu_\epsilon}(J)(x) \leq \begin{cases} T(J)(x) + \epsilon & \text{if } T(J)(x) > -\infty, \\ -1/\epsilon & \text{if } T(J)(x) = -\infty. \end{cases}$$

*Assumption C:* If  $J \in F$  as given in Assumption A and the infimum in (12) is achieved for every  $x \in S$ , then there exists  $\mu \in M$  such that

$$T_\mu(J)(x) = T(J)(x) \quad \forall x \in S.$$

*Assumption D:* For  $J \in F$  as given in Assumption A, define

$$A(J) = \{(x, u) \in \Gamma : p^*(\{w : J[f(x, u, w)] = -\infty\} | x, u) > 0\},$$

where  $p^*(\cdot | x, u)$  represents outer measure. For each  $J \in F$ , there is a  $\mu_J \in M$  such that  $(x, \mu_J(x)) \in A(J)$  whenever  $x \in \text{proj}_S A(J)$ . Furthermore if  $\mu \in M$  and  $\mu_J$  is as above, then  $\hat{\mu}$  defined by

$$\hat{\mu}(x) = \begin{cases} \mu_J(x) & \text{if } x \in \text{proj}_S A(J), \\ \mu(x) & \text{otherwise,} \end{cases}$$

is in  $M$ . Also if  $J \in F$  and  $\mu_1, \mu_2 \in M$ , then  $\bar{\mu}$  defined by

$$\bar{\mu}(x) = \begin{cases} \mu_1(x) & \text{if } T_{\mu_1}(J)(x) \leq T_{\mu_2}(J)(x), \\ \mu_2(x) & \text{otherwise,} \end{cases}$$

is in  $M$ .

Assumption A will be used to show properties of  $T^N(J_0)$ , which is often identical to  $J_N^*$ . By choosing  $F$  to be the set of functions having measurability or continuity properties and showing that Assumption A holds, we can immediately deduce properties of  $T^N(J_0)$ . We will find it very important to be able to choose a control function which nearly achieves the infimum in the definition of  $T(J)$  for  $J \in F$ . This is the condition given in Assumption B. Assumption C states that  $M$  is rich enough to allow exact selection of this infimum if it is achieved. This is necessary in order to construct an optimal policy. Assumption D states that  $M$  contains enough functions to allow certain constructions necessary for the proof of Proposition 1 below.

The following lemma provides some properties of the mappings  $T_\mu$  and  $T$  that we shall need.

LEMMA 1. (a) If  $J_1, J_2 \in F_S$  and  $J_1 \leq J_2$ , then  $T(J_1) \leq T(J_2)$ , and for all  $\mu \in M$   $T_\mu(J_1) \leq T_\mu(J_2)$ .

(b) If  $J_1, J_2 \in F_S$  and  $J_2 \leq J_1 + \varepsilon$  for some  $\varepsilon > 0$ , then  $T_\mu(J_2) \leq T_\mu(J_1) + 2\varepsilon$  for all  $\mu \in M$ .

(c) If  $J_1, J_2 \in F_S$  and for some  $\varepsilon > 0$  we have

$$(17) \quad J_2(x) \leq J_1(x) + \varepsilon \quad \text{if } J_1(x) > -\infty,$$

then for all  $\mu \in M$

$$(18) \quad T_\mu(J_2)(x) \leq T_\mu(J_1)(x) + 2\varepsilon \quad \text{if } T_\mu(J_1)(x) > -\infty.$$

*Proof.* (a) and (b) follow directly from Lemma A.3(a), (b), so we concentrate on proving (c). Let  $x \in S$  be such that  $T_\mu(J_1)(x) > -\infty$ . Then either  $T_\mu(J_1)(x) = +\infty$ , in which case (18) is trivial or else from Lemma A.3(g) we obtain

$$(19) \quad p^*(A | x, \mu(x)) = 0,$$

where  $A = \{w | J_1[f(x, \mu(x), w)] = -\infty\}$ . From (17) we obtain for all  $w \notin A$

$$g[x, \mu(x), w] + J_2[f(x, \mu(x), w)] \leq g[x, \mu(x), w] + J_1[f(x, \mu(x), w)] + \varepsilon$$

and (18) follows from Lemma A.3(b), (e) and (19). Q.E.D.

We now arrive at our first main result:

PROPOSITION 1. *Let Assumptions A and B hold.*

(a) *If  $J_k^*(x) > -\infty$  for all  $x \in S$  and  $k = 1, 2, \dots, N$ , then*

$$J_N^* = T^N(J_0),$$

and for every  $\varepsilon > 0$  there exists an  $\varepsilon$ -optimal policy.

(b) *If Assumption D holds and  $J_{k,\pi}(x) < +\infty$  for all  $x \in S$ ,  $\pi \in \Pi_k$  and  $k = 1, 2, \dots, N$ , then*

$$J_N^* = T^N(J_0),$$

and for every sequence  $\{\varepsilon_n\}$  with  $\varepsilon_n > 0$ ,  $n = 1, 2, \dots$ ,  $\varepsilon_n \downarrow 0$ , there exists a sequence of policies exhibiting  $\{\varepsilon_n\}$  dominated convergence to optimality.

*Proof.* (a) For any  $\pi = (\mu_0, \dots, \mu_{k-1}) \in \Pi_k$ ,

$$\begin{aligned} J_{k,\pi} &= (T_{\mu_0} \cdots T_{\mu_{k-2}} T_{\mu_{k-1}})(J_0) \\ &\cong (T_{\mu_0} \cdots T_{\mu_{k-2}} T)(J_0) && \text{(by Lemma 1(a))} \\ &\cong \cdots \\ &\cong T^k(J_0). \end{aligned}$$

Hence

$$(20) \quad J_k^* \cong T^k(J_0), \quad k = 1, 2, \dots.$$

We conclude the proof by induction. Assumption B and (20) guarantee that when  $N = 1$ , (a) holds. Suppose (a) holds for  $N - 1$ . Then for  $\varepsilon > 0$  there exists  $\pi \in \Pi_{N-1}$  such that

$$J_{N-1,\pi} \cong J_{N-1}^* + \varepsilon/4.$$

From the induction assumption and Lemma 1(a), (b),

$$\begin{aligned} T^N(J_0) &= T(J_{N-1}^*) \\ &\cong T(J_{N-1,\pi} - \varepsilon/4) \\ &\cong T(J_{N-1,\pi}) - \varepsilon/2 \\ &= \inf_{\mu} T_{\mu}(J_{N-1,\pi}) - \varepsilon/2 \\ &\cong J_N^* - \varepsilon/2. \end{aligned}$$

Combining this with (20) we obtain

$$J_N^* = T^N(J_0).$$

Use Assumptions A and B to find  $\mu \in M$  for which

$$T_{\mu}(J_{N-1}^*) \cong T(J_{N-1}^*) + \varepsilon/2.$$

With  $\pi$  as above, we have

$$\begin{aligned} J_{N,(\mu,\pi)} &= T_{\mu}(J_{N-1,\pi}) \\ &\cong T_{\mu}(J_{N-1}^*) + \varepsilon/2 \\ &\cong T(J_{N-1}^*) + \varepsilon \\ &= J_N^* + \varepsilon, \end{aligned}$$



so  $(\mu, \pi)$  is an  $\varepsilon$ -optimal  $N$ -stage policy.

(b) The proof proceeds by induction. Let  $\{\varepsilon_n\}$  be a sequence with  $\varepsilon_n > 0, \varepsilon_n \downarrow 0$ . For  $N = 1$ , Assumptions A and B imply the existence of a sequence of policies  $\pi_n = (\mu_0^n) \in \Pi_1$  for which

$$(21) \quad T_{\mu_0^n}(J_0)(x) \cong \begin{cases} T(J_0)(x) + \varepsilon_n & \text{if } T(J_0)(x) > -\infty, \\ -1/\varepsilon_n & \text{if } T(J_0)(x) = -\infty. \end{cases}$$

By the last part of Assumption D we can assume without loss of generality that

$$(22) \quad T_{\mu_0^n}(J_0)(x) \downarrow -\infty \quad \text{if } T(J_0)(x) = -\infty.$$

This implies  $J_1^* \cong T(J_0)$ , which together with (20) establishes

$$(23) \quad J_1^* = T(J_0).$$

From (21)–(23) we see that  $\{\pi_n\}$  exhibits  $\{\varepsilon_n\}$  dominated convergence to optimality.

Suppose the result holds for  $N - 1$ . Let  $\pi_n = (\mu_1^n, \dots, \mu_{N-1}^n)$  be a sequence of  $(N - 1)$ -stage policies exhibiting  $\{\varepsilon_n/4\}$  dominated convergence to optimality, i.e.,

$$(24) \quad J_{N-1, \pi_n} \rightarrow J_{N-1}^*,$$

$$(25) \quad J_{N-1, \pi_n}(x) \cong J_{N-1}^*(x) + \varepsilon_n/4 \quad \text{if } J_{N-1}^*(x) > -\infty,$$

$$(26) \quad J_{N-1, \pi_n}(x) \cong J_{N-1, \pi_{n-1}}(x) + \varepsilon_n/4 \quad \text{if } J_{N-1}^*(x) = -\infty.$$

We assume without loss of generality that  $\sum_{n=1}^\infty \varepsilon_n < \infty$ . By the induction hypothesis and Assumption A

$$(27) \quad J_{N-1}^* = T^{N-1}(J_0) \in F,$$

so by Assumption B there is a sequence  $\{\mu^n\} \subset M$  such that

$$(28) \quad T_{\mu^n}(J_{N-1}^*)(x) \cong \begin{cases} T^N(J_0)(x) + \varepsilon_n/2 & \text{if } T^N(J_0)(x) > -\infty, \\ -2/\varepsilon_n & \text{if } T^N(J_0)(x) = -\infty. \end{cases}$$

By the last part of Assumption D we can assume without loss of generality that

$$(29) \quad T_{\mu^n}(J_{N-1}^*) \cong T_{\mu^{n-1}}(J_{N-1}^*), \quad n = 2, 3, \dots$$

By Assumption D there is a  $\mu \in M$  such that  $(x, \mu(x)) \in A(J_{N-1}^*)$  whenever  $x \in \text{proj}_S A(J_{N-1}^*)$ , i.e.,

$$(30) \quad p^* (\{w : J_{N-1}^* [f(x, \mu(x), w)] = -\infty\} | x, \mu(x)) > 0$$

whenever for some  $u \in \Gamma_x$

$$p^* (\{w : J_{N-1}^* [f(x, u, w)] = -\infty\} | x, u) > 0.$$

Define

$$\hat{\mu}^n(x) = \begin{cases} \mu(x) & \text{if } x \in \text{proj}_S A(J_{N-1}^*), \\ \mu^n(x) & \text{otherwise.} \end{cases}$$

Then  $\{\hat{\mu}^n\} \subset M$  by Assumption D and

$$\hat{\pi}_n = (\hat{\mu}^n, \pi_n) \in \Pi_N.$$

For  $x \in \text{proj}_S A(J_{N-1}^*)$ , we have

$$\begin{aligned} \limsup_{n \rightarrow \infty} J_{N, \hat{\pi}_n}(x) &= \limsup_{n \rightarrow \infty} T_\mu(J_{N-1, \pi_n})(x) \\ &= \limsup_{n \rightarrow \infty} \int^* \{g[x, \mu(x), w] + J_{N-1, \pi_n}[f(x, \mu(x), w)]\} p(dw | x, \mu(x)) \\ &= \int^* \{g[x, \mu(x), w] + J_{N-1}^*[f(x, \mu(x), w)]\} p(dw | x, \mu(x)) \end{aligned}$$

by (24)–(26), the fact that  $J_{N, \hat{\pi}_1}(x) < +\infty$  for every  $x \in S$ , and Corollary A.1.1. Relation (30) and Lemma A.3(g) imply that  $T_\mu(J_{N-1}^*)(x) = \pm\infty$ . But  $T_\mu(J_{N-1}^*)(x) \leq J_{N, \hat{\pi}_1}(x) < +\infty$ , so

$$\begin{aligned} \limsup_{n \rightarrow \infty} J_{N, \hat{\pi}_n}(x) &= T_\mu(J_{N-1}^*)(x) \\ (31) \qquad \qquad \qquad &= -\infty \\ &\leq T^N(J_0)(x). \end{aligned}$$

For  $x \notin \text{proj}_S A(J_{N-1}^*)$ , we have for each  $u \in \Gamma_x$

$$(32) \qquad p^* (\{w : J_{N-1}^*[f(x, u, w)] = -\infty\} | x, u) = 0.$$

If  $u = \mu^n(x)$  satisfies (32), then denoting  $E = \{w : J_{N-1}^*[f(x, u, w)] = -\infty\}$ , we have from Lemma A.3(e)

$$\begin{aligned} J_{N, \hat{\pi}_n}(x) &= T_{\mu^n}(J_{N-1, \pi_n})(x) \\ &= \int^* \chi_{W-E}(w) \{g[x, \mu^n(x), w] + J_{N-1, \pi_n}[f(x, \mu^n(x), w)]\} p(dw | x, \mu^n(x)) \\ (33) \qquad &\cong \int^* \chi_{W-E}(w) \{g[x, \mu^n(x), w] + J_{N-1}^*[f(x, \mu^n(x), w)]\} p(dw | x, \mu^n(x)) + \varepsilon_n/2 \\ &\qquad \qquad \qquad \text{(by (25) and Lemma A.3(b))} \\ &= T_{\mu^n}(J_{N-1}^*)(x) + \varepsilon_n/2. \end{aligned}$$

Inequality (33) implies for  $x \notin \text{proj}_S A(J_{N-1}^*)$

$$\begin{aligned} \limsup_{n \rightarrow \infty} J_{N, \hat{\pi}_n}(x) &\leq \limsup_{n \rightarrow \infty} T_{\mu^n}(J_{N-1}^*)(x) \\ (34) \qquad \qquad &\leq T^N(J_0)(x) \quad \text{by (28)}. \end{aligned}$$

Combining (31) and (34) we have

$$(35) \qquad \limsup_{n \rightarrow \infty} J_{N, \hat{\pi}_n}(x) \leq T^N(J_0)(x)$$

for every  $x \in S$ , and this proves

$$(36) \qquad J_N^* \leq T^N(J_0).$$

Combining (36) and (20), we obtain  $J_N^* = T^N(J_0)$  and (35) can be replaced by

$$(37) \quad \lim_{n \rightarrow \infty} J_{N, \hat{\pi}_n}(x) = J_N^*(x)$$

for every  $x \in S$ .

To see that the convergence to optimality given in (37) is  $\{\varepsilon_n\}$  dominated, note that if  $T^N(J_0)(x) > -\infty$ , then  $T_{\hat{\mu}^n}(J_{N-1}^*)(x) > -\infty$  for every  $n$ . By Lemma 1(c) and (25)

$$(38) \quad T_{\hat{\mu}^n}(J_{N-1, \pi_n})(x) \leq T_{\hat{\mu}^n}(J_{N-1}^*)(x) + \varepsilon_n/2 \quad \text{if } T^N(J_0)(x) > -\infty.$$

If  $x \notin \text{proj}_S A(J_{N-1}^*)$ , then

$$(39) \quad \begin{aligned} T_{\hat{\mu}^n}(J_{N-1}^*)(x) &= T_{\mu^n}(J_{N-1}^*)(x) \\ &\leq T^N(J_0)(x) + \varepsilon_n/2 \quad \text{if } T^N(J_0)(x) > -\infty \quad \text{by (28)} \\ &= J_N^*(x) + \varepsilon_n/2 \quad \text{if } J_N^*(x) > -\infty \quad \text{by the fact } T^N(J_0) = J_N^*. \end{aligned}$$

Combining (38) and (39), we have for  $x \notin \text{proj}_S A(J_{N-1}^*)$

$$(40) \quad J_{N, \hat{\pi}_n}(x) \leq J_N^*(x) + \varepsilon_n \quad \text{if } J_N^*(x) > -\infty.$$

If  $x \in \text{proj}_S A(J_{N-1}^*)$ , then it is clear from (31) that

$$J_N^*(x) = T(J_{N-1}^*)(x) = -\infty,$$

and so (40) is true for all  $x \in S$ . If  $x \notin \text{proj}_S A(J_{N-1}^*)$ , (33) and (29) can be used to show for  $n \geq 2$

$$\begin{aligned} J_{N, \hat{\pi}_n}(x) &\leq T_{\mu^n}(J_{N-1}^*)(x) + \varepsilon_n/2 \\ &\leq T_{\mu^{n-1}}(J_{N-1}^*)(x) + \varepsilon_n/2 \\ &\leq T_{\mu^{n-1}}(J_{N-1, \pi_{n-1}})(x) + \varepsilon_n/2 \\ &= J_{N, \hat{\pi}_{n-1}}(x) + \varepsilon_n/2. \end{aligned}$$

If  $x \in \text{proj}_S A(J_{N-1}^*)$ ,

$$\begin{aligned} J_{N, \hat{\pi}_n}(x) &= \int^* \{g[x, \mu(x), w] + J_{N-1, \pi_n}[f(x, \mu(x), w)]\} p(dw | x, \mu(x)) \\ &\leq \int^* \{g[x, \mu(x), w] + J_{N-1, \pi_{n-1}}[f(x, \mu(x), w)]\} p(dw | x, \mu(x)) + \varepsilon_n/2 \\ &\hspace{15em} \text{(by (25), (26) and Lemma A.3(b))} \\ &= J_{N, \hat{\pi}_{n-1}}(x) + \varepsilon_n/2, \end{aligned}$$

so for every  $x \in S$

$$J_{N, \hat{\pi}_n}(x) \leq J_{N, \hat{\pi}_{n-1}}(x) + \varepsilon_n/2. \quad \text{Q.E.D.}$$

To see that an assumption such as  $J_k^*(x) > -\infty$  or  $J_{k, \pi}(x) < +\infty$  for all  $x, \pi$  and  $k$  is necessary in order for  $J_N^* = T^N(J_0)$  to hold, consider the following example:

*Example.* Let  $N = 2$ ,  $S = \{\alpha, \beta\}$ ,  $C = (-\infty, +\infty)$ ,  $W = \{1, 2, \dots\}$ ,  $M = \{\mu : \mu(\alpha), \mu(\beta) \in (-\infty, +\infty)\}$ ,  $p(w = k | x, u) = 1/(k^2 \sum_{j=1}^{\infty} 1/j^2)$ ,  $k = 1, 2, \dots$ ,  $f(\alpha, u, w) = f(\beta, u, w) = \beta \ \forall u \in C, w \in W$ ,  $g(\alpha, u, w) = w$ ,  $g(\beta, u, w) = u, \ \forall u \in C, w \in W$ . Here there are two states,  $\alpha$  and  $\beta$ , and we always have  $x_1 = \beta$  so that the cost of the second stage is  $\mu_1(\beta)$  and can be made arbitrarily small. On the other hand,

$\int w dp = +\infty$ , so that  $J_{2,\pi}(\alpha) = \int \{g[\alpha, \mu_0(\alpha), w] + \mu_1(\beta)\} dp = \int \{w + \mu_1(\beta)\} dp = +\infty$  for all  $\pi \in \Pi_2$ . We have by a straightforward calculation  $J_2^*(\alpha) = +\infty$ ,  $J_2^*(\beta) = -\infty$ , while  $T^2(J_0)(\alpha) = -\infty$ ,  $T^2(J_0)(\beta) = -\infty$ .

Despite the need for various assumptions in order to show the equality  $J_N^* = T^N(J_0)$ , the following result, which establishes the validity of the DP algorithm as a means for obtaining optimal policies, requires none of the assumptions of Proposition 1. We say that a policy  $\pi = (\mu_0, \dots, \mu_{N-1}) \in \Pi_N$  is *uniformly N-stage optimal* if for  $k = 1, \dots, N$ ,  $\pi^k = (\mu_{N-k}, \dots, \mu_{N-1})$  is optimal in the k-stage problem of minimizing  $J_{k,\pi}$  over  $\Pi_k$ .

PROPOSITION 2. (a) *If there exists a uniformly N-stage optimal policy then*

$$(41) \quad J_k^* = T^k(J_0), \quad \forall k = 1, \dots, N.$$

(b) *A policy  $\pi = (\mu_0, \dots, \mu_{N-1}) \in \Pi_N$  is uniformly N-stage optimal if and only if*

$$(42) \quad (T_{\mu_{N-k}} T^{k-1})(J_0) = T^k(J_0), \quad \forall k = 1, \dots, N.$$

*Proof.* (a) Let  $\pi = (\mu_0, \dots, \mu_{N-1})$  be uniformly N-stage optimal. Then

$$(43) \quad T(J_0) = J_1^* = T_{\mu_{N-1}}(J_0)$$

by definition. For every  $\mu \in M$ ,

$$(T_\mu T)(J_0) = (T_\mu T_{\mu_{N-1}})(J_0),$$

which implies

$$T^2(J_0) = \inf_{\mu \in M} (T_\mu T)(J_0) = \inf_{\mu \in M} (T_\mu T_{\mu_{N-1}})(J_0) \geq J_2^* = (T_{\mu_{N-2}} T_{\mu_{N-1}})(J_0) \geq T^2(J_0).$$

Therefore

$$(44) \quad T^2(J_0) = J_2^* = (T_{\mu_{N-2}} T)(J_0) = (T_{\mu_{N-2}} T_{\mu_{N-1}})(J_0).$$

Replace (43) by (44) and continue. This proves (41).

(b) This follows from (a) and (20). Q.E.D.

It follows from Proposition 2(b) that when  $M$  is rich enough so that Assumption C holds, then existence of a uniformly N-stage optimal policy is equivalent to attainment of the infimum in the DP algorithm. We state this as a separate proposition.

PROPOSITION 3. *Let Assumption C hold. A uniformly N-stage optimal policy exists if and only if the infimum in*

$$(45) \quad \inf_{u \in \Gamma_x} \int^* \{g(x, u, w) + T^{k-1}(J_0)[f(x, u, w)]\} p(dw | x, u)$$

*is achieved for every  $x \in S$ ,  $k = 1, \dots, N$ .*

**4. The model without topological assumptions.** The simplest special case of our model is when we take  $F = F_S$  in Assumption A and place restrictions on  $M$  only as follows. A subset  $\Gamma$  of  $SC$  is given with the property  $\text{proj}_S \Gamma = S$  and  $M$  is taken to be the set of all mappings from  $S$  to  $C$  whose graphs lie in  $\Gamma$ . Note that  $\Gamma$  and  $M$  correspond as in (5). It is easy to see that Assumptions A, B, C, and D hold when  $F$  and  $M$  have been so chosen. Hence the results of Propositions 1–3 apply.

It is not customary to use outer integration in connection with Dynamic Programming, so the model outlined here is somewhat unusual. A special case of this model often considered is the case of a countable disturbance space [1]. As mentioned

earlier, integration reduces to summation in such a model. The countable disturbance space model is more restrictive than the Borel space model of § 6 in that a countable disturbance space is a special case of a Borel space. It is more general in that  $S$  and  $C$  are not required to have a Borel structure. The model described in the first paragraph of this section is, of course, more general than both. The main advantage that it offers is simplicity—there is no need to introduce elaborate topological assumptions in order to ascertain the validity of the DP algorithm. There are, however, inherent limitations centering around the pathologies of outer integration (Appendix A) in the nontopological model. Topological assumptions also play an important role in the treatment of problems with imperfect state information (see [2], [29]).

**5. Borel space models with semicontinuity assumptions.** We first introduce some notation and definitions. For any topological space  $Y$  we denote by  $\mathcal{B}_Y$  the Borel  $\sigma$ -algebra generated by the open sets. A *Borel space*  $X$  is a topological space such that there exists a complete separable metric space  $Y$  and a homeomorphism  $\varphi$  of  $X$  into  $Y$  with  $\varphi(X) \in \mathcal{B}_Y$ .

It follows that a Borel space  $X$  is metrizable and separable. Note that if  $X$  and  $Y$  are Borel spaces, then the product space  $XY$  equipped with the product topology is also a Borel space and  $\mathcal{B}_{XY}$  equals the product  $\sigma$ -algebra  $\mathcal{B}_X \mathcal{B}_Y$  on  $XY$  [21, Chap. 1]. Also, every Borel subset of a Borel space becomes a Borel space when endowed with the relative topology. The extended real line  $[-\infty, +\infty]$  with the topology generated by the open real intervals together with the sets  $[-\infty, \alpha)$ , and  $(\alpha, +\infty]$ ,  $\alpha$  real, is a Borel space. In what follows we implicitly assume that every Borel subset of  $[-\infty, +\infty]$  is endowed with the corresponding relative topology and is thus a Borel space. If  $X$  and  $Y$  are Borel spaces and  $f: X \rightarrow Y$  is such that  $f^{-1}(B) \in \mathcal{B}_X$  for each  $B \in \mathcal{B}_Y$ , then we say that  $f$  is *Borel measurable*.

If  $X$  is a Borel space, we denote by  $P(X)$  the set of probability measures on  $\mathcal{B}_X$ . We take the topology on  $P(X)$  to be the weakest with respect to which all mappings of the form  $p \rightarrow \int f dp$  are continuous, as  $f$  ranges over the set of bounded continuous real-valued functions on  $X$ . With this topology  $P(X)$  becomes a Borel space [21, Chap. 2]. Let  $X$  and  $Y$  be Borel spaces and for each  $x \in X$ , let  $q(dy|x)$  be a probability measure on  $\mathcal{B}_Y$ . If the mapping  $x \rightarrow q(dy|x)$  is continuous from  $X$  to  $P(Y)$ , we will say that  $q(dy|x)$  is a *continuous stochastic kernel on  $Y$  given  $X$* .

We now define two special cases of the model of § 2. In both cases,  $S$ ,  $C$  and  $W$  are Borel spaces,  $\mathcal{F} = \mathcal{B}_W$ ,  $p(dw|x, u)$  is a continuous stochastic kernel on  $W$  given  $SC$ , and  $f(x, u, w)$  is continuous from  $SCW$  to  $S$ .

*Lower semicontinuous model.* Here  $C$  is compact, and  $g$  is lower semi-continuous and bounded below. A subset  $\Gamma$  of  $SC$  is given and is assumed to be of the form

$$\Gamma = \bigcup_{j=1}^{\infty} \Gamma_j,$$

where for all  $j$ ,  $\Gamma_j \subset \Gamma_{j+1}$ ,  $\Gamma_j$  is a closed subset of  $SC$ , and for all  $(x, u) \in SC$

$$(46) \quad \lim_{j \rightarrow \infty} \inf_{(x,u) \in \Gamma_j - \Gamma_{j-1}} \int g(x, u, w) p(dw|x, u) = +\infty.$$

(By convention the infimum over the empty set is  $+\infty$ . Thus we allow the possibility that for some  $\bar{j}$ ,  $\Gamma_j = \Gamma_{\bar{j}}$  for all  $j \geq \bar{j}$ .) It is also assumed that  $\text{proj}_S \Gamma = S$ . The set  $M$  of admissible control functions is taken to be the set of all Borel measurable functions from  $S$  to  $C$  whose graphs lie in  $\Gamma$ . (Notice that if the sets  $\Gamma_j$ ,  $j = 1, 2, \dots$ , are compact then there is no loss of generality in assuming that  $C$  is compact. This is true because if

$C$  is not compact, it can be homeomorphically embedded in a compact Borel space  $\tilde{C}$  [7, Chap. 9, Cor. 9.2] and the images of  $\Gamma_j$  are compact and hence closed in  $S\tilde{C}$ . There is no need to extend  $f$  and  $g$  to  $S\tilde{C}W$  nor  $p(dw|x, u)$  to  $S\tilde{C}$  for the proof we give of Proposition 4.)

*Upper semicontinuous model.* Here  $g$  is upper semicontinuous and bounded above. An open subset  $\Gamma$  of  $SC$  is given and it is assumed that  $\text{proj}_S \Gamma = S$ . The set  $M$  is the set of all Borel measurable functions from  $S$  to  $C$  whose graphs lie in  $\Gamma$ .

By selecting an appropriate subset  $F \subset F_S$  for each model, we show now that some of the Assumptions A, B, C and D are satisfied. This in turn will allow application of some of the results of § 3.

PROPOSITION 4. (a) *In the lower semicontinuous model Assumptions A, B and C are satisfied with  $F$  being the set of lower semicontinuous functions  $J: S \rightarrow (-\infty, +\infty)$  which are bounded below.*

(b) *In the upper semicontinuous model Assumptions A and B are satisfied with  $F$  being the set of upper semicontinuous functions  $J: S \rightarrow [-\infty, +\infty)$  which are bounded above.*

*Proof.* (a) If  $J: S \rightarrow (-\infty, +\infty)$  is lower semicontinuous and bounded below, then

$$H(x, u) = \int \{g(x, u, w) + J[f(x, u, w)]\}p(dw|x, u)$$

is also [25, Lemma 3.4]. Condition (46) guarantees that  $\{(x, u) \in \Gamma: H(x, u) \leq \alpha\}$  is closed for each real  $\alpha$ . The result follows from a simple modification of Lemma 3.4 and the Selection Theorem of [15].

(b) This follows from (17) of [9] and the fact that  $\text{proj}_S$  is an open mapping. Q.E.D.

By combining Proposition 4 with the results of § 2, we obtain the following:

COROLLARY 4.1. *In the lower semicontinuous model, we have  $J_N^* = T^N(J_0)$  and there exists a uniformly  $N$ -stage optimal policy.*

COROLLARY 4.2. *In the upper semicontinuous model, if  $J_k^*(x) > -\infty$  for all  $x \in S$  and  $k = 1, 2, \dots, N$ , then  $J_N^* = T^N(J_0)$ , and for every  $\varepsilon > 0$  there exists an  $\varepsilon$ -optimal policy.*

David Freedman [9] has proved results quite similar to Corollaries 4.1 and 4.2 by placing control constraints, not directly on the control  $u$  as we have done by requiring  $(x, u) \in \Gamma$ , but rather on the pair  $(x, P)$ , where  $P$  is the distribution of the subsequent state. Since the mapping  $(x, u) \rightarrow (x, P)$  is continuous in the semicontinuous models, requiring  $(x, u)$  to be in an open set (our upper semicontinuous model) is slightly more general than requiring  $(x, P)$  to be in an open set (Freedman's model), while requiring  $(x, u)$  to be in the union of an increasing sequence of closed sets (our lower semicontinuous model) is significantly more general than requiring  $(x, P)$  to be in a closed set (Freedman's model). Our lower semicontinuous model does not require a compact state space. For example, we can take  $S = R^n$ ,  $C$  to be the one point compactification of  $R^m$ ,

$$\Gamma_j = \{(x, u): u'u \leq j\}, \quad g(x, u, w) = x'Qx + u'Ru,$$

where  $Q$  is a positive semidefinite and  $R$  is a positive definite matrix of appropriate dimension.

**6. General Borel space models with perfect state information.** For the models of this section we shall need the notions of analytic sets, universally measurable sets and related facts. For more detailed treatments we refer the reader to [2], [6], [11], [14], [16], [21].

Let  $\mathcal{N}$  be the cross product of countably many copies of the positive integers. Let the set of positive integers have the discrete topology and  $\mathcal{N}$  the product topology. A separable metric space  $A$  is *analytic* if there is a continuous function  $f$  mapping  $\mathcal{N}$  onto  $A$ . In what follows the empty set will also be considered analytic.

We list some properties of analytic sets that we shall be using:

(a) Every Borel space is analytic but in every uncountable Borel space there exist analytic subsets which are not Borel spaces [14], § 38VI.

(b) The countable union, intersection, and cross product of analytic sets is analytic [21, Chap. 1, Thms. 3.1 and 3.2].

(c) If  $X$  and  $Y$  are Borel spaces,  $A \subset X$  and  $B \subset Y$  are analytic sets, and  $f$  is a Borel measurable function from  $X$  to  $Y$ , then  $f(A)$  and  $f^{-1}(B)$  are analytic [21, Chap. 1, Thm. 3.5]. As a consequence, if  $D$  is an analytic subset of  $XY$ , then  $\text{proj}_X D$  is analytic.

In addition to the Borel  $\sigma$ -algebra, we are interested in two more  $\sigma$ -algebras that arise naturally in a Borel space  $X$ . The *analytic  $\sigma$ -algebra*, denoted  $\mathcal{A}_X$ , is the  $\sigma$ -algebra generated by the analytic subsets of  $X$ . The *universal  $\sigma$ -algebra*, denoted  $\mathcal{U}_X$ , is the intersection of all completions with respect to finite measures of the Borel  $\sigma$ -algebra  $\mathcal{B}_X$ . We have

$$\mathcal{B}_X \subset \mathcal{A}_X \subset \mathcal{U}_X,$$

and if  $X$  is uncountable both inclusions are strict. In fact, it is possible to prove that if  $X$  is uncountable, then under the continuum hypothesis  $\mathcal{U}_X$  has a larger cardinality than both  $\mathcal{B}_X$  and  $\mathcal{A}_X$ . (We are indebted to Professor J. Doob for pointing out this fact to us.)

Let  $X$  and  $Y$  be Borel spaces,  $D$  be a subset of  $X$ , and  $f: D \rightarrow Y$ . If  $D \in \mathcal{A}_X$  and  $f^{-1}(B) \in \mathcal{A}_X$  for all  $B \in \mathcal{B}_Y$ , we say that  $f$  is *analytically measurable*. If  $D \in \mathcal{U}_X$  and  $f^{-1}(B) \in \mathcal{U}_X$  for all  $B \in \mathcal{B}_Y$ , we say that  $f$  is *universally measurable*. If  $D$  is analytic,  $Y = [-\infty, +\infty]$  and the set  $\{x \in D: f(x) < \alpha\}$  is analytic for every real  $\alpha$ , we say that  $f$  is *lower semianalytic*. For a lower semianalytic  $f$ , the sets  $\{x \in D: f(x) \leq \alpha\}$  are also analytic for every  $\alpha \in [-\infty, +\infty]$ . Note that a lower semianalytic function is analytically measurable and hence also universally measurable, the sum of two lower semianalytic functions is lower semianalytic, and a Borel measurable function from  $X$  to  $[-\infty, +\infty]$  is lower semianalytic.

If  $X$  is a Borel space and  $p \in P(X)$ , then  $p$  has a unique extension to a probability measure on  $\mathcal{U}_X$ . We denote this extension by  $p$  also, and we write  $p(E)$  instead of  $p^*(E)$  when  $E \in \mathcal{U}_X$ . Likewise, if  $f: X \rightarrow [-\infty, +\infty]$  is a universally measurable function we will write  $\int f dp$  in place of  $\int^* f dp$ . Under these circumstances  $\int f dp$  obeys the rules of classical integration, provided we take care in handling the expression  $+\infty - \infty$ .

If  $X$  and  $Y$  are Borel spaces,  $q(dy|x)$  is a probability measure on  $\mathcal{B}_Y$  for each  $x \in X$ , and the function  $q(B|\cdot)$  is Borel measurable from  $X$  to  $[0, 1]$  for all  $B \in \mathcal{B}_Y$ , we say that  $q(dy|x)$  is a *Borel measurable stochastic kernel on  $Y$  given  $X$* . The stochastic kernel  $q(dy|x)$  is Borel measurable if and only if the mapping  $x \rightarrow q(\cdot|x)$  is Borel measurable from  $X$  to  $P(Y)$ .

We now specify the two special cases of the problem of § 2 to be considered in this section. In both cases,  $S$ ,  $C$  and  $W$  are Borel spaces,  $\mathcal{F} = \mathcal{B}_W$ ,  $p(dw|x, u)$  is a Borel measurable stochastic kernel on  $W$  given  $SC$ ,  $f$  is Borel measurable, and  $g$  is lower semianalytic. In both models an analytic subset  $\Gamma$  of  $SC$  is given with  $\text{proj}_S \Gamma = S$ . The models differ only in the specification of  $M$ .

*Borel model with analytically measurable policies (BAP for short):* Here  $M$  consists of all analytically measurable functions from  $S$  to  $C$  whose graphs lie in  $\Gamma$ .

*Borel model with universally measurable policies (BUP for short):* Here  $M$  consists of all universally measurable functions from  $S$  to  $C$  whose graphs lie in  $\Gamma$ .

Our main result of this section is the following:

**PROPOSITION 5.** (a) *In BAP Assumptions A and B are satisfied with  $F$  being the class of lower semianalytic functions  $J: S \rightarrow [-\infty, +\infty]$ .*

(b) *In BUP Assumptions A, B, C and D are satisfied with  $F$  being the class of lower semianalytic functions  $J: S \rightarrow [-\infty, +\infty]$ .*

We postpone the proof of Proposition 5 until we develop some further machinery. By combining Proposition 5 with the results of § 3 we obtain the following:

**COROLLARY 5.1.** *In BAP if  $J_k^*(x) > -\infty$  for all  $x \in S$  and  $k = 1, \dots, N$ , then  $J_N^* = T^N(J_0)$  and for every  $\varepsilon > 0$  there exists an  $\varepsilon$ -optimal policy.*

**COROLLARY 5.2.** *Consider BUP.*

(a) *If  $J_k^*(x) > -\infty$  for all  $x \in S$  and  $k = 1, \dots, N$ , then  $J_N^* = T^N(J_0)$  and for every  $\varepsilon > 0$  there exists an  $\varepsilon$ -optimal policy.*

(b) *If  $J_{k,\pi}(x) < +\infty$  for all  $x \in S$ ,  $\pi \in \Pi_k$  and  $k = 1, \dots, N$ , then  $J_N^* = T^N(J_0)$  and for every sequence  $\{\varepsilon_n\}$ ,  $\varepsilon_n > 0$ ,  $n = 1, 2, \dots$ ,  $\varepsilon_n \downarrow 0$ , there exists a sequence of policies exhibiting  $\{\varepsilon_n\}$  dominated convergence to optimality. If in addition  $J_N^*(x) > -\infty$  for all  $x \in S$ , then for every  $\varepsilon > 0$  there exists an  $\varepsilon$ -optimal policy.*

(c) *If the infimum in*

$$\inf_{u \in \Gamma_x} \int^* \{g(x, u, w) + T^{k-1}(J_0)[f(x, u, w)]\} p(dw | x, u)$$

*is attained for every  $x \in S$  and  $k = 1, \dots, N$ , then  $J_k^* = T^k(J_0)$ ,  $k = 1, \dots, N$ , and there exists a uniformly  $N$ -stage optimal policy.*

We now provide two results that are crucial in our development. The first is often attributed to von Neuman [17], but was also proved by Jankov [13]. A proof of the version given here may be found in Blackwell, Freedman and Orkin [4]. Part (a) of the second result is contained in a proof given by Blackwell, Freedman and Orkin [4, Thm. (43)]. Part (b) is an extension of a selection theorem of Brown and Purves [5, Thm. 2] in that  $f$  is allowed to be lower semianalytic rather than Borel measurable. Our proof parallels the proofs of [4] and [5].

**JANKOV-VON NEUMANN LEMMA.** *Let  $X$  and  $Y$  be Borel spaces and  $A \subset XY$  be an analytic set. Then there exists an analytically measurable function  $\varphi: \text{proj}_X A \rightarrow Y$  such that  $(x, \varphi(x)) \in A$  for every  $x \in \text{proj}_X A$ .*

**SELECTION THEOREM.** *Let  $X$  and  $Y$  be Borel spaces,  $D \subset XY$  be an analytic set, and  $f: D \rightarrow [-\infty, +\infty]$  be a lower semianalytic function. Define  $g: \text{proj}_X D \rightarrow [-\infty, +\infty]$  by*

$$(47) \quad g(x) = \inf_{y \in D_x} f(x, y),$$

*where  $D_x = \{y: (x, y) \in D\}$ . Then  $g$  is lower semianalytic. Furthermore:*

(a) *For every  $\varepsilon > 0$  there exists an analytically measurable function  $\varphi: \text{proj}_X D \rightarrow Y$  such that for all  $x \in \text{proj}_X D$*

$$\begin{aligned} f[x, \varphi(x)] &\leq g(x) + \varepsilon && \text{if } g(x) > -\infty, \\ f[x, \varphi(x)] &\leq -\frac{1}{\varepsilon} && \text{if } g(x) = -\infty. \end{aligned}$$



(b) *The set*

$$I = \{x \in \text{proj}_X D : \text{for some } y_0 \in D_x, f(x, y_0) = g(x)\}$$

is universally measurable, and for every  $\varepsilon > 0$  there exists a universally measurable function  $\varphi : \text{proj}_X D \rightarrow Y$  such that for all  $x \in \text{proj}_X D$

$$\begin{aligned} f[x, \varphi(x)] &= g(x) && \text{if } x \in I, \\ f[x, \varphi(x)] &\leq g(x) + \varepsilon && \text{if } x \notin I, g(x) > -\infty, \\ f[x, \varphi(x)] &\leq -\frac{1}{\varepsilon} && \text{if } x \notin I, g(x) = -\infty. \end{aligned}$$

*Proof.* (a) Since  $\{x : g(x) < \alpha\} = \text{proj}_X \{(x, y) \in D : f(x, y) < \alpha\}$ ,  $g$  is lower semi-analytic. For  $k = 0, \pm 1, \pm 2, \dots$  define

$$\begin{aligned} A(k) &= \{(x, y) \in D : f(x, y) < k\varepsilon\}, \\ B(k) &= \{x \in \text{proj}_X D : (k - 1)\varepsilon \leq g(x) < k\varepsilon\}, \\ B(-\infty) &= \{x \in \text{proj}_X D : g(x) = -\infty\}, \\ B(+\infty) &= \{x \in \text{proj}_X D : g(x) = +\infty\}. \end{aligned}$$

The sets  $A(k)$ ,  $k = 0, \pm 1, \pm 2, \dots$  and  $B(-\infty)$  are analytic, while the sets  $B(k)$ ,  $k = 0, \pm 1, \pm 2, \dots$  and  $B(+\infty)$  are analytically measurable. By the Jankov-Von Neumann Lemma there exists, for each  $k = 0, \pm 1, \pm 2, \dots$ , an analytically measurable  $\varphi_k : \text{proj}_X A(k) \rightarrow C$  with  $(x, \varphi_k(x)) \in A(k)$  for all  $x \in \text{proj}_X A(k)$ , and an analytically measurable  $\bar{\varphi} : \text{proj}_X D \rightarrow C$  such that  $(x, \bar{\varphi}(x)) \in D$  for all  $x \in \text{proj}_X D$ . Let  $k^*$  be an integer such that  $k^* \leq -1/\varepsilon^2$ . Define  $\varphi : \text{proj}_X D \rightarrow C$  by

$$\varphi(x) = \begin{cases} \varphi_k(x) & \text{if } x \in B(k), \quad k = 0, \pm 1, \pm 2, \dots, \\ \bar{\varphi}(x) & \text{if } x \in B(+\infty), \\ \varphi_{k^*}(x) & \text{if } x \in B(-\infty). \end{cases}$$

Since  $B(k) \subset \text{proj}_X A(k)$  and  $B(-\infty) \subset \text{proj}_X A(k)$  for all  $k$  this definition is possible. Then  $\varphi$  has the required properties.

(b) Denote by  $Q$  the set of rationals and let  $Q^* = Q \cup \{-\infty, +\infty\}$ . Denote also by  $R^*$  the extended reals. Consider the set  $E \subset XYR^*$  defined by

$$E = \{(x, y, b) : (x, y) \in D, f(x, y) \leq b\}.$$

Since

$$E = \bigcap_{k=1}^{\infty} \bigcup_{r \in Q^*} \left\{ (x, y, b) : (x, y) \in D, f(x, y) \leq r, r \leq b + \frac{1}{k} \right\},$$

it follows that  $E$  is analytic in  $XYR^*$  and hence the set

$$A = \text{proj}_{XR^*}(E)$$

is analytic in  $XR^*$ . The mapping  $T : \text{proj}_X D \rightarrow XR^*$  defined by

$$T(x) = (x, g(x))$$

is analytically measurable and

$$I = \{x : (x, g(x)) \in A\} = T^{-1}(A).$$

Since the inverse image under a universally measurable function of a universally measurable set is universally measurable,  $I$  is universally measurable.

Since  $E$  is analytic, by the Jankov–Von Neumann Lemma there is an analytically measurable  $\rho: A \rightarrow Y$  such that  $(x, \rho(x, b), b) \in E$  for every  $(x, b) \in A$ . Define  $\Psi: I \rightarrow Y$  by

$$\Psi(x) = \rho(x, g(x)) = (\rho \circ T)(x) \quad \forall x \in I.$$

Then  $\Psi$  is universally measurable and by construction

$$(48) \quad f[x, \Psi(x)] = g(x) \quad \forall x \in I.$$

By part (a) there exists an analytically measurable  $\Psi_\varepsilon: \text{proj}_X D \rightarrow Y$  such that

$$(49) \quad f[x, \Psi_\varepsilon(x)] \leq g(x) + \varepsilon \quad \text{if } g(x) > -\infty,$$

$$(50) \quad f[x, \Psi_\varepsilon(x)] \leq -1/\varepsilon \quad \text{if } g(x) = -\infty.$$

Define  $\varphi: \text{proj}_X D \rightarrow Y$  by

$$\varphi(x) = \begin{cases} \Psi(x) & \text{if } x \in I, \\ \Psi_\varepsilon(x) & \text{if } x \in \text{proj}_X D - I. \end{cases}$$

Then  $\varphi$  is universally measurable and, by (48)–(50), it has the required properties. Q.E.D.

Suppose  $X$  and  $Y$  are Borel spaces,  $f: XY \rightarrow [-\infty, +\infty]$  is universally measurable (i.e., measurable with respect to  $\mathcal{U}_{XY}$ ), and  $q(dy|x)$  is a Borel measurable stochastic kernel. Then it can be shown that  $g(x) = \int f(x, y)q(dy|x)$  is universally measurable. If  $f$  is actually Borel measurable, so is  $g$ . If  $f$  is lower semianalytic, then  $g$  is also lower semianalytic. This last fact can be obtained by modifying Lemma (29) of [4] (see [29]).

We are now ready to prove Proposition 5.

*Proof of Proposition 5.* Let  $J: S \rightarrow [-\infty, +\infty]$  be lower semianalytic. Then the function  $H: SC \rightarrow [-\infty, +\infty]$  defined by

$$H(x, u) = \int \{g(x, u, w) + J[f(x, u, w)]\}p(dw|x, u)$$

is lower semianalytic. It follows that

$$T(J)(x) = \inf_{u \in \Gamma_x} H(x, u),$$

is lower semianalytic. Since  $J_0$  is lower semianalytic, Assumption A is satisfied for both BAP and BUP. The Selection Theorem guarantees that Assumption B is satisfied for BAP, while Assumptions B and C are satisfied for BUP. It remains to verify Assumption D for BUP. We first show that the function

$$(51) \quad (x, u) \rightarrow -p(\{w: J[f(x, u, w)] = -\infty\} | x, u)$$

is lower semianalytic whenever  $J$  is. Define  $t(dx'|x, u)$  by

$$t(E|x, u) = p(\{w: f(x, u, w) \in E\} | x, u) \quad \forall E \in \mathcal{B}_S.$$

We will show that  $t(dx'|x, u)$  is a Borel measurable stochastic kernel. Clearly for fixed  $(x, u)$ ,  $t(\cdot|x, u)$  is a probability measure on  $S$ . We need to show that  $p(B_{(x,u)}|x, u)$  is Borel measurable for each Borel subset  $B$  of  $SCW$  [ $B_{(x,u)}$  is the cross section

$\{w: (x, u, w) \in B\}$ . It is easy to show that the sets  $B \in \mathcal{B}_{SCW}$  for which  $p(B_{(x,u)}|x, u)$  is measurable form a Dynkin system, so that by the Dynkin system theorem, we need only verify that  $p((B_1 B_2 B_3)_{(x,u)}|x, u)$  is measurable for all  $B_1 \in \mathcal{B}_S, B_2 \in \mathcal{B}_C, B_3 \in \mathcal{B}_W$ . But

$$p((B_1 B_2 B_3)_{(x,u)}|x, u) = \begin{cases} p(B_3|x, u) & \text{if } (x, u) \in B_1 B_2, \\ 0 & \text{otherwise,} \end{cases}$$

and this is measurable since  $p(B_3|x, u)$  is measurable. Hence  $t(dx'|x, u)$  is a Borel measurable stochastic kernel. We have for all  $(x, u) \in SC$

$$(52) \quad \begin{aligned} -p(\{w: J[f(x, u, w)] = -\infty\}|x, u) &= -t(\{x': J(x') = -\infty\}|x, u) \\ &= \int -\chi_{\{J = -\infty\}} t(dx'|x, u). \end{aligned}$$

The function  $-\chi_{\{J = -\infty\}}$  can be easily seen to be lower semianalytic. It follows from the remark preceding the proof that the function of (51) is lower semianalytic. Hence the set

$$A(J) = \{(x, u) \in \Gamma: p(\{w: J[f(x, u, w)] = -\infty\}|x, u) > 0\}$$

is analytic, and by the Jankov-Von Neumann Lemma there is a  $\mu_J \in M$  such that  $(x, \mu_J(x)) \in A(J)$  whenever  $x \in \text{proj}_S A(J)$ . If  $\mu \in M$  and  $\hat{\mu}$  is as in Assumption D, it follows from the fact that  $\text{proj}_S A(J)$  is analytic and hence universally measurable that  $\hat{\mu} \in M$ . If  $\mu_1, \mu_2 \in M$ , then  $T_{\mu_1}(J)$  and  $T_{\mu_2}(J)$  can be easily shown to be universally measurable. Hence if  $\mu$  is as in Assumption D it follows that  $\bar{\mu} \in M$ . Q.E.D.

*Remark.* In the models in which  $C$  is equipped with a  $\sigma$ -algebra, one can speak of randomized policies  $\pi = (\mu_0, \dots, \mu_{N-1})$ , where  $\mu_k(du_k|x_k)$  is an appropriately measurable stochastic kernel on  $C$  given  $S$ . Control constraints can be introduced by requiring that  $\mu_k(\Gamma_x|x) = 1$  for every  $x \in S, k = 0, \dots, N-1$ , where  $\Gamma = \{(x, u): u \in \Gamma_x\}$  is some prescribed subset of  $SC$ . The cost corresponding to such a policy is

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

where

$$T_{\mu_k}(J)(x) = \int_C \int_W \{g(x, u, w) + J[f(x, u, w)]\} p(dw|x, u) \mu_k(du|x).$$

It is clear that  $J_{N,\pi}$  is bounded below by  $T^N(J_0)$ , so if  $J_N^* = T^N(J_0)$ , the admission of randomized policies to the models considered does not alter the optimal cost function. Note however that in the example of § 3 if randomized policies are admitted, then the optimal cost function becomes  $J_2^*(\alpha) = J_2^*(\beta) = -\infty$  and is different from the one corresponding to nonrandomized policies. Furthermore an optimal randomized policy exists.

*Remark.* There is the  $\sigma$ -algebra of “ $C$ -sets” studied by Selivanovskij [26]. This  $\sigma$ -algebra, which we call the limit  $\sigma$ -algebra, is contained strictly between the analytic and universal  $\sigma$ -algebras in Borel spaces and has the property that all the results of this section remain valid if the words “universally measurable” are replaced by “limit measurable”. This  $\sigma$ -algebra is the minimal acceptable  $\sigma$ -algebra for DP in the sense that the composition of limit measurable functions is limit measurable and every analytically measurable function is limit measurable, but no smaller  $\sigma$ -algebra has these two properties. The limit  $\sigma$ -algebra is discussed more fully in [2], [30].

**Appendix A: The outer integral.** Throughout this appendix,  $(X, \mathcal{B}, p)$  is a probability space. Unless otherwise specified  $f, g,$  and  $h$  are functions from  $X$  to  $[-\infty, +\infty]$ .

DEFINITION A.1. If  $f \geq 0$ , the *outer integral* of  $f$  with respect to  $p$  is defined by

$$(A.1) \quad \int^* f dp = \inf \left\{ \int g dp : f \leq g, g \text{ is } \mathcal{B}\text{-measurable} \right\}.$$

If  $f$  is arbitrary, define

$$(A.2) \quad \int^* f dp = \int^* f^+ dp - \int^* f^- dp.$$

If  $\int^* f^+ dp < +\infty$ , we say  $f$  is *outer summable above*. If  $\int^* f^- dp < +\infty$ , we say  $f$  is *outer summable below*. If  $f$  is outer summable above or outer summable below, we say  $f$  is *outer summable*. In the following discussion, simple proofs are omitted.

LEMMA A.1. If  $f \geq 0$ , then there exists a  $\mathcal{B}$ -measurable  $g$  with  $g \geq f$ , such that

$$(A.3) \quad \int^* f dp = \int g dp.$$

LEMMA A.2. If  $f \geq 0$ ,  $h \geq 0$ , then

$$(A.4) \quad \int^* (f+h) dp \leq \int^* f dp + \int^* h dp.$$

If either  $f$  or  $h$  is  $\mathcal{B}$ -measurable, then equality holds in (A.4).

We provide an example to show that strict inequality can occur in (A.4), even if  $f+h$  is  $\mathcal{B}$ -measurable. For this and subsequent examples we will need the following observation: For any  $E \subset X$ ,

$$(A.5) \quad \int^* \chi_E dp = p^*(E),$$

where  $p^*(E)$  is the  $p$ -outer measure defined by

$$(A.6) \quad p^*(E) = \inf \{ p(B) : E \subset B, B \in \mathcal{B} \}.$$

This follows from the fact that for any set  $E$  there exists a set  $A \in \mathcal{B}$  such that  $E \subset A$  and  $p(A) = p^*(E)$ .

*Example A.1.* Let  $X = [0, 1]$ ,  $\mathcal{B}$  be the Borel  $\sigma$ -algebra, and  $p$  be Lebesgue measure restricted to  $\mathcal{B}$ . Let  $E \subset X$  be a set for which  $p^*(X-E) = 1$  [10, § 16, Thm. E]. Then

$$\int (\chi_E + \chi_{X-E}) dp = \int 1 dp = 1, \quad \int^* \chi_E dp + \int^* \chi_{X-E} dp = 2,$$

and strict inequality holds in (A.4).

Lemma A.2 cannot be extended to (possibly negative) bounded functions even if  $h$  is  $\mathcal{B}$ -measurable, as the following example demonstrates.

*Example A.2.* Let  $(X, \mathcal{B}, p)$  and  $E$  be as before. Let  $f = \chi_E - \chi_{X-E}$ ,  $h = 1$ . Then

$$\begin{aligned} \int^* (f+h) dp &= \int^* 2\chi_E dp = 2, \\ \int^* f dp + \int h dp &= \int^* \chi_E dp - \int^* \chi_{X-E} dp + 1 = 1. \end{aligned}$$

LEMMA A.3. (a) If  $f \leq g$  then  $\int^* f dp \leq \int^* g dp$ .

(b) If  $\epsilon > 0$  and  $f \leq g \leq f + \epsilon$ , then

$$(A.7) \quad \int^* f dp \leq \int^* g dp \leq \int^* f dp + 2\epsilon.$$

(c) If  $f$  is outer summable, then

$$(A.8) \quad \int^* (-f) dp = - \int^* f dp.$$

(d) If  $A, B \in \mathcal{B}$  are disjoint, then for any  $f$

$$(A.9) \quad \int^* \chi_{A \cup B} f dp = \int^* \chi_A f dp + \int^* \chi_B f dp.$$

(e) If  $E \subset X$  satisfies  $p^*(E) = 0$ , then for every  $f$

$$\int^* f dp = \int^* \chi_{X-E} f dp.$$

(f) If  $p^*({x: f(x) = +\infty}) > 0$  then, for every  $g$ ,  $\int^* (g + f) dp = +\infty$ .

(g) If  $p^*({x: f(x) = -\infty}) > 0$  then, for every  $g$ , either  $\int^* (g + f) dp = +\infty$  or  $\int^* (g + f) dp = -\infty$ .

*Proof.* (b) In light of (a), it remains only to show that

$$(A.10) \quad \int^* (f + \epsilon) dp \leq \int^* f dp + 2\epsilon.$$

For  $g_1 \geq f^+$ ,  $g_1$   $\mathcal{B}$ -measurable and

$$\int^* f^+ dp = \int g_1 dp,$$

we have

$$(f + \epsilon)^+ \leq g_1 + \epsilon$$

so

$$(A.11) \quad \int^* (f + \epsilon)^+ dp \leq \int g_1 dp + \epsilon = \int^* f^+ dp + \epsilon.$$

For  $g_2 \geq (f + \epsilon)^-$ ,  $g_2$   $\mathcal{B}$ -measurable and

$$(A.12) \quad \int^* (f + \epsilon)^- dp = \int g_2 dp,$$

we have

$$\begin{aligned} g_2 + \epsilon &\geq (f + \epsilon)^- + \epsilon \\ &= \max \{f^- - \epsilon, 0\} + \epsilon \\ &\geq f^-, \end{aligned}$$

so

$$\begin{aligned}
 \varepsilon + \int^* (f + \varepsilon)^- dp &= \varepsilon + \int g_2 dp \\
 \text{(A.13)} \qquad \qquad \qquad &= \int (g_2 + \varepsilon) dp \\
 &\cong \int^* f^- dp \quad \text{by (a).}
 \end{aligned}$$

Combine (A.11) and (A.13) to conclude (A.10).

(f) We have  $(g + f)^+(x) = +\infty$  if  $f(x) = +\infty$ , so that  $p^*({x : (g + f)^+(x) = +\infty}) > 0$ . Hence  $\int^* (g + f)^+ dp = +\infty$  and it follows that  $\int^* (g + f) dp = +\infty$ .

(g) Consider the sets  $E = {x : f(x) = -\infty}$  and  $E_g = {x : f(x) = -\infty, g(x) < +\infty}$ . If  $p^*(E_g) = 0$  then we have

$$\text{(A.14)} \qquad p^*(E - E_g) = p^*(E - E_g) + p^*(E_g) \cong p^*(E) > 0.$$

Since we have  $f(x) + g(x) = +\infty$  for  $x \in E - E_g$ , it follows from (f) that  $\int^* (g + f) dp = +\infty$ . If  $p^*(E_g) > 0$ , then  $p^*({x : (g + f)^-(x) = +\infty}) \cong p^*(E_g) > 0$  and hence, by (f),  $\int^* (g + f)^- dp = +\infty$ . Hence if  $\int^* (g + f)^+ dp = +\infty$  then  $\int^* (g + f) dp = +\infty$ , while if  $\int^* (g + f)^+ dp < +\infty$  then  $\int^* (g + f) dp = -\infty$ . Q.E.D.

The bound given in (A.7) is the sharpest possible. To see this, let  $f$  be as defined in Example A.2,  $g = f + 1$  and  $\varepsilon = 1$ . Despite these pathologies of outer integration, there is a monotone convergence theorem, which we now prove.

**THEOREM A.1.** *If  $\{f_n\}$  is a sequence of nonnegative functions and  $f_n \uparrow f$ , then*

$$\text{(A.15)} \qquad \int^* f_n dp \uparrow \int^* f dp.$$

*If  $\{f_n\}$  is a sequence of nonpositive functions and  $f_n \downarrow f$ , then*

$$\int^* f_n dp \downarrow \int^* f dp.$$

*Proof.* We prove the first statement of the theorem. The second follows from the first by Lemma A.3(c). Assume  $f_n \cong 0$  and  $f_n \uparrow f$ . Let  $\{g_n\}$  be a sequence of  $\mathcal{B}$ -measurable functions such that  $g_n \cong f_n$  and

$$\text{(A.16)} \qquad \int^* f_n dp = \int g_n dp.$$

If for some  $n$ ,  $\int g_n dp = \int^* f_n dp = +\infty$ , then (A.15) is assured. If not, then for every  $n$

$$\text{(A.17)} \qquad \int g_n dp < \infty.$$

Suppose (A.17) holds for every  $n$  and for some  $n$ ,

$$p({x : g_n(x) > g_{n+1}(x)}) > 0.$$

Then since  $g_{n+1} \cong f_{n+1} \cong f_n$ , we have that  $\bar{g}$  defined by

$$\bar{g}(x) = \begin{cases} g_n(x) & \text{if } g_n(x) \leq g_{n+1}(x), \\ g_{n+1}(x) & \text{if } g_n(x) > g_{n+1}(x), \end{cases}$$

satisfies  $g_n \cong \bar{g} \cong f_n$  everywhere and  $\bar{g} < g_n$  on a set of positive measure. This contradicts (A.16). We may therefore assume without loss of generality that  $g_1 \cong g_2 \cong \dots$ . Let  $g = \lim_{n \rightarrow \infty} g_n$ . Then  $g \cong f$  and

$$\lim_{n \rightarrow \infty} \int^* f_n dp = \lim_{n \rightarrow \infty} \int g_n dp = \int g dp \cong \int^* f dp.$$

But  $f_n \cong f$  for every  $n$ , so the reverse inequality holds as well. Q.E.D.

One might hope that if  $\{f_n\}$  is a sequence of functions which are bounded below and  $f_n \uparrow f$ , then (A.15) remains valid. This is not the case, as the following example shows.

*Example A.3.* Let  $X = [0, 1)$ ,  $\mathcal{B}$  be the Borel  $\sigma$ -algebra, and  $p$  be Lebesgue measure restricted to  $\mathcal{B}$ . Define an equivalence relation  $\sim$  on  $X$  by

$$x \sim y \Leftrightarrow x - y \text{ is rational.}$$

Let  $F_0$  be constructed by choosing one representative from each equivalence class. Let  $Q = \{q_0, q_1, \dots\}$  be an enumeration of the rationals in  $[0, 1)$  with  $q_0 = 0$  and define

$$F_k = F_0 + q_k [\text{mod } 1], \quad k = 0, 1, \dots$$

Then  $F_0, F_1, \dots$  is a sequence of disjoint sets with

$$(A.18) \quad \bigcup_{k=0}^{\infty} F_k = [0, 1).$$

If for some  $n < \infty$ , we have  $p^*(\bigcup_{k=n}^{\infty} F_k) < 1$ , then  $E = \bigcup_{k=0}^{n-1} F_k$  contains a  $\mathcal{B}$ -measurable set with measure  $\delta > 0$ . For  $k = 1, \dots, n-1$ , let  $q_k = r_k/s_k$ , where  $r_k$  and  $s_k$  are integers and  $r_k/s_k$  is reduced to lowest terms. Let  $\{p_1, p_2, \dots\}$  be a sequence of prime numbers such that

$$\max_{1 \leq k \leq n-1} s_k < p_1 < p_2 < \dots$$

Then the sets  $E, E + p_1^{-1} [\text{mod } 1], E + p_2^{-1} [\text{mod } 1], \dots$  are disjoint, and by the translation invariance of  $p$ , each contains a  $\mathcal{B}$ -measurable set with measure  $\delta > 0$ . It follows that  $[0, 1)$  must contain a  $\mathcal{B}$ -measurable set of infinite measure. This contradiction implies

$$(A.19) \quad p^*\left(\bigcup_{k=n}^{\infty} F_k\right) = 1$$

for every  $n$ . Define

$$f_n = -\chi_{\bigcup_{k=n}^{\infty} F_k}, \quad n = 0, 1, \dots$$

Then  $f_n \uparrow 0$ , but (A.5) and (A.19) imply that for every  $n$

$$\int^* f_n dp = -1.$$

By a change of sign in Example A.3, we see that the second part of Theorem A.1 cannot be extended to functions which are bounded above unless additional conditions are imposed. We impose such conditions in order to prove a corollary.

**COROLLARY A.1.1.** *Let  $\{\epsilon_n\}$  be a sequence of positive numbers with  $\sum_{n=1}^{\infty} \epsilon_n < \infty$ .*

Let  $\{f_n\}$  be a sequence with

$$(A.20) \quad \lim_{n \rightarrow \infty} f_n = f,$$

$$(A.21) \quad f \leq f_n, \quad n = 1, 2, \dots,$$

$$(A.22) \quad f_n(x) \leq f(x) + \varepsilon_n \quad \text{if } f(x) > -\infty,$$

$$(A.23) \quad f_n(x) \leq f_{n-1}(x) + \varepsilon_n \quad \text{if } f(x) = -\infty, \quad n = 2, 3, \dots,$$

$$(A.24) \quad \int^* f_1 dp < +\infty.$$

Then

$$(A.25) \quad \lim_{n \rightarrow \infty} \int^* f_n dp = \int^* f dp.$$

*Proof.* From (A.20) we have  $\lim_{n \rightarrow \infty} f_n^+ = f^+$  and  $\lim_{n \rightarrow \infty} f_n^- = f^-$ . Now

$$\inf_{k \geq n} f_k^- \leq f_n^- \leq f^-$$

and

$$\inf_{k \geq n} f_k^- \uparrow f^-$$

as  $n \rightarrow \infty$ . By the theorem

$$\int^* f^- dp = \lim_{n \rightarrow \infty} \int^* \inf_{k \geq n} f_k^- dp \leq \lim_{n \rightarrow \infty} \int^* f_n^- dp \leq \int^* f^- dp,$$

so

$$(A.26) \quad \lim_{n \rightarrow \infty} \int^* f_n^- dp = \int^* f^- dp.$$

Let  $A = \{x : f(x) = -\infty\}$ . If  $p^*(A) = 0$ , then (A.21), (A.22), (A.24) and Lemma A.3(b) and (d) imply

$$\int^* f^+ dp \leq \int^* f_n^+ dp \leq 2\varepsilon_n + \int^* f^+ dp < +\infty,$$

so

$$(A.27) \quad \lim_{n \rightarrow \infty} \int^* f_n^+ dp = \int^* f^+ dp < +\infty.$$

Combine (A.26) and (A.27) to conclude (A.25). If  $p^*(A) > 0$ , then  $\int^* f^- dp = -\infty$  and (A.26) will imply (A.25) provided that

$$(A.28) \quad \int^* f^+ dp < +\infty$$

and

$$(A.29) \quad \limsup_{n \rightarrow \infty} \int^* f_n^+ dp < +\infty.$$



Conditions (A.21) and (A.24) imply (A.28). Conditions (A.21), (A.22) and (A.23) imply for every  $x \in X$ ,

$$f_n(x) \leq f_{n-1}(x) + \varepsilon_n, \quad n = 2, 3, \dots,$$

so

$$\int^* f_n^+ dp \leq 2\varepsilon_n + \int^* f_{n-1}^+ dp$$

and

$$\int^* f_n^+ dp \leq 2 \sum_{k=2}^n \varepsilon_k + \int^* f_1^+ dp.$$

The finiteness of  $\sum_{k=2}^{\infty} \varepsilon_k$  and (A.24) imply (A.29). Q.E.D.

#### REFERENCES

- [1] D. P. BERTSEKAS, *Dynamic Programming and Stochastic Control*, Academic Press, New York, 1976.
- [2] D. P. BERTSEKAS AND S. SHREVE, *Stochastic Optimal Control: The Discrete Time Case*, Academic Press, New York, 1978.
- [3] D. BLACKWELL, *Discounted dynamic programming*, Ann. Math. Statist., 36 (1965), pp. 226–235.
- [4] D. BLACKWELL, D. FREEDMAN AND M. ORKIN, *The optimal reward operator in dynamic programming*, Ann. Probability, 2 (1974), pp. 926–941.
- [5] L. D. BROWN AND R. PURVES, *Measurable selections of extrema*, Ann. Statist., 1 (1973), pp. 902–912.
- [6] C. DELLACHERIE, *Ensembles Analytiques, Capacites, Mesures des Hausdorff*, Springer-Verlag, New York, 1972.
- [7] J. DUNGUNDJI, *Topology*, Allyn and Bacon, Boston, 1966.
- [8] E. B. DYNKIN AND A. A. JUSKEVIC, *Controlled Markov Processes and their Applications*, Moscow, 1975. English translation to be published by Springer-Verlag.
- [9] D. FREEDMAN, *The optimal reward operator in special classes of dynamic programming problems*, Ann. Probability, 2 (1974), pp. 992–994.
- [10] P. R. HALMOS, *Measure Theory*, Van Nostrand, New York, 1950.
- [11] F. HAUSDORFF, *Set Theory*, Chelsea, New York, 1957.
- [12] K. HINDERER, *Foundations of Nonstationary Dynamic Programming with Discrete Time Parameter*, Springer-Verlag, New York, 1970.
- [13] B. JANKOV, *On the uniformisation of A-sets*, Doklady Acad. Nauk. SSSR, 30 (1941), pp. 591–592 (in Russian).
- [14] K. KURATOWSKI, *Topology I*, Academic Press, New York, 1966.
- [15] A. MAITRA, *Discounted dynamic programming on compact metric spaces*, Sankhyā Ser. A, 30 (1968), pp. 211–216.
- [16] P. A. MEYER, *Probability and Potentials*, Blaisdell, Waltham, MA, 1966.
- [17] J. VON NEUMANN, *On rings of operators. Reduction theory*, Ann. Math., (1949), pp. 401–485.
- [18] P. OLSEN, *Multistage stochastic programming with recourse: The equivalent deterministic problem*, this Journal, 14 (1976), pp. 495–517.
- [19] ———, *When is a multistage stochastic programming problem well-defined?* this Journal, 14 (1976), pp. 518–527.
- [20] ———, *Multistage stochastic programming with recourse as mathematical programming in an  $L_p$  space*, this Journal, 14 (1976), pp. 528–537.
- [21] K. PARTHASARATHY, *Probability Measures on Metric Spaces*, Academic Press, New York, 1967.
- [22] R. T. ROCKAFELLAR AND R. WETS, *Stochastic convex programming: Relatively complete recourse and induced feasibility*, this Journal, 14 (1976), pp. 574–589.
- [23] ———, *Stochastic convex programming: Basic duality*, Pacific J. Math., 62 (1976), pp. 173–195.
- [24] M. SCHÄL, *Conditions for optimality in dynamic programming and for the limit of n-stage optimal policies to be optimal*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete, 32 (1975), pp. 179–196.
- [25] ———, *On dynamic programming: Compactness of the space of policies*, Stochastic Processes Appl., 3 (1975), pp. 345–364.

- [26] E. SELIVANOVSKIJ, *Ob odnom klasse effektivnyh množestv (množestva C)*, Mat. Sbornik, 35 (1928), pp. 379–413.
- [27] S. SHREVE AND D. P. BERTSEKAS, *A new theoretical framework for finite horizon stochastic control*, Proc. 14th Annual Allerton Conf. on Circuit and System Theory (Univ. of Ill., Urbana, 1976), pp. 336–343.
- [28] ———, *Equivalent stochastic and deterministic optimal control problems*, Proc. 1976 IEEE Conf. on Decision and Control, Institute of Electrical and Electronics Engineers, Inc., New York, 1976, pp. 705–709.
- [29] S. SHREVE, *Dynamic programming in complete separable spaces*, Ph.D. thesis, Dept. of Math., Univ. of Ill., Urbana, IL, 1977.
- [30] ———, *Probability measures and the C-sets of Selivanovskij*, Pacific J. Math., to appear.
- [31] R. E. STRAUCH, *Negative dynamic programming*, Ann. Math. Statist., 37 (1966), pp. 871–890.