References

[1] S. C. Pincura, "A stability criterion for certain multiplicative nonlinear control systems," in *Proc. 1968 Joint Automatic Control Conf.*, Ann Arbor, Mich., pp. 787–796.

[2] N. E. Nahi and S. Partovi, "On the absolute stability of a dynamic system with a nonlinear element function of two state variables," *IEEE Trans. Automat. Contr.* (Short Paper), vol. AC-13, pp. 573–575, Oct. 1968.

[3] N. Satyanarayana, M. A. L. Thathachar, and M. D. Srinath, "Stability of a class of multiplicative nonlinear systems," *IEEE Trans. Automat. Contr.* (Short Paper), vol. AC-15, pp. 647–649, Dec. 1970.

[4] N. Satyanarayana and M. D. Srinath, "Criteria for stability of a class of multiplicative nonlinear systems," *IEEE Trans. Automat. Contr.* (Corresp.), vol. AC-16, pp. 75–76, Feb. 1971.

[5] M.-Y. Wu, "Stability criteria for a class of multiplicative time-varying nonlinear systems," *IEEE Trans. Automat. Contr.* (Corresp.), vol. AC-17, pp. 141–142, Feb. 1972.

[6] M. K. Sundareshan, "L_2-stability analysis of feedback systems via positive operator theory," Ph.D. dissertation, Dep. of Elec. Eng., Indian Inst. of Science, Bangalore, India, Nov. 1972.

[7] J. C. Willems and R. W. Brockett, "Some new rearrangement inequalities having application in stability analysis," *IEEE Trans. Automat. Contr.*, vol. AC-13, pp. 539–549, Oct. 1968.

[8] M. K. Sundareshan and M. A. L. Thathachar, "Time-domain criteria for the L_2-stability of nonstationary feedback systems," *IEEE Trans. Automat. Contr.*, vol. AC-18, pp. 80–81, Feb. 1973.

[9] ——, "L_2-stability of nonstationary feedback systems; Frequency-domain criteria," *IEEE Trans. Automat. Contr.*, vol. AC-19, pp. 217–224, June 1974.

[10] G. Zames, "On the input-output stability of time-varying nonlinear feedback systems—Part I: Conditions derived using concepts of loop gain, conicity, and positivity," *IEEE Trans. Automat. Contr.*, vol. AC-11, pp. 228–238, Apr. 1966.
——, "On the input-output stability of time-varying nonlinear feedback systems—Part II: Conditions involving circles in the frequency plane and sector nonlinearities," *IEEE Trans. Automat. Contr.*, vol. AC-11, pp. 465–476, July 1966.

[11] M. K. Sundareshan and M. A. L. Thathachar, "Improved conditions for the L_2-stability of nonstationary feedback systems," *IEEE Trans. Automat. Contr.* (Corresp.), vol. AC-18, pp. 674–675, Dec. 1973.

[12] ——, "L_2-stability of linear time-varying systems—Conditions involving noncausal multipliers," *IEEE Trans. Automat. Contr.*, vol. AC-17, pp. 504–510, Aug. 1972.

Convergence of Discretization Procedures in Dynamic Programming

DIMITRI P. BERTSEKAS, MEMBER, IEEE

***Abstract*—The computational solution of discrete-time stochastic optimal control problems by dynamic programming requires, in most cases, discretization of the state and control spaces whenever these spaces are infinite. In this short paper we consider a discretization procedure often employed in practice. Under certain compactness and Lipschitz continuity assumptions we show that the solution of the discretized algorithm converges to the solution of the continuous algorithm, as the discretization grids become finer and finer. Furthermore, any control law obtained from the discretized algorithm results in a value of the cost functional which converges to the optimal value of the problem.**

I. Introduction

It is well known that the principal framework for analysis and solution of sequential stochastic optimization problems is that of dynamic programming as developed and popularized principally by Bellman [2], [3]. In lack of an analytical solution to the problem under consideration a computer solution is required. Under these circumstances whenever some of the spaces of definition of the system are infinite, discretization of these spaces becomes necessary. In practice one hopes that if there is sufficient continuity present in the problem the computer solution will approximate closely the true solution of the problem if a suitable discretization grid with a sufficiently large number of points is used. It is thus worthwile to have precise theoretical results which guarantee convergence of various discretization procedures under concrete assumptions. Estimates of the convergence rate may also be useful. While it is unclear that such theoretical results will have significant impact on the way dynamic programming is currently employed, they will, if nothing else, help alleviate some of the nagging fears in the practitioner's mind.

The question of convergence of discretization procedures has been raised by Bellman and Dreyfus [3]. However, to the author's knowledge, no related theoretical results have appeared in the literature with the exception of a recent paper by Fox [10]. In the present paper results in a similar vein as those of Fox are obtained. The two papers are complementary however, since the analytical approach, the assumptions, the problem formulation, and the discretization procedure are all different. In particular, in [10] the case of discrete probability distributions (including deterministic problems) is ruled out in an essential way while in our case we allow the presence of discrete distributions at the outset. Also in [10] discretization is limited to the state space while we consider discretization of both state and control spaces.

Some of the ideas in the paper were clarified during the course of a tutorial with T. J. Lee. This interaction is gratefully acknowledged.

Manuscript received March 7, 1974; revised August 1, 1974 and January 17, 1975. Paper recommended by E. R. Barnes, Past Chairman of the IEEE S-CS Computational Methods Committee. This work was supported by the Joint Services Electronics Program under Contract DAAB-07-72-C-0259.

The author is with the Department of Electrical Engineering and the Coordinated Science Laboratory, University of Illinois, Urbana, Ill. 61801.

II. Discretization Procedures—Finite Horizon Problems

Consider the following dynamic programming algorithm:

$$J_N(x) = g_N(x) \qquad x \in S_N \subset R^{s_N} \tag{1}$$

$$J_k(x) = \sup_{u \in U_k(x)} E_w\{ g_k(x,u,w) + J_{k+1}[f_k(x,u,w)] | x,u,k\}$$
$$x \in S_k \subset R^{s_k}, \qquad k = 0,1,\cdots,N-1. \tag{2}$$

This algorithm is associated with a stochastic optimal control problem involving the discrete time dynamic system

$$x_{k+1} = f_k(x_k,u_k,w_k), \qquad k=0,1,\cdots,N-1, \quad x_0\text{:given} \tag{3}$$

and the cost functional

$$\mathop{E}_{w_0,\cdots,w_{N-1}}\left\{\sum_{k=0}^{N-1} g_k(x_k,u_k,w_k) + g_N(x_N)\right\}. \tag{4}$$

In the above equation x_k is the system state-element of a Euclidean space R^{s_k}, $k=0,1,\cdots,N$. The algorithm (1), (2) is defined over given *compact* subsets $S_k \subset R^{s_k}$, $k=0, 1,\cdots,N-1$. The control input at time k is denoted by u_k and is an element of some space C_k, $k=0, 1,\cdots,N-1$. In what follows we shall assume that C_k is either a subset of a Euclidean space or a finite set. The sets $U_k(x_k) \subset C_k$ are given for each $x_k \in S_k$ and represent a state-dependent control constraint.

We denote by w_k the input disturbance which is assumed to be an element of a set W_k, $k=0, 1,\cdots,N-1$. We assume in this section that *each set W_k has a finite number (say I_k) of elements*. This assumption is valid in many problems of interest, most notably in deterministic problems where the set W_k consists of a single element. In problems where the sets W_k are infinite, our assumption amounts to replacing the dynamic programming algorithm (1), (2) by another algorithm whereby the expected value (integral) in (2) is approximated by a finite sum. For most problems of interest this finite sum approximation may be justified in the sense that the resulting error can be made arbitrarily small by taking a sufficiently large number of terms in the finite sum. The reader may easily provide relatively mild assumptions under which the approximation is valid in the above sense. A discretization procedure involving the state and control spaces as well as the disturbance space, together with a corresponding convergence result may be found in an unpublished report by the author. Concerning the probabilities of the elements of W_k, denoted by $p_k^i(x_k,u_k)$, $i=1,\cdots,I_k$, we assume that they depend on the current state x_k and control u_k but they do not explicitly depend on the previous values of input disturbances $w_0,w_1,\cdots,w_{k-1}$.

The functions $g_N, g_k, f_k, k=0,1,\cdots,N-1$ in (3), (4) are given. Concerning f_k, S_k, $U_k(x)$, and W_k we make the following assumption which is necessary in order that the algorithm (1), (2) be well posed:

$$\{z | z = f_k(x,u,w), x \in S_k, u \in U_k(x), w \in W_k\} \subset S_{k+1},$$
$$k = 0,1,\cdots,N-1. \tag{5}$$

In many problems the above assumption is satisfied automatically while in other problems it is necessary to reformulate the problem so that (5) holds. We also assume that all given sets are nonempty.

We shall consider two different sets of assumptions in addition to the ones already made. In the first set of assumptions the control space C_k is assumed to be a finite set for each k. Some examples of problems in this category are hypothesis testing problems in statistics [1], [7] where a finite number of actions are of interest (accept hypothesis i, $i=1,\cdots,I$, or take another sample), asset selling and purchasing problems [16], [15] (accept the current offer, reject the offer and wait for the next), and other problems in a similar vein. In the second set of assumptions the control space C_k is assumed to be a Euclidean space. Such problems abound in stochastic control, inventory control, planning and scheduling problems, etc., and require discretization of both the state space and the control space. The reader may easily extend our analysis and results to cases where the control space is the union or the Cartesian product of a finite set and a Euclidean space.

Assumptions A

Assumption A.1: The control spaces C_k, $k=0,1,\cdots,N-1$ are finite sets and

$$U_k(x)=C_k \qquad \forall x\in S_k, k=0,1,\cdots,N-1. \tag{6}$$

Assumption A.2: The functions f_k, g_k satisfy the following Lipschitz conditions for all $x,x'\in S_k, u\in C_k, w\in W_k, k=0,1,\cdots,N-1$

$$\|f_k(x,u,w)-f_k(x',u,w)\| \leqslant L_k\|x-x'\| \tag{7}$$

$$\|g_k(x,u,w)-g_k(x',u,w)\| \leqslant M_k\|x-x'\| \tag{8}$$

$$\|g_N(x)-g_N(x')\| \leqslant M_N\|x-x'\| \qquad \forall x,x'\in S_N \tag{9}$$

where M_N, M_k, L_k, $k=0,1,\cdots,N-1$ are positive constants and $\|\cdot\|$ denotes the usual Euclidean norm.

Assumption A.3: The probabilities $p_k^i(x,u)$, $i=1,2,\cdots,I_k$ of the elements of the finite set $W_k=\{1,2,\cdots,I_k\}$ satisfy for all k the Lipschitz condition

$$|p_k^i(x,u)-p_k^i(x',u)| \leqslant N_k\|x-x'\| \qquad \forall x,x'\in S_k, u\in C_k, i\in W_k \tag{10}$$

where $N_k, k=0,\cdots,N-1$, are positive constants. (This assumption is satisfied in particular if the probabilities p_k^i do not depend on the state.)

Assumptions B

Assumption B.1: The control space $C_k, k=0,1,\cdots,N-1$ is a compact subset of a Euclidean space. The sets $U_k(x)$ are compact for every $x\in S_k$ and in addition the set

$$U_k=\bigcup_{x\in S_k} U_k(x) \qquad k=0,1,\cdots,N-1 \tag{11}$$

is compact. Furthermore the sets $U_k(x)$ satisfy

$$U_k(x)\subset U_k(x')+\{u \mid \|u\|\leqslant P_k\|x-x'\|\} \qquad \forall x,x'\in S_k, \quad k=0,1,\cdots,N-1 \tag{12}$$

where P_k are positive constants. (This last assumption, (12), is equivalent to assuming that the point-to-set map $x\to U_k(x)$ is Lipschitz continuous in the Hausdorff metric sense [9].)

Assumption B.2: The functions f_k, g_k satisfy the following Lipschitz conditions for all $x,x'\in S_k, u,u'\in U_k, w\in W_k$, $k=0,1,\cdots,N-1$

$$\|f_k(x,u,w)-f_k(x',u',w)\| \leqslant \bar{L}_k(\|x-x'\|+\|u-u'\|) \tag{13}$$

$$\|g_k(x,u,w)-g_k(x',u',w)\| \leqslant \bar{M}_k(\|x-x'\|+\|u-u'\|) \tag{14}$$

$$\|g_N(x)-g_N(x')\| \leqslant \bar{M}_N\|x-x'\| \qquad \forall x,x'\in S_N \tag{15}$$

where $\bar{M}_N, \bar{M}_k, \bar{L}_k, k=0,1,\cdots,N-1$ are positive constants.

Assumption B.3: The probabilities $p_k^i(x,u)$, $i=1,\cdots,I_k$ of the elements of the finite set $W_k=\{1,2,\cdots,I_k\}$ satisfy for all k the Lipschitz condition

$$|p_k^i(x,u)-p_k^i(x',u')| \leqslant \bar{N}_k(\|x-x'\|+\|u-u'\|) \qquad \forall x,x'\in S_k, u,u'\in U_k, i\in W_k \tag{16}$$

where $\bar{N}_k, k=0,1,\cdots,N-1$ are positive constants.

Prior to considering discretization of the dynamic programming algorithm we establish the Lipschitz continuity of the "cost-to-go" functions $J_k: S_k\to R$ of (1), (2).

Proposition 1: Under Assumptions A or Assumptions B the functions $J_k: S_k\to R$, $k=0,1,\cdots,N-1$, given by (1), (2), satisfy

$$|J_k(x)-J_k(x')| \leqslant A_k\|x-x'\| \qquad \forall x,x'\in S_k, k=0,1,\cdots,N \tag{17}$$

where $A_k, k=0,1,\cdots,N$, are some positive constants.

Proof: Under Assumptions A we have by (9) that (17) holds for $k=N$ with $A_N=M_N$. For $k=N-1$ we have that for each $x,x'\in S_{N-1}$

$$\begin{aligned}|J_{N-1}(x)-J_{N-1}(x')| &= \Big|\max_{u\in C_{N-1}} \sum_{i=1}^{I_{N-1}} \{g_{N-1}(x,u,i) + J_N[f_{N-1}(x,u,i)]\}\, p_{N-1}^i(x,u) \\ &\quad - \max_{u\in C_{N-1}} \sum_{i=1}^{I_{N-1}} \{g_{N-1}(x',u,i)+J_N[f_{N-1}(x',u,i)]\}\, p_{N-1}^i(x',u)\Big| \\ &\leqslant \max_{u\in C_{N-1}} \Big|\sum_{i=1}^{I_{N-1}} [g_{N-1}(x,u,i)p_{N-1}^i(x,u)-g_{N-1}(x',u,i)p_{N-1}^i(x',u)]\Big| \\ &\quad + \max_{u\in C_{N-1}} \Big|\sum_{i=1}^{I_{N-1}} [J_N[f_{N-1}(x,u,i)]p_{N-1}^i(x,u) \\ &\quad - J_N[f_{N-1}(x',u,i)]p_{N-1}^i(x',u)]\Big|.\end{aligned}$$

Now we use the fact that if $\alpha: S\to R$, $\beta: S\to R$ are Lipschitz continuous functions over a compact subset S of a Euclidean space with Lipschitz constants μ_α, μ_β the product function $\alpha(\cdot)\beta(\cdot)$ is also Lipschitz continuous satisfying for all $t_1,t_2\in S$

$$|\alpha(t_1)\beta(t_1)-\alpha(t_2)\beta(t_2)| \leqslant \Big[\mu_\alpha \max_{t\in S}|\beta(t)|+\mu_\beta\max_{t\in S}|\alpha(t)|\Big]\|t_1-t_2\|. \tag{18}$$

Then the earlier estimate is strengthened to yield

$$|J_{N-1}(x)-J_{N-1}(x')| \leqslant A_{N-1}\|x-x'\| \qquad \forall x,x'\in S_{N-1}$$

where A_{N-1} is given by

$$A_{N-1}=I_{N-1}(M_{N-1}+L_{N-1}A_N+B_{N-1}N_{N-1})$$

$$\begin{aligned}B_{N-1}&=\max\{|g_{N-1}(x,u,w)| \mid x\in S_{N-1}, u\in C_{N-1}, w\in W_{N-1}\} \\ &\quad +\max\{|J_N[f_{N-1}(x,u,w)]| \mid x\in S_{N-1}, u\in C_{N-1}, w\in W_{N-1}\}.\end{aligned}$$

Thus the result is proved for $k=N-1$ and similarly it is proved for every k.

We turn now to proving the result under Assumptions B. Again the result holds for $k=N$ with $A_N=M_N$. For $k=N-1$ we have for each $x,x'\in S_{N-1}$

$$\begin{aligned}|J_{N-1}(x)-J_{N-1}(x')| &= \Big|\max_{u\in U_{N-1}(x)} \sum_{i=1}^{I_{N-1}} \{g_{N-1}(x,u,i)+J_N[f_{N-1}(x,u,i)]\}\, p_{N-1}^i(x,u) \\ &\quad - \max_{u\in U_{N-1}(x')} \sum_{i=1}^{I_{N-1}} \{g_{N-1}(x',u,i)+J_N[f_{N-1}(x',u,i)]\}\, p_{N-1}^i(x',u)\Big|.\end{aligned}$$

Now using (18), the Lipschitz condition Assumptions B.2, B.3, and the above equality it is straightforward to show that

$$|J_{N-1}(x)-J_{N-1}(x')|$$

$$\leqslant \max_{u\in U_{N-1}(x)\cup U_{N-1}(x')} \Big| \sum_{i=1}^{I_{N-1}} [g_{N-1}(x,u,i)p_{N-1}^i(x,u) - g_{N-1}(x',u,i)p_{N-1}^i(x',u)]\Big|$$

$$+ \max_{u\in U_{N-1}(x)\cup U_{N-1}(x')} \Big| \sum_{i=1}^{I_{N-1}} [J_N[f_{N-1}(x,u,i)]p_{N-1}^i(x,u) - J_N[f_{N-1}(x',u,i)]p_{N-1}^i(x',u)\Big|$$

$$+2I_{N-1}\big(\overline{M}_{N-1}+\overline{L}_{N-1}A_N+\overline{B}_{N-1}\overline{N}_{N-1}\big)P_{N-1}\|x-x'\|$$

where

$$\overline{B}_{N-1}=\max\{|g_{N-1}(x,u,w)|\,|x\in S_{N-1},u\in U_{N-1},w\in W_{N-1}\}$$

$$+\max\{|J_N[f_{N-1}(x,u,w)]|\,|x\in S_{N-1},u\in U_{N-1},w\in W_{N-1}\}. \quad (19)$$

Strengthening the above estimate and using (18) and our assumptions we obtain

$$|J_{N-1}(x)-J_{N-1}(x')|\leqslant A_{N-1}\|x-x'\|$$

where

$$A_{N-1}=I_{N-1}(1+2P_{N-1})\big(\overline{M}_{N-1}+\overline{L}_{N-1}A_N+\overline{B}_{N-1}\overline{N}_{N-1}\big)$$

and the result is proved for $k=N-1$. Similarly the result is proved under Assumptions B for all k. Q.E.D.

We now proceed to describe procedures for discretizing the algorithm (1), (2) under Assumptions A and Assumptions B.

Discretization Procedure Under Assumptions A

We partition each compact set S_k into n_k mutually disjoint sets $S_k^1,S_k^2,\cdots,S_k^{n_k}$ such that $S_k=\cup_{i=1}^{n_k}S_k^i$, and select arbitrary points $x_k^i\in S_k^i$, $i=1,\cdots,n_k$. We approximate the dynamic programming algorithm (1), (2), by the following algorithm which is defined on the finite grids G_k where

$$G_k=\{x_k^1,x_k^2,\cdots,x_k^{n_k}\} \qquad k=0,1,\cdots,N-1. \quad (20)$$

We have

$$\hat{J}_N(x)=g_N(x) \qquad \text{if} \quad x\in G_N \quad (21)$$

$$\hat{J}_N(x)=g_N(x_N^i) \qquad \text{if} \quad x\in S_N^i, i=1,2,\cdots,n_N \quad (22)$$

$$\hat{J}_k(x)=\max_{u\in C_k} \mathop{E}_{w}\{g_k(x,u,w)+\hat{J}_{k+1}[f_k(x,u,w)]|x,u,k\} \text{ if } x\in G_k \quad (23)$$

$$\hat{J}_k(x)=\hat{J}_k(x_k^i) \qquad \text{if} \quad x\in S_k^i, i=1,2,\cdots,n_k, k=0,1\ldots,N-1. \quad (24)$$

The algorithm above corresponds to computing the "cost-to-go" functions $\hat{J}_k$ on the finite grid by means of the dynamic programming algorithm (21), (23), and extending their definition on the whole compact set S_k by making them constant on each section S_k^i of S_k. Thus $\hat{J}_k$ may be viewed as a piecewise-constant approximation of J_k. An alternative way of viewing the discretized algorithm (21), (23) is to observe that it corresponds to a stochastic control problem involving a certain finite state system (defined over the finite state spaces $G_0,\cdots,G_N$) and an appropriately reformulated cost functional.

Carrying out the dynamic programming algorithm (21), (23) involves a finite number of operations. Simultaneously we obtain an optimal control law as a sequence of functions $\hat{\mu}_k: G_k\to C_k$

$$\hat{\mu}_0(x),\hat{\mu}_1(x),\cdots,\hat{\mu}_{N-1}(x)$$

defined on the respective grids G_k, $k=0,\cdots,N-1$, where $\hat{\mu}_k(x_k^i)$ maximizes the right-hand side of (23) when $x=x_k^i$, $i=1,2,\cdots,n_k$. We extend the definition of this control law over the whole state space by defining for every $x\in S_k$, $k=0,1,\cdots,N-1$

$$\mu_k(x)=\hat{\mu}_k(x_k^i) \qquad \text{if} \quad x\in S_k^i, i=1,\cdots,n_k. \quad (25)$$

Thus we obtain a piecewise-constant control law $\{\mu_0,\mu_1,\cdots,\mu_{N-1}\}$ defined over the whole space. The value of the cost functional corresponding to $\{\mu_0,\mu_1,\cdots,\mu_{N-1}\}$ is denoted by $\tilde{J}_0(x_0)$, and is obtained by the last step of the algorithm

$$\tilde{J}_N(x)=g_N(x), \qquad x\in S_N \quad (26)$$

$$\tilde{J}_k(x)=\mathop{E}_{w}\{g_k[x,\mu_k(x),w]+\tilde{J}_{k+1}[f_k[x,\mu_k(x),w]]|x,\mu_k(x),k\}$$

$$x\in S_k \qquad k=0,1,\cdots,N-1. \quad (27)$$

Denote by d_s the maximum diameter of the sets S_k^i

$$d_s=\max_{k=0,1,\cdots,N-1}\ \max_{i=1,\cdots,n_k}\ \sup_{x\in S_k^i}\|x-x_k^i\|. \quad (28)$$

We shall be interested in whether $\hat{J}_k$ and $\tilde{J}_k$ converge in some sense to J_k for each k as d_s tends to zero.

Discretization Procedure Under Assumptions B

Here the state spaces S_k are discretized in the same way as under Assumptions A. In addition finite grids H_k of points in U_k are selected

$$H_k=\{u_k^1,\cdots,u_k^{p_k}\}\subset U_k, \qquad k=0,1,\cdots,N-1.$$

We assume that

$$U_k(x_k^i)\cap H_k\neq\varnothing \qquad \forall i=1,\cdots,n_k, k=0,1,\cdots,N-1$$

where $\varnothing$ denotes the empty set.

We now approximate the algorithm (1), (2) by the following algorithm:

$$\hat{J}_N(x)=g_N(x) \qquad \text{if} \quad x\in G_N \quad (31)$$

$$\hat{J}_N(x)=g_N(x_N^i) \qquad \text{if} \quad x\in S_N^i, i=1,\cdots,n_N \quad (32)$$

$$\hat{J}_k(x)=\max_{u\in U_k(x)\cap H_k} \mathop{E}_{w}\{g_k(x,u,w)+\hat{J}_{k+1}[f_k(x,u,w)]|x,u,k\}$$

$$\text{if} \quad x\in G_k \quad (33)$$

$$\hat{J}_k(x)=\hat{J}_k(x_k^i) \qquad \text{if} \quad x\in S_k^i, i=1,2,\cdots,n_k$$

$$k=0,1,\cdots,N-1. \quad (34)$$

Similarly as under Assumptions A we obtain a control law $\{\hat{\mu}_0,\cdots,\hat{\mu}_{N-1}\}$ defined on the grids $G_k, k=0,1,\cdots,N-1$ which is extended over the whole state space to yield the control law $\{\mu_0,\mu_1,\cdots,\mu_{N-1}\}$ by means of the piecewise-constant approximation (25). The corresponding value $\tilde{J}_0(x_0)$ of the cost functional is given by equations identical to (26), (27).

Again we are interested in the question whether $\hat{J}_k$ and $\tilde{J}_k$ converge in some sense to J_k for each k as both d_s and d_c tend to zero where

$$d_s=\max_{k=0,1,\cdots,N-1}\ \max_{i=1,\cdots,n_k}\ \sup_{x\in S_k^i}\|x-x_k^i\| \quad (28)$$

$$d_c=\max_{k=0,1,\cdots,N-1}\ \max_{i=1,\cdots,n_k}\ \max_{u\in U_k(x_k^i)}\ \min_{u'\in U_k(x_k^i)\cap H_k}\|u-u'\|. \quad (35)$$

This question is answered in the affirmative in the next section.

III. Convergence Results

The following proposition is the main result of this short paper. It shows convergence of the discretization procedures and justifies the employment of the control law obtained from the discretized algorithm as a suboptimal control law.

Proposition 2: There exist positive constants $\alpha_0,\alpha_1,\cdots,\alpha_N$, $\beta_0,\beta_1,\cdots,\beta_N$ (independent of the grids $G_0,\cdots,G_N,H_0,\cdots,H_{N-1}$ used in the discretization procedure) such that under Assumptions A

$$|J_k(x)-\hat{J}_k(x)|\leq \alpha_k d_s \qquad \forall x\in S_k,\ k=0,1,\cdots,N \tag{36}$$

$$|J_k(x)-\tilde{J}_k(x)|\leq \alpha_k d_s \qquad \forall x\in S_k,\ k=0,1,\cdots,N \tag{37}$$

and under Assumptions B

$$|J_k(x)-\hat{J}_k(x)|\leq \beta_k(d_s+d_c) \qquad \forall x\in S_k, k=0,1,\cdots,N \tag{38}$$

$$|J_k(x)-\tilde{J}_k(x)|\leq \beta_k(d_s+d_c) \qquad \forall x\in S_k, k=0,1,\cdots,N \tag{39}$$

where $J_k,\hat{J}_k,\tilde{J}_k,d_s,d_c$ are given by (1), (2), (21)–(24) [or (31)–(34)], (26)–(28), (35).

Proof: We first prove the proposition under Assumptions A. We have by (21), (22), $J_N(x)=\hat{J}_N(x)$ for all $x\in G_N$ while for any $x\epsilon S_N^i$, $i=1,\cdots,n_N$

$$|J_N(x)-\hat{J}_N(x)|=|g_N(x)-g_N(x_N^i)|\leq M_N\|x-x_N^i\|\leq M_N d_s. \tag{40}$$

Hence (36) holds for $k=N$ with $\alpha_N=M_N$. Also $J_N(x)=\tilde{J}_N(x)$, $\forall x\in S_N$ and hence (37) also holds for $k=N$.

To prove (36) for $k=N-1$ we have by (23) for any $i=1,2,\cdots,n_{N-1}$

$$\begin{aligned}
&|J_{N-1}(x_{N-1}^i)-\hat{J}_{N-1}(x_{N-1}^i)|\\
&=|\max_{u\in C_{N-1}} \mathop{E}_{w}\{g_{N-1}(x_{N-1}^i,u,w)\\
&\quad+J_N[f_{N-1}(x_{N-1}^i,u,w)]|x_{N-1}^i,u,N-1\}\\
&\quad-\max_{u\in C_{N-1}} \mathop{E}_{w}\{g_{N-1}(x_{N-1}^i,u,w)\\
&\quad+\hat{J}_N[f(x_{N-1}^i,u,w)]|x_{N-1}^i,u,N-1\}|\\
&\leq \max_{u\in C_{N-1}} |\mathop{E}_{w}\{J_N[f_{N-1}(x_{N-1}^i,u,w)]\\
&\quad-\hat{J}_N[f_{N-1}(x_{N-1}^i,u,w)]|x_{N-1}^i,u,N-1\}|\leq \alpha_N d_s
\end{aligned} \tag{41}$$

where the last step follows by (40).

Also for any $x\in S_{N-1}^i, i=1,\cdots,n_{N-1}$ we have using (41) and Proposition 1

$$\begin{aligned}
|J_{N-1}(x)-\hat{J}_{N-1}(x)|&=|J_{N-1}(x)-\hat{J}_{N-1}(x_{N-1}^i)|\\
&\leq|J_{N-1}(x)-J_{N-1}(x_{N-1}^i)|+|J_{N-1}(x_{N-1}^i)-\hat{J}_{N-1}(x_{N-1}^i)|\\
&\leq A_{N-1}\|x-x_{N-1}^i\|+\alpha_N d_s\leq(A_{N-1}+\alpha_N)d_s.
\end{aligned}$$

Hence (36) holds for $k=N-1$ with $\alpha_{N-1}=A_{N-1}+\alpha_N$, and similarly it is shown to hold for all k.

To prove (37) for $k=N-1$ let $x\in S_{N-1}^i$. We have by (24) and the previous inequality

$$\begin{aligned}
|J_{N-1}(x)-\tilde{J}_{N-1}(x)|&\leq|J_{N-1}(x)-\hat{J}_{N-1}(x)|+|\hat{J}_{N-1}(x)-\tilde{J}_{N-1}(x)|\\
&\leq(A_{N-1}+\alpha_N)d_s+|\hat{J}_{N-1}(x_{N-1}^i)-\tilde{J}_{N-1}(x)|.
\end{aligned}$$

For notational convenience we write $\mu_{N-1}(x)=\mu_{N-1}(x_{N-1}^i)=\mu_{N-1}^i$ [cf. (25)]. By using (23), (27), and (18), we have

$$\begin{aligned}
&|\hat{J}_{N-1}(x_{N-1}^i)-\tilde{J}_{N-1}(x)|\\
&\leq|\sum_{j=1}^{I_{N-1}} g_{N-1}(x_{N-1}^i,\mu_{N-1}^i,j)p_{N-1}^j(x_{N-1}^i,\mu_{N-1}^i)\\
&\quad-\sum_{j=1}^{I_{N-1}} g_{N-1}(x,\mu_{N-1}^i,j)p_{N-1}^j(x,\mu_{N-1}^i)|\\
&\quad+|\sum_{j=1}^{I_{N-1}} \hat{J}_N[f_{N-1}(x_{N-1}^i,\mu_{N-1}^i,j)]p_{N-1}^j(x_{N-1}^i,\mu_{N-1}^i)\\
&\quad-\sum_{j=1}^{I_{N-1}} \tilde{J}_N[f_{N-1}(x,\mu_{N-1}^i,j)]p_{N-1}^j(x,\mu_{N-1}^i)|\\
&\leq I_{N-1}(M_{N-1}+N_{N-1}\max\{|g_{N-1}(x,u,w)||x\in S_{N-1},\\
&\quad u\in C_{N-1},w\in W_{N-1}\})\|x-x_{N-1}^i\|\\
&\quad+|\sum_{j=1}^{I_{N-1}} \{\hat{J}_N[f_{N-1}(x_{N-1}^i,\mu_{N-1}^i,j)]\\
&\quad-J_N[f_{N-1}(x_{N-1}^i,\mu_{N-1}^i,j)]\}p_{N-1}^j(x_{N-1}^i,\mu_{N-1}^i)|\\
&\quad+|\sum_{j=1}^{I_{N-1}} \{J_N[f_{N-1}(x_{N-1}^i,\mu_{N-1}^i,j)]p_{N-1}^j(x_{N-1}^i,\mu_{N-1}^i)\\
&\quad-\sum_{i=1}^{I_{N-1}} J_N[f_{N-1}(x,\mu_{N-1}^i,j)]p_{N-1}^j(x,\mu_{N-1}^i)|\\
&\quad+|\sum_{j=1}^{I_{N-1}} \{J_N[f_{N-1}(x,\mu_{N-1}^i,j)]\\
&\quad-\tilde{J}_N[f_{N-1}(x,\mu_{N-1}^i,j)]\}p_{N-1}^j(x,\mu_{N-1}^i)|.
\end{aligned}$$

From the above inequality, Proposition 1, (36), (37) as proved for $k=N$, and the Lipschitz Conditions A.2 and A.3, we easily obtain

$$|\hat{J}_{N-1}(x_{N-1}^i)-\tilde{J}_{N-1}(x)|\leq\delta_N\|x-x_{N-1}^i\|\leq\delta_N d_s$$

where δ_N is a positive scalar not depending on the grid G_{N-1}. Using the above inequality in (42)

$$|J_{N-1}(x)-\tilde{J}_{N-1}(x)|\leq(A_{N-1}+\alpha_N+\delta_N)d_s. \tag{43}$$

Thus (37) holds for $k=N-1$ with $\alpha_{N-1}=A_{N-1}+\alpha_N+\delta_N$. Similarly (37) is shown to hold for all k.

The proof of (38), (39) is similar to the proof of (36), (37) and is left to the reader. Q.E.D.

IV. Infinite Horizon Problems with Discounted Cost Functionals

In this section we obtain a convergence result for the case of an infinite horizon problem with a discounted cost functional by making use of the results of the previous two sections. Consider the functional equation for $J_\infty: S\to R$

$$J_\infty(x)=\sup_{u\in U(x)} \mathop{E}_{w}\{g(x,u,w)+cJ_\infty[f(x,u,w)]|x,u\}, \qquad x\in S \tag{44}$$

where c is the discount factor, $0<c<1$. This equation is associated with a stochastic control problem over an infinite horizon involving the stationary system

$$x_{k+1}=f(x_k,u_k,w_k) \qquad k=0,1,\cdots \tag{45}$$

with x_0: given and the discounted cost functional

$$\lim_{N\to\infty} \mathop{E}_{w_0,\cdots,w_{N-1}} \left\{ \sum_{k=0}^{N-1} c^k g(x_k,u_k,w_k) \right\}. \tag{46}$$

For a discussion of such problems we refer to [11]–[13], [15]. The notation adopted corresponds in the obvious manner to the notation of Section II. Time indices are dropped in view of stationarity. All assumptions of Section II made prior to Assumptions A or B are in effect with obvious modifications to account for stationarity. Again we may introduce assumptions analogous to Assumptions A and B for $S, U(x), f, g$ and W (call them Assumptions A′ and B′) and corresponding discretization procedures. The assumption (5) in the infinite horizon setting is equivalent to assuming strong reachability of the set S as defined in [4], [5]. Many problems of practical interest must be appropriately reformulated in order for this assumption to be satisfied (see [5], [6]). The corresponding discretization grids are denoted by

$$G=\{x^1,x^2,\cdots,x^n\}\subset S=\bigcup_{i=1}^{n} S^i$$

$$H=\{u^1,u^2,\cdots,u^p\}\subset U=\bigcup_{x\in S} U(x).$$

The discretized functional equation under Assumptions A' is given by

$$\hat{J}_\infty(x)=\max_{u\in C} \mathop{E}_{w}\{ g(x,u,w)+c\hat{J}_\infty[f(x,u,w)]|x,u\} \quad \text{if} \quad x\in G \tag{47}$$

$$\hat{J}_\infty(x)=\hat{J}_\infty(x^i) \quad \text{if} \quad x\in S^i, i=1,\cdots,n. \tag{48}$$

Under Assumptions B′ the discretized functional equation becomes

$$\hat{J}_\infty(x)=\max_{u\in U(x)\cap H} \mathop{E}_{w}\{ g(x,u,w)+c\hat{J}_\infty[f(x,u,w)]|x,u\} \quad \text{if} \quad x\in G \tag{49}$$

$$\hat{J}_\infty(x)=\hat{J}_\infty(x^i) \quad \text{if} \quad x\in S^i, i=1,\cdots,n. \tag{50}$$

Under either Assumptions A′ or B′ each of the functional equations (44), (47)–(50) has a unique solution in the normed space of all bounded real valued functions over S with the sup-norm, which may be obtained from the fixed point of certain corresponding contraction mappings [8], [15]. Furthermore the solution of (47)–(50) together with associated stationary control laws can be conveniently calculated by Howard's policy iteration algorithm [11] or linear programming [14], [15], which require a finite number of arithmetical operations.

Consider now for concreteness Assumptions A′ and let $J^m(x)$ denote the optimal value function corresponding to an m-stage truncation of the infinite horizon problem, i.e., corresponding to the cost functional

$$\mathop{E}_{w_0,\cdots,w_m} \left\{ \sum_{k=0}^{m} c^k g(x_k,u_k,w_k) \right\}.$$

In view of the fact that $g(x,u,w)$ is bounded above and below over $S\times U\times W$ it may be easily shown that for every m

$$|J_\infty(x)-J^m(x)|\leq \frac{rc^m}{1-c}, \qquad \forall x\in S \tag{51}$$

where r is some positive constant. Consider also the discretized algorithm for the m-stage truncated problem. We have again

$$|\hat{J}_\infty(x)-\hat{J}^m(x)|\leq \frac{rc^m}{1-c}, \qquad \forall x\in S \tag{52}$$

while by Proposition 2

$$|J^m(x)-\hat{J}^m(x)|\leq \alpha_m d_s, \qquad \forall x\in S$$

where

$$d_s=\max_{i=1,\cdots,n} \sup_{x\in S^i} \|x-x^i\|$$

and α_m is a positive scalar depending on m but not depending on the grid G and hence on d_s. Combining (51)–(53) and using the triangle inequality we have

$$\sup_{x\in S} |J_\infty(x)-\hat{J}_\infty(x)|\leq \frac{2rc^m}{1-c}+\alpha_m d_s, \qquad \forall m\in\{1,2,\cdots\}. \tag{54}$$

It follows that given any $\epsilon>0$ there exists a $\delta>0$ such that $d_s<\delta$ implies

$$\sup_{x\in S} |J_\infty(x)-\hat{J}_\infty(x)|<\epsilon.$$

Equivalently it follows that

$$\lim_{d_s\to 0} \sup_{x\in S} |J_\infty(x)-\hat{J}_\infty(x)|=0. \tag{55}$$

It is also evident that (55) can be established under Assumptions B′ (analogous to Assumptions B of Section II) in a similar manner. Equation (55) shows the uniform convergence of the discretized algorithm and constitutes the basic result of this section.

References

[1] K. J. Arrow, D. Blackwell, and M. A. Girshick, "Bayes and minimax solutions of sequential design problems," *Econometrica*, vol. 17, pp. 213–244, 1949.
[2] R. Bellman, *Dynamic Programming*. Princeton, N.J.: Princeton Univ. Press, 1957.
[3] R. Bellman and S. Dreyfus, *Applied Dynamic Programming*. Princeton, N.J.: Princeton Univ. Press, 1962.
[4] D. P. Bertsekas, "Infinite time reachability of state-space regions by using feedback control," *IEEE Trans. Automat. Contr.*, vol. AC-17, pp. 604–612, Oct. 1972.
[5] ——, "Convergence of the feasible region in infinite horizon optimization problems," in *Proc. 1972 Joint Automatic Control Conf.*, Stanford, Calif., June 1972.
[6] ——, "Linear convex stochastic control problems over an infinite horizon," *IEEE Trans. Automat. Contr.* (Tech. Notes and Corresp.), vol. AC-18, pp. 314–315, June 1973.
[7] M. H. DeGroot, *Optimal Statistical Decisions*. New York: McGraw-Hill, 1970.
[8] E. V. Denardo, "Contraction mappings in the theory underlying dynamic programming," *SIAM Rev.*, vol. 9, no. 2, 1967.
[9] J. Dugundji, *Topology*. Boston, Mass.: Allyn and Bacon, 1966.
[10] B. L. Fox, "Discretizing dynamic programs," *J. Optimiz. Theory Appl.*, vol. 11, no. 3, pp. 228–234, 1973.
[11] R. Howard, *Dynamic Programming and Markov Processes*. Cambridge, Mass.: M.I.T. Press, 1960.
[12] A. Kauffman and R. Cryon, *Dynamic Programming*. New York: Academic, 1967.
[13] H. Kushner, *Introduction to Stochastic Control*. New York: Holt, Rinehart, and Winston, 1971.
[14] A. Manne, "Linear programming and sequential decisions," *Management Sci.*, vol. 6, no. 3, pp. 259–267, 1960.
[15] R. Ross, *Applied Probability Models with Optimization Applications*. San Francisco, Calif.: Holden-Day, 1970.
[16] D. White, *Dynamic Programming*. San Francisco, Calif.: Holden-Day, 1969.