TECHNICAL REPORT 67-7

METHODS FOR THE NUMERICAL SOLUTION OF DEGENERATE LINEAR AND NONLINEAR ELLIPTIC BOUNDARY VALUE PROBLEMS.

HAROLD J. KUSHNER

NOVEMBER, 1967

CENTER FOR DYNAMICAL SYSTEMS

METHODS FOR THE NUMERICAL SOLUTION OF DEGENERATE LINEAR AND NONLINEAR ELLIPTIC BOUNDARY VALUE PROBLEMS

by

Harold J. Kushner*

Center for Dynamical Systems
Brown University
Providence, Rhode Island

*This research was supported in part by the Air Force Office of Scientific Research, Office of Aerospace Research, United States Air Force, under AFOSR Grant No. AF-AFOSR 693-67, in part by the National Aeronautics and Space Administration under Grant No. NGR 40-002-015, and in part by the National Science Foundation under Grant No. GK-967.

Abstract

Two distinct but related results are obtained. First, an iterative method is derived for obtaining the solution of optimal control problems for Markov chains. The method usually converges much faster, and requires less computer storage space, than the methods of Howard or Eaton and Zadeh. Second, non-linear finite difference equations which 'approximate' the non-linear degenerate elliptic equation (*) arising out of the stochastic optimization problem (for diffusion processes) are found. The difference equations, and their solution, may have a meaning for the control problem even when the elliptic equation for the cost is degenerate. The iterative methods (version of the Jacobi and Gauss-Seidel) for the iterative solution of these non-linear systems are discussed and compared. Both converge to the solution (provided that the difference equations were derived using the method introduced in the paper), one (new to this paper) often much faster than the other (Theorem 2). In fact, the typical time required for the numerical solution is about the time required for a related linear problem. The method of obtaining the difference equations, and the proof of convergence of the associated iterative procedures, is illustrated by a detailed example. For a wide variety of non-linear degenerate elliptic boundary value problems, the method yields a suitable set of non-linear finite difference equations and implies that the associated iterative procedures converge.

METHODS FOR THE NUMERICAL SOLUTION OF DEGENERATE LINEAR AND NONLINEAR ELLIPTIC BOUNDARY VALUE PROBLEMS

Harold J. Kushner

The paper is concerned with the formation of, and iterative solution to, finite difference equations for non-linear degenerate boundary value problems which, loosely speaking, have the form (*) in a domain D, where boundary values $V(x) = B(x)$ on ∂D (the boundary of D) are assigned.

$$\sum_{i,j} a_{ij}(x)\frac{\partial^2 V(x)}{\partial x_i \partial x_j} + \min_u \{\sum_i f_i(x,u)\frac{\partial V(x)}{\partial x_i} + k(x,u)\} = 0$$
$$\sum_{i,j} a_{ij}(x)\frac{\partial^2 V(x)}{\partial x_i \partial x_j} + \max_u \{\sum_i f_i(x,u)\frac{\partial V(x)}{\partial x_i} + k(x,u)\} = 0 \qquad (*)$$

u is a vector parameter which varies in a compact set U, and $\{a_{ij}(x)\}$ is non-negative definite matrix which may possibly not be of full rank.

If the matrix $\{a_{ij}(x)\}$ is not of full rank, there are questions concerning the existence, uniqueness and smoothness of solutions to (*). In particular, even if there is a solution, it may possibly be insensitive to the values of $B(x)$ on some part of ∂D. Nevertheless, under the mild conditions imposed, the set of difference equations which we obtain do have a unique solution and, moreover, can be given a physical interpretation as a solution to a physical problem which is closely related to a physical problem from which the equation (*) may arise (see Part I).

In Part II of this paper, we describe a technique for choosing a difference scheme, and prove that the (Jacobi[†]) iterative procedure for the set of non-linear difference equations converges to a unique solution, regardless of the initial guess. In Section IV, we prove that, under certain conditions the corresponding Gauss Seidel iterative method for the non-linear difference system converges faster than the Jacobi method, exactly as for the linear case [1].

The method of Section II draws upon a similarity of form between certain difference approximations to (*) and certain equations obtained in the optimal control of Markov chains. Once the similarity is clear, we need only refer to the relevant control theory literature for the proofs of convergence (for the Jacobi method). Although relevant results from control theory [2], [3] have been available for several years, it appears that even workers concerned with the control of continuous time Markov processes (which often yield equations of the type (*)) do not seem to have made use of them in studying the convergence properties of numerical schemes for (*). In Part I, the relevant Markov chain results are stated. In Part III, we give a very simple development (using only simple properties of Markov chains) of the usual results on the Jacobi and Gauss-Seidel [1] matrix iterative techniques.. Aside from providing a simple alternative derivation, the point of view is useful in that it assigns an interesting intuitive interpretation to the methods, and provides some insight into the

[†]The terminology is that of Varga [1]. The exact schemes are described in Sections I and II.

problem of ordering the grid points (for the Gauss-Seidel procedure) in order to achieve a more rapid convergence. Numerical results indicate that, generally speaking, the iterative procedures for the nonlinear problems converge about as fast as the corresponding iterative procedures for the linear problems obtained by fixing the u in (*) to be some function of x with values in U.

I. Results From Stochastic Control Theory.[†]

The techniques of the sequel involve simple calculations with the transition probabilities of Markov chains. Let $X_0, X_1, \ldots$ be a sequence of random variables which take values in the state space $S = (0,1,\ldots,N)$. Each element of S is called a state. If $X_n = i$, we say that the process $\{X_m\}$ is in state i at time n. Let $P\{X_{n+m} \in \Gamma | X_o = i_o, \ldots, X_n = i_n\}$ denote the probability that X_{n+m} is in Γ(in S) given that $X_o = i_o, \ldots, X_n = i_n$. The process $\{X_m\}$ is called a Markov process (see [4] for more details) if the 'Markov property' $P\{X_{n+m} \in \Gamma | X_o = i_o, \ldots, X_n = i_n\} = P\{X_{n+m} \in \Gamma | X_n = i_n\}$ holds for all $\Gamma \subset S$, $m > 0$ and $i_o, \ldots, i_n$ in S. Define the transition probabilities $p_{ij} = P\{X_{n+1} = j | X_n = i\}$, where p_{ij} does not depend on n. Define the m-step transition probability $p_{ij}^m = P(X_{n+m} = j | X_n = i)$. The Markov property implies that (see [4])

$$p_{ij}^{n+m} = \sum_{k=0}^{N} p_{ik}^n p_{kj}^m$$

for any $n \geqq 0$, $m \geqq 0$. Let state 0 be an 'absorbing' state in that $p_{00} = 1$ (once in state zero, the process never leaves it). Suppose that there is some $\alpha > 0$ so that[††] $p_{i0}^N \geqq \alpha$ for all i. Then p_{i0}^n

---

[†] The paper has been written with the hope that the ideas would be accessible to workers in numerical analysis who are not particularly well versed in stochastic control theory. A modified version, containing numerical results and some discussion of the control origin of the boundary value problem (*) has been written and will eventually appear in an appropriate engineering journal.

[††] This condition is clearly satisfied if, and only if, for each i there is a chain of states $i_1, \ldots, i_r$ so that $p_{ii_1} p_{i_1 i_2} \cdots p_{i_{r-1} i_r} p_{i_r 0} > 0$.

is nondecreasing† in n and

$$p_{i0}^{mN} \geqq 1-(1-\alpha)^m. \qquad (2)$$

Suppose that there is a real valued function $K(\cdot)$ on S, called the 'cost' and $K(0) = 0$. Let $E_i Y$ denote the 'conditional' expectation of Y, given $X_o = i$; i.e., $E_i Y = E[Y|X_0 = i]$. Then, the average cost, defined by

$$C(i) = E_i \sum_{n=0}^{\infty} K(X_n)$$

is finite†† by (2) and satisfies††† [2], [3],

$$C(i) = \sum_{1}^{N} p_{ij} C(j) + K(i) = E_i C(X_1) + K(i) \qquad (3)$$

---

† p_{i0}^n is nondecreasing since state 0 is absorbing and $p_{i0}^{m+n} \geqq p_{i0}^n p_{00}^m \geqq p_{i0}^n$. (2) is true for $m = 1$. Suppose it is true for $m = r$. Then, by the Markov property, $p_{i0}^{(r+1)N} = p_{i0}^{rN} p_{00}^N + \sum_{j=1}^{N} p_{ij}^{rN} p_{j0}^N \geqq 1-(1-\alpha)^r + \alpha \sum_{j=1}^{N} p_{ij}^{rN} \geqq 1-(1-\alpha)^r + \alpha(1-\alpha)^r = 1-(1-\alpha)^{r+1}$.

††Let $K = \max_i K(i)$. Then $C(i) \leqq K \sum_{n=0}^{\infty} P(X_n \neq 0)$. But, by (2), the sum is finite.

†††$C(i) = K(i) + E_i \sum_{n=1}^{\infty} K(X_n)$. But the latter sum equals $E_i C(X_1)$ which, in turn equals $\sum_{j=1}^{N} p_{ij} C(j)$.

To each state $i = 1,\ldots,N$ we now associate a parameter u_i taking values in a compact set U. Let the $p_{ij}(u_i)$ and $K(i,u_i)$ be continuous in the u_i. Let $K(0,u) \equiv 0$ and $p_{00}(u) \equiv 1$. u_i is called a control, and $\pi \equiv (u_1,\ldots,u_N)$ a control policy [2], [3]. To each fixed π (which fixes the $\{p_{ij}(u_i)\}$). There is a corresponding cost which we denote as $\{C^\pi(i)\}$. The object of stochastic control theory (for discrete Markov chains) is the selection of the π (or, equivalently, the transition probabilities $\{p_{ij}(u_i),\ u_i \in U\}$) which minimizes† the cost ($u(X_n)$ is written in lieu of u_{X_n})

$$C^\pi(i) \equiv E_i^\pi \sum_0^\infty K(X_n, u(X_n)). \tag{4}$$

Further, let $p_{ij}^n(\pi_0,\ldots,\pi_{n-1}) = P\{X_n = j \mid X_0 = i,$ and π_k used at time $k, k = 0,\ldots,n-1\}$. Suppose that

$$\inf_{\{\pi_k\}} p_{i0}^N(\pi_0,\ldots,\pi_{N-1}) \geq \alpha > 0.$$

This condition holds if to each i there is some chain of states $i_1,\ldots,i_r$ so that

$$\inf_{\{u_j\}} p_{ii_1}(u_i)p_{i_1i_2}(u_{i_1})\cdots p_{i_{r-1}i_r}(u_{i_{r-1}})p_{i_r0}(u_{i_r}) > 0.$$

---

†The minimization could be replaced by a maximization, as noted in the remark following Theorem 2. For definiteness, except in Example 2, a minimization is used.

Write $p_{ij}^n(\pi,\ldots,\pi)$ as $p_{ij}^n(\pi)$. Under our assumptions, the optimum control exists† and the corresponding minimum cost $C(i) \equiv \inf_{\pi} C^{\pi}(i)$ satisfies (see [2], [3] and Section IV);

$$C(i) = \min_{u_i \in U} \{ \sum_1^N p_{ij}(u_i)C(j) + K(i,u_i)\}. \tag{5}$$

Furthermore, for <u>any</u> initial $C_o(i)$, the sequence $C_n(i)$ defined by

$$C_{n+1}(i) = \min_{u_i \in U} \{ \sum_1^N p_{ij}(u_i)C_n(i) + K(i,u_i)\}. \tag{6}$$

converges [3] to $C(i)$. This is the crucial fact for Part II. Some insight into the reasonably simple derivation will be given in Parts III and IV. In Part IV, we extend the convergence theorem to the Gauss-Seidel form

$$\tilde{C}_{n+1}(i) = \min_{u_i \in U} \{ \sum_{j=1}^{i-1} p_{ij}(u_i)\tilde{C}_{n+1}(j) + \sum_{j=i}^{N} p_{ij}(u_i)\tilde{C}_n(j) + K(i,u_i)\}, \tag{7}$$

and give conditions under which (7) is strictly better than (6). (7) always converges if (6) does, and to the same unique solution. Furthermore, (7) has only about half the memory requirements of (6).

<u>A Remark on the Physical Interpretation of</u> (*) <u>and</u> (5).

Suppose that a random diffusion process is governed by the (Itô) vector stochastic differential equation [5]

---

†Existence is essentially proved in the reference [3] and is proved in Section IV. As noted there, the proof can easily be changed to correspond to the 'maximization' problem.

$$dx = f(x,u)dt + v(x)dz, \qquad a_{ij}(x) = \frac{1}{2}\sum_k v_{ik}(x)v_{kj}(x).$$

$u = u(x)$ is a 'control', taking values in U. Define the cost, corresponding to a control u,

$$C^u(x) = E_x B(X_\tau) + E_x \int_0^\tau k(x_s, u(x_s))ds,$$

where E_x is the expectation conditioned in the event that the initial condition is $x \in D$, and τ is the first arrival time on ∂D. Purely formally, define

$$V(x) = \inf_u C^u(x),$$

Then $V(x)$ formally satisfies (*). Under certain conditions, there is an optimal $u = u^o$, and $C^{u^o}(x) = V(x)$ which, in turn, satisfies (*). See [6], [7]. In Examples 1 and 2, we show how to determine difference approximations to (*) which correspond to a form such as (5). In such a case, whether or not the original optimal control problem is meaningful, it turns out that the solution to (5) (which exists and is unique) is the solution to an optimization problem for a Markov chain closely related to the process X_t.

II. A Method for the Numerical Solution of Elliptic Nonlinear Boundary Value Problems.

Two simple examples will suffice to illustrate the method.

Example 1. Let $h > 0$ be the step size and e_i the unit vector in the i^{th} coordinate direction. For notational convenience, it is supposed that the step size is the same in each direction. The method is clearly valid if the step sizes are different; of course, the conditions (9) would then be different. Let

$$a_{11} \frac{\partial^2 V}{\partial x_1^2} + 2a_{12} \frac{\partial^2 V}{\partial x_1 \partial x_2} + a_{22} \frac{\partial^2 V}{\partial x_2^2} + f\frac{\partial V}{\partial x_1} - \left|\frac{\partial V}{\partial x_2}\right| + k(x) = 0 \qquad (8)$$

or, equivalently,

$$\sum a_{ij} \frac{\partial^2 V}{\partial x_i \partial x_j} + f\frac{\partial V}{\partial x_1} + \min_{|u| \leq 1} u\frac{\partial V}{\partial x_2} + k(x) = 0 \qquad (8')$$

be defined in the rectangle $R = \{x: -A_1 \leqq x_1 \leqq A_1, -A_2 \leqq x_2 \leqq A_2\}$, $n_1 h = 2A_1$, $n_2 h = 2A_2$. Then there are $(n_1-1)(n_2-1)$ points of the grid internal to R and $(n_1+1)(n_2+1) \equiv N$ total grid points. The a_{ij} and f are functions of x and we only suppose that†

$$\begin{aligned} &a_{11} \geqq 0, \ a_{12} \geqq 0 \\ &a_{22} > 2a_{12} + h/2 \\ &V(x) = B(x) \quad \text{on} \quad \partial R. \end{aligned} \qquad (9)$$

† If $a_{12} < 0$, the difference scheme (10) for $\partial^2 V/\partial x_1 \partial x_2$ must be changed.

First, a difference scheme for

$$\sum a_{ij}\partial^2 V/\partial x_i \partial x_j + f\partial V/\partial x_1 + u\partial V/\partial x_2 + k(x) = 0 \qquad (8'')$$

will be given, where u is some function of x satisfying $|u| \leq 1$. Then the difference equations will be identified with a particular optimal stochastic control problem, and convergence of the corresponding iterative procedures (6) or (7) shown - for any initial guess. Finally, the limit of the sequence of iterations will be the unique solution of a non-linear difference scheme for (8). The reason for our particular choice of difference approximations will be seen in the sequel. At each grid point internal to R, let

$$\begin{aligned} \partial^2 V/\partial x_1^2 &\to [V(x+e_1h)+V(x-e_1h)-2V(x)]/h^2 \\ \partial^2 V/\partial x_1 \partial x_2 &\to [V(x+e_2h+e_1h)-V(x+e_2h)-V(x)+V(x-e_1h)]/h^2 \qquad (10) \\ \partial^2 V/\partial x_2^2 &\to [V(x+e_2h)+V(x-e_2h)-2V(x)]/h^2. \end{aligned}$$

$$\partial V/\partial x_2 \to [V(x+e_2h)-V(x-e_2h)]/2h \qquad (11)$$

$$\partial V/\partial x_1 \to [V(x+e_1h)-V(x)]/h \quad \text{if} \quad f(x) \geq 0 \qquad (12a)$$

$$\partial V/\partial x_1 \to [V(x)-V(x-e_1h)]/h \quad \text{if} \quad f(x) < 0. \qquad (12b)$$

Then, using the symbol $V(x)$ for the solution of the difference equations also, the substitution of (10)-(12) into (8") gives

$$\begin{aligned} 0 = {} & a_{11}[V(x+e_1h)+V(x-e_1h)-2V(x)]+2a_{12}[V(x+e_2h+e_1h) \\ & \qquad -V(x+e_2h)-V(x)+V(x-e_1h)] \\ & + a_{22}[V(x+e_2h)+V(x-e_2h)-2V(x)]+k(x)h^2 \qquad (13) \\ & + hf\left\{\begin{matrix} V(x+e_1h)-V(x) \\ V(x)-V(x-e_1h) \end{matrix}\right\} + \frac{hu}{2}[V(x+e_2h)-V(x-e_2h)] \end{aligned}$$

where the upper entry is used if $f \geq 0$ and vice versa. Collecting terms and dividing by the coefficient of $V(x)$ yields, for x internal to R,

$$V(x) = \frac{V(x+e_1h)}{Q}\begin{Bmatrix} a_{11} + h|f| \\ a_{11} \end{Bmatrix} + \frac{V(x-e_1h)}{Q}\begin{Bmatrix} a_{11} + 2a_{12} \\ a_{11} + 2a_{12} + h|f| \end{Bmatrix}$$

$$+ \frac{V(x+e_2h)}{Q}[a_{22}-2a_{12}+hu/2] + \frac{V(x-e_2h)}{Q}[a_{22}-hu/2] \qquad (14)$$

$$+ V(x+e_2h+e_1h)\frac{2a_{12}}{Q} + \frac{k(x)h^2}{Q},$$

where

$$Q = 2a_{11} + 2a_{22} + 2a_{12} + h|f|.$$

Now we relate (14) to a control problem. The coefficients of the $V(\cdot)$ on the right of (14) have the character of probabilities; they are non-negative and sum to unity. Now, order the N grid points on R, and associate a state of a Markov process with each point. We use both notations 'x' and 'i' to refer to the grid points. If x is internal to R, define $K(x,u)$ by $h^2k(x)/Q(x)$. If $x \in \partial R$, define $K(x,u)$ by $B(x)$. _Define an absorbing state_ 0 _and define_ $p_{x0} = 1$ for any x on ∂R, and any control $u(x)$. For x internal to R, the $p_{xy}(u)$ are given by the coefficients of $V(y)$ on the right of (14). For each x internal to R, only $p_{x,x+e_2h}(u)$ and $p_{x,x-e_2h}(u)$ depend on the control parameter u, and $u \in U \equiv [-1,1]$. A control policy $\pi = (u_1,\ldots,u_N)$ for this problem is merely an association of some number in $[-1,1]$ with each state on R.

For the moment suppose that the transition probabilities satisfy

$$\inf_{\{\pi_j\}} p^N_{x0}(\pi_0,\ldots,\pi_{N-1}) \geq \alpha > 0. \tag{15}$$

Then by the results in Part I, the π minimizing the $C(x) = E_{x_0}\sum^{\infty} K(X_n,u(X_n))$ exists and satisfies

$$\begin{aligned} C(x) = \min_{|u|\leq 1} \Bigg\{ & \frac{C(x+e_1h)}{Q}\begin{Bmatrix} a_{11} + h|f| \\ a_{11} \end{Bmatrix} + \frac{C(x-e_1h)}{Q}\begin{Bmatrix} a_{11} + 2a_{12} \\ a_{11} + 2a_{12} + h|f| \end{Bmatrix} \\ & + \frac{C(x+e_2h)}{Q}[a_{22}-2a_{12}+hu/2] + \frac{C(x-e_2h)}{Q}[a_{22}-hu/2]+ \\ & + C(x+e_2h+e_1h)\frac{2a_{12}}{Q} + \frac{k(x)h^2}{Q} \Bigg\} \end{aligned} \tag{16}$$

for x internal to R. Since $p_{x0} \equiv 1$ and $K(x,U(x)) = B(x)$ for x on ∂R, we assign

$$C(x) = B(x) \equiv K(x,U(x)),$$

for x on ∂R.

Furthermore, replacing $C(\cdot)$ on the left and right of (16) by $C_{n+1}(\cdot)$ and $C_n(\cdot)$, resp, the resulting iterative process converges to the unique solution of (16) for any $C_o(\cdot)$. Performing the minimization in (16) yields $u(x) = -\text{sign}[C(x+e_2h) - C(x-e_2h)]$. It is clear that, if $u(x) = -\text{sign}[(V(x+e_2h) - V(x-e_2h)]$ is substituted into (13), then the $C(x)$ of (16) satisfy (13). There can be no other $V(x)$ satisfying (13), by virtue of the uniqueness of $C(x)$. Now, writing the last term in (13) as $-h|V(x+e_2h)-V(x-e_2h)|/2$, we note that (13) is a difference equation for (8). Since $C(x)$ also satisfies the boundary conditions, the demonstration of convergence is complete, except for (15).

Property (15) is proved as follows. Fix x and let $x + e_2nh$ be on the boundary ∂R. Denote the state $x + e_2ih$ by i, $i \geq 1$. Since $p_{x,x\pm e_2h}(u) \geq (a_{22}-2a_{12}-h/2)/Q \geq \gamma > 0$, and $p_{y0} = 1$ for y on ∂R, we have

$$p^N_{x0}(\pi_o,\dots,\pi_{N-1}) \geq \min p_{x1}(u(x))p_{12}(u(1))\cdots p_{n-1,n}(u(n-1))p_{n0}(u(n)) \tag{17}$$

$$\geq [a_{22}-2a_{12}-h/2]^n \geq [a_{22}-2a_{12}-h/2]^N = \gamma^N = \alpha.$$

The difference schemes (10), (11), (12) are selected to assure that the <u>coefficients in</u> (14) <u>would be non-negative and sum to at most unity</u>. Other choices are certainly possible, but it must be noted that the difference equations corresponding to any arbitrary difference scheme may not correspond to a control problem and, consequently, the above proof of convergence may not be valid. Finally, if $k(x) \geq \epsilon > 0$ in D, Theorem 2 implies that the Gauss-Seidel procedure is strictly preferable to the Jacobi procedure - at least if $C_0(i) = 0$. Experimentally, it has been found to be generally preferable.

<u>Example</u> 2. We would like to consider the 2-dimensional problem (18), in the domain R of Example 1

$$\sum a_{ij}\frac{\partial V^2}{\partial x_i \partial x_j} + f\frac{\partial V}{\partial x_1} + (\partial V/\partial x_2)^2/4 + k(x) = 0 \tag{18}$$

or, equivalently,

$$\sum a_{ij}\partial^2 V/\partial x_i \partial x_j + f\frac{\partial V}{\partial x_1} + \max_u\,[u\,\frac{\partial V}{\partial x_2} - u^2] + k = 0. \tag{18'}$$

In order to prove convergence, we will need a bound on the $(\partial V/\partial x_2)^2$ term. Hence, we will actually obtain a numerical scheme for

$$\sum a_{ij}\frac{\partial^2 V}{\partial x_i \partial x_i} + f\frac{\partial V}{\partial x_1} + \max_{|u|\leq c}\,[u\,\frac{\partial V}{\partial x} - u^2] + k = 0. \tag{19}$$

Bounding by c linearizes the $(\partial^2 V/\partial x_2^2)/4$ term for large values: and the maximized term in (18') equals

$$\begin{array}{ll} (\partial^2 V_2/\partial x_2)/4 & \text{if } |\partial V/\partial x_2| \leq 2c \\ c|\partial V/\partial x_2| - c^2 & \text{if } |\partial V/\partial x_2| > 2c \end{array}$$

The difference schemes (10)-(12) will be used. Then for any fixed $u = u(x)$,

$$V(x)Q(x) = V(x+e_1h)\left\{\begin{array}{l} a_{11} + h|f| \\ a_{11} \end{array}\right\} \tag{20}$$

$$+ V(x-e_1h)\left\{\begin{array}{l} a_{11} + 2a_{12} \\ a_{11} + 2a_{12} + h|f| \end{array}\right\} + V(x+e_2h)[a_{22}-2a_{12}+hu/2]$$

$$+ V(x-e_2h)[a_{22}-hu/2] + V(x+e_2h+e_1h)2a_{12} - u^2h^2 + kh^2 = 0.$$

Let

$$a_{22} > hc/2 + 2a_{12},\ a_{11} \geqq 0 .$$

Then the coefficients on the right of (20) are non-negative and sum to $Q(x)$. Using the same identifications with quantities in the control problem as was done in Example 1, it can easily be verified that (15) holds here also. The optimal control problem corresponding to (20) is represented by (21):

$$\begin{aligned} V(x) = {} & \frac{V(x+e_1h)}{Q}\left\{\begin{matrix} a_{11} + h|f| \\ a_{11} \end{matrix}\right\} + \frac{V(x-e_1h)}{Q}\left\{\begin{matrix} a_{11} + 2a_{12} \\ a_{11} + 2a_{12} + h|f| \end{matrix}\right\} \\ & + \frac{V(x+e_2h)}{Q}[a_{22}-2a_{12}] + \frac{V(x-e_2h)}{Q}a_{22} \qquad (21) \\ & + \max_{|u|\leq c} \frac{1}{Q}\left\{\frac{h}{2}u\cdot V(x+e_2h) - \frac{h}{2}u\cdot V(x-e_2h) - u^2h^2\right\} + V(x+e_2h+e_1h)\frac{2a_{12}}{Q} \\ & + \frac{kh^2}{Q} , \end{aligned}$$

where the maximized term equals the smallest of

$$[V(x+e_2h)-V(x-e_2h)]^2/16$$
$$ch[|V(x+e_2h)-V(x-e_2h)|/2-ch].$$

Upon substituting this term in (21) we get a difference scheme for

(18'). Replacing V in the left and right of (21) by V_{n+1} and V_n, resp., gives an iterative procedure which converges to the unique solution of (21) for any V_o. The details are exactly as in Example 1.

III. A Simple Probabilistic Derivation of Some Results in Matrix Iterative Analysis.

Much of matrix iterative analysis is devoted to the system $AY = b$, where A is diagonally dominant and strictly diagonally dominant for some row. Let $A = -D + F$, where D is the diagonal part of A and the elements of D and F are non-negative. Such systems commonly arise in the numerical solution of boundary value problems. Let A be $N \times N$ and write $D^{-1}F = P = \{p_{ij}\}$ and $-D^{-1}b = K = \{K(i)\}$. Two well known iterative techniques for the numerical solution of $Ax = b$ are the Jacobi

$$Y_{n+1} = D^{-1}FY_n - D^{-1}b$$

or, equivalently,

$$Y_{n+1}(i) = \sum_{j=1}^{N} p_{ij}Y_n(j) + K(i) \tag{22}$$

and the Gauss-Seidel, which can be written as

$$\tilde{Y}_{n+1}(i) = \sum_{j=1}^{i-1} p_{ij}\tilde{Y}_{n+1}(j) + \sum_{j=i}^{N} p_{ij}\tilde{Y}_n(j) + K(i), \tag{23}$$

and their relative merits are well known [1]. Here we offer a probabilistic proof of their relative properties which, aside from its simplicity, provides some insight into the preferred ordering of the rows of A and the choice of a difference scheme, and also leads to a proof that (7) is at least as good as (6) for the non-linear problem.

The p_{ij} have the character of transition probabilities. They are non-negative and $\sum_{j=1}^{N} p_{ij} \le 1$. As before, let $X_o, X_1, \ldots$ be a Markov chain with state space $S = (0,1,\ldots,N)$ and transition probabilities p_{ij}. Define $p_{00} = 1$, $p_{i0} = 1 - \sum_{j=1}^{N} p_{ij}$ and $k(0) = 0$. Let $y = (y_1,\ldots,y_N)$ and define the norm $\|y\| = \max_i |y_i|$.

By successive substitutions, (23) can be written as

$$\tilde{Y}_{n+1}(i) = \sum_{j=1}^{N} q_{ij}\tilde{Y}_n(j) + \tilde{K}(i) \qquad (24)$$

where the q_{ij} and $\tilde{K}(i)$ (obtained by substituting the $\tilde{Y}_{n+1}(j)$, $j < i$, of (24) into (23), reordering and equating coefficients with (24)) are

$$\begin{aligned} q_{ij} &= p_{ij} + \sum_{k=1}^{i-1} p_{ik}q_{kj} && j \ge i \\ &= \sum_{k=1}^{i-1} p_{ik}q_{kj} && j < i \qquad (25) \\ &= p_{ij} && i = 1 \end{aligned}$$

$$\tilde{K}(i) = \sum_{j=1}^{i-1} p_{ij}\tilde{K}(j) + K(i)$$

The $\{q_{ij}\}$ also have the character of transition probabilities for some Markov chain $\tilde{X}_o, \ldots, \tilde{X}_n, \ldots$ on the state space $S = (0,1,\ldots,N)$. Define $q_{i0} = 1 - \sum_1^N q_{ij}$ and $q_{00} = 1$.

Now, let $TY = (T_1Y, \ldots, T_NY)$ be a map of R^N into R^N (Euclidean N-space) with i^{th} component $T_iY = \sum_{j=1}^N p_{ij}Y_j + K(i)$. Then

$$\|TY-TZ\| = \max_i \left| \sum_{j=1}^N p_{ij}(Y_j-Z_j) \right|$$

and, by iterating,

$$\|T^nY-T^nZ\| = \max_i \left| \sum_{j=1}^N p_{ij}^n(Y_j-Z_j) \right|$$

(where p_{ij}^n is, again the n step transition probability $P\{X_{k+n} = j | X_k = i\}$, and is the i,j^{th} element of the matrix P^n). Thus, if $\max_i \sum_{j=1}^N p_{ij}^n \to 0$ as $n \to \infty$, (22) (or, equivalently, $PY + K = TY$) represents a contraction mapping with a unique fixed point. Similarly, let $QY = (Q_1Y, \ldots, Q_NY)$ with $Q_iY = \sum_{j=1}^N q_{ij}Y_j + \tilde{K}(i)$. Then

$$\|Q^nY-Q^nZ\| = \max_i \left| \sum_{j=1}^N q_{ij}^n(Y_j-Z_j) \right|$$

and it is clear that the relative rates of convergence, to the fixed point, of the Y_n and $\tilde{Y}_n$ whose components are given by (22) and (23) depend on the relative rates of convergence to zero of $\sum_{j=1}^N p_{ij}^n$ and

$\sum_{j=1}^{N} q_{ij}^{n}$, resp., as $n \to \infty$. Theorem 1 (see (28) in particular) is the probabilistic proof that (23) is preferable to (22).

Theorem 1. Suppose that the rows can be and are ordered so that

$$\sum_{j=1}^{N} p_{1j} < 1$$

$$\max \{p_{i1}, \ldots, p_{i,i-1}\} > 0 \qquad (i > 1) \tag{26}$$

$$p_{ij} > 0 \quad \text{for some} \quad j(i) > 1.$$

Then (22) and (23) converge to the same unique solution of

$$C(i) = \sum_{j=1}^{N} p_{ij} C(j) + K(i). \tag{27}$$

In addition

$$\sum_{j=1}^{N} q_{ij}^{n} < \sum_{j=1}^{N} p_{ij}^{n}, \tag{28}$$

if either $n > 1$ or $i > 1$. If $n = 1$ and $i = 1$ hold simultaneously, then there is strict equality in (28). For the Markov chain introduced in this section, we have

$$C(i) \equiv E_i \sum_{o}^{\infty} K(X_n) < \infty \tag{29}$$

If $Y_o(i) \equiv 0$ in (22), then $Y_m(i) = E_i \sum_{o}^{m-1} K(X_n)$.

Remark. A weakening of (26) is possible, but (26) allows for a proof in an unburdened notation. The condition (26) means simply that (referring to the original form $AY = b$) $a_{11} > \sum_{2}^{N} a_{1j}$, $a_{ij} > 0$ for some $j < i$ and $a_{ij} > 0$ for some $j > 1$. These conditions are usually satisfied by the matrix problems arising from boundary value problems.

Remark. Since the effective norms of the contraction operators T^n or Q^n are the probabilities of not being absorbed ($\sum_{1}^{N} p_{ij}^{n} = 1 - p_{i0}^{n}$ or $\sum_{1}^{N} q_{ij}^{n} = 1 - q_{i0}^{n}$), if there is a choice of difference schemes, it is preferable to choose one with the higher absorbtion probabilities for the range of n of interest. In fact, the absorbtion probabilities can often be estimated by merely observing the directed graph corresponding to the process transition probabilities. The probabilities of moving toward the boundary ∂R should be maximized by the choice of the difference scheme. Furthermore, the q_{io}^{n} depend on the ordering of the states, and the graph may yield some useful information concerning preferable orderings.

We note also that the method is applicable if the boundary conditions are of the form $\partial V(x)/\partial n + \beta V(x) = B(x)$ where n is an outward normal and $\beta \neq 0$. Then, the process, on reaching the boundary, has some probability of being absorbed at state 0 at the next step (proportional to β) and some probability of being reflected back into the interior of R. The analysis is quite similar to that given here.

Proof. (29) and the line below it are obvious and we will only prove (28). (25) and (26) imply that

$$\sum_{j=1}^{N} p_{1j} = \sum_{j=1}^{N} q_{1j} < 1. \tag{30}$$

Let $i = m = 2$ in (25). Then, for this value of m, and using (see (26)) the fact that $p_{21} > 0$, and (30), we obtain

$$\begin{aligned}\sum_{j=1}^{N} q_{mj} &= \sum_{j=m}^{N} p_{mj} + \sum_{j=1}^{N}\sum_{k=1}^{m-1} p_{mk}q_{kj} \\ &= \sum_{j=m}^{N} p_{mj} + \sum_{k=1}^{m-1} p_{mk}\Big(\sum_{j=1}^{N} q_{kj}\Big) \\ &< \sum_{j=m}^{N} p_{mj} + \sum_{k=1}^{m-1} p_{mk} < \sum_{j=1}^{N} p_{mj}\end{aligned} \tag{31}$$

(The latter term on the second line of (31) is merely $p_{21}\sum_{j=1}^{N} p_{1j}$.) Now, supposing that (28) is true for $n = 1$ and $i = m-1$ we show that it is true for $n = 1$ and $i = m$. The first two lines of (31) are an identity for any m. The third line holds also since, by the induction hypothesis and (26), $\sum_{j=1}^{N} q_{kj} < 1$ for $k = 1,\ldots,m-1$, and $p_{mk} > 0$ for some $k \leq m-1$.

Thus, by induction, (28) is true for $n = 1$. We now prove it for $n > 1$. Since the q_{ij} are transition probabilities for a Markov chain, we have†

---

† (32) uses only (25) and the Markov property $q_{ij}^{n+m} = \sum_{k=0}^{N} q_{ik}^{n} q_{kj}^{m}$ [4]. In (32), the summation need be over $j \in [1,N]$ only since $j \neq 0$ and $q_{0j}^{n} = 0$ by the previous definition $q_{00} = 1$.

$$F(i,r+1) \equiv \sum_{j=1}^{N} q_{ij}^{r+1} = \sum_{k=1}^{N} \sum_{j=1}^{N} q_{ik} q_{kj}^{r} = \sum_{k=1}^{N} q_{ik} \left(\sum_{j=1}^{N} q_{kj}^{r} \right). \tag{32}$$

$$\sum_{j=1}^{N} q_{sj}^{r+1} = \sum_{k=1}^{N} \sum_{j=1}^{N} q_{sk}^{r} q_{kj} \leqq \sum_{k=1}^{N} q_{sk}^{r}. \tag{33}$$

Substituting (25) for the left hand q_{ik} in the far right term of (32) gives

$$F(i,r+1) = \sum_{k=i}^{N} p_{ik} \left(\sum_{j=1}^{N} q_{kj}^{r} \right) + \sum_{k=1}^{N} \left(\sum_{s=1}^{i-1} p_{is} q_{sk} \right) \sum_{j=1}^{N} q_{kj}^{r}. \tag{34}$$

The last term on the right side of (34) is $\sum_{s=1}^{i-1} p_{is} \sum_{j=1}^{N} q_{sj}^{r+1}$. Then using the bound (33) in (34) and rearranging (34) gives

$$F(i,r+1) \leqq \sum_{k=1}^{N} p_{ik} \left(\sum_{j=1}^{N} q_{kj}^{r} \right). \tag{35}$$

Let $r = 1$. Then since $\sum_{j=1}^{N} q_{kj} \leqq \sum_{j=1}^{N} p_{kj}$ with a strict inequality for $k > 1$, and $p_{ik} > 0$ for some $k > 1$ (by hypothesis) we have

$$F(i,2) < \sum_{k=1}^{N} p_{ik} \sum_{j=1}^{N} p_{kj} = \sum_{j=1}^{N} p_{ij}^{2}.$$

Then (28) is valid for $n = 2$. Suppose it is valid for $n = r$. Then by the hypothesis that $p_{ik} > 0$ for some $k \geqq 1$ and the induction hypothesis, (35) is bounded by

$$F(i,r+1) < \sum_{k=1}^{N} p_{ik}\left(\sum_{j=1}^{N} p_{kj}^{r}\right) = \sum_{j=1}^{N} p_{ij}^{r+1},$$

and the theorem is proved.

IV. The Gauss-Seidel Method for the Non-Linear Problem.

Let $\pi_n = (u_1^n,\ldots,u_N^n)$, and $\tilde{\pi}_n = (\tilde{u}_1^n,\ldots,\tilde{u}_N^n)$ be the policies determined by the minimizing operations in (6), and (7), resp., at the n^{th} iteration. Let $P(\pi_n) = \{p_{ij}(u_i^n)\}$, and $K(\pi_n) = \{K(i,u_i^n)\}$. Then $C_{n+1} = P(\pi_n)C_n + K(\pi_n)$. Similarly, once the $\tilde{\pi}_n$ is determined by (7), we may write $\tilde{C}_{n+1} = Q(\tilde{\pi}_n)\tilde{C}_n + \tilde{K}(\tilde{\pi}_n)$, where $\tilde{Q}(\tilde{\pi}_n) = \{q_{ij}(\tilde{\pi}_n)\}$, $\tilde{K}(\tilde{\pi}_n) = \{\tilde{K}(i,\tilde{\pi}_n)\}$, and the $q_{ij}(\tilde{\pi}_n)$ and $\tilde{K}(i,\tilde{\pi}_n)$ are determined by (25) in terms of the $\{p_{ij}(\tilde{u}_i^n)\}$ and $\{K(i,\tilde{u}_i^n)\}$.

Let $\pi' = (u_1',\ldots,u_N')$ be an arbitrary policy. Define I_0 as the set of states i, $1 \leq i \leq N$, which satisfies the condition: for each π' and i in I_0, there is a chain of states $i > i_1 > i_2 > \cdots > i_r \geq 1$ (depending on π' and i) for which $K(i_r,u_{i_r}') > 0$ and $p_{ii_1}(u_1')\cdots p_{i_{r-1}i_r}(u_{i_{r-1}}') > 0$. Thus I_0 contains those states which can reach (for any control) some lower state, via a 'decreasing' chain, where the lower state has a positive cost for the chosen control. (Both the chain and the lower state may depend on the control.) Define the set of states I_n, $n \geq 1$, as those satisfying the condition: i is in I_n if for any control u_i', there is some state j in I_{n-1} (depending on u_i') for which $p_{ij}(u_i') > 0$. Then, I_{n-1} is reachable from I_n in one step for any control. As will appear in the proof of Theorem 2, if $(1,\ldots,N) = \hat{S} = I_0+\cdots+I_N$, then (7) is preferable to (6) (under the other conditions of the theorem). This condition says that given any control sequence, some state in I_0 will eventually be reached. This is clearly a necessary condition for (7) to be strictly preferable to (6), as can be seen from the way the successive substitutions in (7) are carried out. In any case (36) holds.

Theorem 2. *Suppose* U *is compact and the* $p_{ij}(\cdot)$ *and* $K(i,\cdot)$ *are continuous. Let* $K(0,\cdot) \equiv 0$ *and* $p_{00}(\cdot) \equiv 1$. *For any* $u \in U$, *let*

$$\sum_{j=1}^{N} p_{1j}(u) < 1 \tag{26'}$$

$$p_{ij}(u) > 0 \quad \textit{for some} \quad 1 \leq j(i,u) \leq i - 1 (\text{for } i > 1)$$

$$p_{ij}(u) > 0 \quad \textit{for some} \quad j(i,u) > 1.$$

Then both (6) *and* (7) *converge to the unique finite solution of* (5). *Let* $C_0(i) = \tilde{C}_0(i) = 0$ *and* $K(i,u) \geq 0$, $i = 1,\ldots,N$. *Then*

$$0 \leq C_n(i) \leq \tilde{C}_n(i) \uparrow C(i). \tag{36}$$

In addition, for i *in* $I_0+\cdots+I_n$,

$$\tilde{C}_{m+n}(i) > C_{m+n}(i), \quad m \geq 1. \tag{36'}$$

i.e., (7) *is strictly preferable to* (6) *for* i *in* $I_0+\cdots+I_N$.

Proof. The existence of a unique finite solution $\{C(i)\}$ to (5) is essentially proved in [2], [3]. For the moment, suppose that there is a solution. Denote the minimizing control by $\pi = (u_1,\ldots,u_N)$.

Then it is easy to verify that

$$
\begin{aligned}
C(i) &= \sum_{j=1}^{N} p_{ij}(u_i)C(j) + K(i,u_i) \\
C(i) &= \sum_{j=1}^{N} q_{ij}(\pi)C(j) + \tilde{K}(i,\pi)
\end{aligned}
\tag{37}
$$

where the $\{q_{ij}(\pi)\}$ and $\{\tilde{K}(i,\pi)\}$ are computed from the $\{p_{ij}(u_i)\}$ and $\{K(i,u_i)\}$ as in (25). Also (37) is not decreased, for any $i \geq 1$, if $\hat{\pi} \neq \pi$ replaces π. Thus, by the minimizing properties of π, π_n and $\tilde{\pi}_n$

$$
\begin{aligned}
C &= P(\pi)C + K(\pi) \leq P(\pi_n)C + K(\pi_n) \\
C_{n+1} &= P(\pi_n)C_n + K(\pi_n) \leq P(\pi)C_n + K(\pi)
\end{aligned}
\tag{38}
$$

$$
\begin{aligned}
C &= Q(\pi)C + \tilde{K}(\pi) \leq Q(\tilde{\pi}_n)C + \tilde{K}(\tilde{\pi}_n) \\
\tilde{C}_{n+1} &= Q(\tilde{\pi}_n)\tilde{C}_n + \tilde{K}(\tilde{\pi}_n) \leq Q(\pi)\tilde{C}_n + \tilde{K}(\pi).
\end{aligned}
\tag{39}
$$

(38) and (39) yield

$$
\begin{aligned}
C_{n+1} - C &\leq P^{n+1}(\pi)(C_0 - C) \\
C - C_{n+1} &\leq P(\pi_n)\cdots P(\pi_0)(C - C_0)
\end{aligned}
\tag{40}
$$

$$
\begin{aligned}
\tilde{C}_{n+1} - C &\leq Q^{n+1}(\pi)(\tilde{C}_0 - C) \\
C - \tilde{C}_{n+1} &\leq Q(\tilde{\pi}_n)\cdots Q(\tilde{\pi}_0)(C - \tilde{C}_0)
\end{aligned}
\tag{41}
$$

By Theorem 1, the sums of the rows in $P^n(\pi)$ and $Q^n(\pi)$ (which are $\sum_{j=1}^{N} p_{ij}^{n+1}(\pi)$ and $\sum_{j=1}^{N} q_{ij}^{n+1}(\pi)$, resp.) tend to zero as $n \to \infty$ monotonically, and faster for $Q(\pi)$ than for $P(\pi)$. Next it will be shown that the row sums of $P(\pi_n)\cdots P(\pi_0)$ and $Q(\tilde{\pi}_n)\cdots Q(\tilde{\pi}_0)$ also tend to zero as $n \to \infty$ for any $\pi_n \cdot\cdot \pi_0$ or $\tilde{\pi}_n \cdot\cdot \tilde{\pi}_0$ sequence. This implies, via (40) and (41), that $C_n \to C$ and $\tilde{C}_n \to C$, resp. Note that the row sums are $\sum_{j=1}^{N} p_{ij}^{n+1}(\pi_n \cdots \pi_0)$ and $\sum_{j=1}^{N} q_{ij}^{n+1}(\tilde{\pi}_n \cdots \tilde{\pi}_0)$, resp.; i.e., the non-absorbtion probabilities - when the process starts in state i - and uses π_n (resp. $\tilde{\pi}_n$) first and π_0 (resp. $\tilde{\pi}_0$) at time n.

First, a result for controls which <u>depend on the past</u> will be obtained. Let λ_s denote the collection $\{u_j(i_0 \cdots i_{s-1}\}$, where $u_j(i_0 \cdots i_{s-1})$ is the value of the control (lies in U) which is used at time s if $X_s = j$ and $X_0 = i_0, \ldots, X_{s-1} = i_{s-1}$. λ_s is a control policy at time S <u>depending on the past</u>. Let $\lambda_0 \lambda_1 \lambda_2 \cdots \lambda_s$ denote the control policy: λ_0 is used at time 0, λ_1 at time $1, \ldots,$ and λ_s at time s; λ_0 is just a π type of policy. Let $p_{ij}^r(\lambda_s \cdots \lambda_{s+r-1})$ denote $P\{X_{r+s} = j | X_s = i$ and λ_k used at time k, $k = s, \ldots, s+r-1\}$. The hypothesis (26') implies that there is some $\alpha > 0$ so that

$$1 - \sum_{j=1}^{N} p_{ij}^{N}(\lambda_s \cdots \lambda_{s+N-1}) \equiv p_{i0}^{N}(\lambda_s \cdots \lambda_{s+N-1}) \geq \alpha$$

for any sequence $\lambda_s \cdots \lambda_{s+n-1}$. Furthermore, since state 0 is absorbing $p_{i0}^N(\lambda_0 \cdots \lambda_{N-1})$ <u>is nondecreasing</u> in N for any fixed

$\lambda_0 \dots \lambda_n \dots$ sequence. It will be proved that

$$p_{i0}^{nN}(\lambda_0 \dots \lambda_{nN-1}) \geq 1 - (1-\alpha)^n. \tag{42}$$

(42) holds for $n = 1$. Suppose that its true for $n = r$. Then, using the induction hypothesis and $p_{00}^{s}(\cdot) \equiv 1$,

$$\begin{aligned}
p_{i0}^{rN+N}(\lambda_0 \dots \lambda_{rN+N-1}) &= p_{i0}^{rN}(\lambda_0 \dots \lambda_{rN-1}) p_{00}^{N} \\
&+ \sum_{j=1}^{N} p_{ij}^{rN}(\lambda_0 \dots \lambda_{rN-1}) p_{j0}^{N}(\lambda_{rN} \dots \lambda_{rN+N-1}) \\
&\geq 1 - (1-\alpha)^r + \alpha \sum_{j=1}^{N} p_{ij}^{rN}(\lambda_0 \dots \lambda_{rN-1}) \geq 1 - (1-\alpha)^r + \alpha(1-\alpha)^r \\
&= 1 - (1-\alpha)^{r+1}
\end{aligned}$$

and (42) is proved. This implies that the row sums $\sum_{j=1}^{N} p_{ij}^{n}(\pi_n \dots \pi_0)$ tend to zero as $n \to \infty$. Next, in order to obtain a similar relation for $\sum_{j=1}^{N} q_{ij}^{n}(\tilde{\pi}_n \dots \tilde{\pi}_0)$, the 'history dependent' control analogy to (32-35) must be given. Using (12) and (32-34),

$$F_1^{r+1}(\tilde{\pi}_r \dots \tilde{\pi}_0) \equiv \sum_{j=1}^{N} q_{ij}^{r+1}(\tilde{\pi}_r \dots \tilde{\pi}_0) = \sum_{k=1}^{N} q_{ik}(\tilde{\pi}_r) \sum_{j=1}^{N} q_{kj}^{r}(\tilde{\pi}_{r-1} \dots \tilde{\pi}_0) \tag{43}$$

$$= \sum_{k=i}^{N} p_{ik}(\tilde{\pi}_r) \sum_{j=1}^{N} q_{kj}^{r}(\tilde{\pi}_{r-1} \dots \tilde{\pi}_0) + \sum_{k=1}^{N} \Big(\sum_{s=1}^{i-1} p_{is}(\tilde{\pi}_r) q_{sk}(\tilde{\pi}_r) \Big) \sum_{j=1}^{N} q_{kj}^{r}(\tilde{\pi}_{r-1} \dots \tilde{\pi}_0)$$

$$= \sum_{k=i}^{N} p_{ik}(\tilde{\pi}_r) \sum_{j=1}^{N} q_{kj}^{r}(\tilde{\pi}_{r-1} \dots \tilde{\pi}_0) + \sum_{k=1}^{i-1} p_{ik}(\tilde{\pi}_r) \sum_{j=1}^{N} q_{kj}^{r+1}(\tilde{\pi}_r \dots \tilde{\pi}_0)$$

It will be shown that for any $\tilde{\pi}_r \ldots \tilde{\pi}_0$, there is some corresponding $\lambda_0 \ldots \lambda_r$, so that

$$\sum_{j=1}^{N} q_{ij}^{r+1}(\tilde{\pi}_r \ldots \tilde{\pi}_0) \leq \sum_{j=1}^{N} p_{ij}^{r+1}(\lambda_0 \ldots \lambda_r). \tag{44}$$

(47) and (45) imply that the left side of (47) goes to zero as $n \to \infty$, and hece, that $\tilde{C}_n \to C$. (47) holds for $r = 0$ since for any π' (Theorem 1)

$$\sum_{j=1}^{N} q_{ij}(\pi') \leq \sum_{j=1}^{N} p_{ij}(\pi'). \tag{45}$$

For $r = 1$, using (43), (45) and

$$\sum_{j=1}^{N} q_{kj}^{r+1}(\tilde{\pi}_r \ldots \tilde{\pi}_0) \leq \sum_{j=1}^{N} q_{kj}^{r}(\tilde{\pi}_r \ldots \tilde{\pi}_1), \tag{46}$$

we obtain

$$\begin{aligned} F_i^2(\tilde{\pi}_1\tilde{\pi}_0) &\leq \sum_{k=i}^{N} p_{ik}(\tilde{\pi}_1) \sum_{j=1}^{N} q_{kj}(\tilde{\pi}_0) + \sum_{k=1}^{i-1} p_{ik}(\tilde{\pi}_1) \sum_{j=1}^{N} q_{kj}(\tilde{\pi}_1) \\ &\leq \sum_{k=i}^{N} p_{ik}(\tilde{\pi}_1) \sum_{j=1}^{N} p_{kj}(\tilde{\pi}_0) + \sum_{k=1}^{i-1} p_{ik}(\tilde{\pi}_1) \sum_{j=1}^{N} p_{kj}(\tilde{\pi}_1) \qquad (47) \\ &\leq \sum_{k=1}^{N} p_{ik}^2(\tilde{\pi}_1\lambda_1) \end{aligned}$$

where $\tilde{\pi}_1\lambda_1$ is the control sequence which uses $\tilde{\pi}_1$ at time 0 and

λ_1 at time 1; where $\lambda_1 = \tilde{\pi}_0$ on the set of states $\{k:k \geq i\}$, and $\lambda_1 = \tilde{\pi}_1$ on the set of states $\{k:k < i\}$. Thus (44) holds for $r = 2$. In general suppose (44) holds for r. Then, repeating the steps (47), and using the induction hypothesis and (46),

$$F_i^{r+1}(\tilde{\pi}_r \dots \tilde{\pi}_0) = \sum_{k=i}^{N} p_{ik}(\tilde{\pi}_r) \sum_{j=1}^{N} q_{kj}^{r}(\tilde{\pi}_{r-1} \dots \tilde{\pi}_0)$$

$$+ \sum_{k=1}^{i-1} p_{ik}(\tilde{\pi}_r) \sum_{j=1}^{N} q_{kj}^{r+1}(\tilde{\pi}_r \dots \tilde{\pi}_0)$$

$$\leq \sum_{k=i}^{N} p_{ik}(\tilde{\pi}_r) \sum_{j=1}^{N} p_{kj}^{r}(\lambda_0' \dots \lambda_{r-1}') + \sum_{k=1}^{r-1} p_{ik}(\tilde{\pi}_r) \sum_{j=1}^{N} p_{kj}^{r}(\lambda_0'' \dots \lambda_{r-1}'')$$

for <u>some</u> $\lambda_0' \dots \lambda_{r-1}'$ and $\lambda_0'' \dots \lambda_{r-1}''$. Finally

$$F_i^{r+1}(\tilde{\pi}_r \dots \tilde{\pi}_0) \leq \sum_{j=1}^{N} p_{ij}^{r+1}(\tilde{\pi}_r \lambda_0 \dots \lambda_{r-1}) \tag{48}$$

where $\lambda_0 \dots \lambda_{r-1} = \lambda_0' \dots \lambda_{r-1}'$ if $X_1 \geq i$ and $\lambda_0 \dots \lambda_{r-1} = \lambda_0'' \dots \lambda_{r-1}''$ if $X_1 < i$. Thus (44) holds for all r.

We now prove the last assertion of the Theorem. Let $\alpha_i \leq C(i)$. Then, since it is now supposed that $K(\cdot, i) \geq 0$, we have

$$\beta_i \equiv \min_u \left[\sum_{j=1}^{N} p_{ij}(u)\alpha_j + K(i,u) \right] \leq C(i). \tag{49}$$

Furthermore the β_i are nondecreasing as the $\{\alpha_j\}$ increase. <u>Let</u> $\tilde{C}_0(i) = C_0(i) \equiv 0$ <u>from here on</u>. It is easily verified that $C_n = \{C_n(i)\}$ is the optimum cost for an $n + 1$ stage control process and is nondecreasing (it tends to C) as n increases. It is next verified, using (49), that $\tilde{C}_n(i) \leq C(i)$, for all n: Note that $C_1(i) = \inf_u K(i,u)$. From the last line of (25), for any $\pi' = (u_1', \ldots, u_N')$

$$\tilde{K}(i,\pi') = \sum_{j=1}^{i-1} p_{ij}(u_i')\tilde{K}(j,\pi') + K(i,u_i') \geq K(i,u_i'). \tag{50}$$

(50), the remark below (49), and $\tilde{C}_0 = C_0 = 0$, imply that $\tilde{C}_n \geq C_n$ for all n. Note that (50) is a <u>strict inequality</u> for i in I_0, and <u>any</u> π'. These remarks together with (50) imply that $\tilde{C}_n(i) > C_n(i)$ for i in I_0 and all $n \geq 1$. Next suppose that $i \in I_1$; i.e., for <u>any</u> u in U, there is <u>some</u> j in I_0 so that $p_{ij}(u) > 0$. Then, since $\tilde{C}_n(j) > C_n(j)$ for $j \in I_0$ and $\tilde{C}_n(j) \geq C_n(j)$ otherwise (for $n \geq 1$), we have $\tilde{C}_{n+1}(i) > C_{n+1}(i)$ for i in $I_1 + I_0$ and $n \geq 1$. Repeating this procedure gives $\tilde{C}_{n+r}(i) > C_{n+r}(i)$ for i in $I_0 + \cdots + I_r$, and all $n \geq 1$.

We must only prove that (5) has a unique finite solution. By what has been said, it is clear that, under the hypotheses and for any <u>fixed</u> $\pi = (u_1, \ldots, u_N)$ with $u_i \in U$, that

$$C^{\pi}(i) = \sum_{j=1}^{N} p_{ij}(u_i)C^{\pi}(j) + K(i,u_i)$$

has a unique solution. Also, for any π' and π'', there is a π such that $C^{\pi}(i) \leq \min\,(C^{\pi'}(i), C^{\pi''}(i))$, $i = 1, \ldots, N$. (See [3], p. 26 for a proof). Thus there is a sequence $\pi_n = (u_1^n, \ldots, u_N^n)$

so that

$$C^{\pi_n}(i) \to \operatorname*{glb}_{\pi} C^{\pi}(i) \equiv C(i), \quad i = 1,\ldots,N.$$

Also

$$\lim_n C^{\pi_n}(i) = \lim_n \left[\sum_{j=1}^{N} p_{ij}(u_i^n) C^{\pi_n}(j) + K(i,u_i^n)\right]$$

$$C(i) = \lim_n \left[\sum_{j=1}^{N} p_{ij}(u_i^n) C(j) + K(i,u_i^n)\right].$$

The continuity of $p_{ij}(\cdot)$ and $K(i,\cdot)$ and compactness of U imply that there is some $\pi = (u_1,\ldots,u_N)$ so that

$$C(i) = \sum_{j=1}^{N} p_{ij}(u_i) C(j) + K(i,u_i).$$

The definition $C(i) \equiv \operatorname*{glb}_{\pi} C^{\pi}(i)$ implies that $C(i)$ satisfies (5). Q.E.D.

Acknowledgement. The author would like to express his appreciation to Professor M. Alan Feldstein for his critical reading of parts of the original manuscript and his suggestions for the improvement of its readability.

REFERENCES

[1] R. S. Varga, Matrix Iterative Analysis, Prentic Hall, 1962.

[2] R. A. Howard, Dynamic Programming and Markov Processes, The Technology Press of M. I. T., 1960.

[3] J. H. Eaton, L. A. Zadeh, "Optimal Pursuit Strategies in Discrete State Probabilistic Systems", J. of Basic Eng., 84, March, 1962, pp. 23-29.

[4] W. Feller, Probability Theory and its Applications, Vol. I, Wiley, 1950.

[5] J. L. Doob, Stochastic Processes, Wiley, 1953.

[6] H. J. Kushner, Stochastic Stability and Control, Academic Press, 1967.

[7] W. H. Fleming, "Some Markovian Optimization Problems", J. Math. and Mech., 12, p. 131-140, 1963.