ERGODIC THEOREMS FOR A CERTAIN CLASS OF MARKOFF PROCESSES

Thesis by

Maurice Kennedy

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1954

Acknowledgements

It is with great pleasure that I acknowledge my debt to Professor Samuel Karlin. He suggested this problem to me, and has been most generous with help and advice throughout the past two years.

I have also great pleasure in expressing my thanks to the United States Government, for having awarded me a grant for graduate study during the year 1951-52.

My thanks are also due to the California Institute of Technology for having provided me with the opportunity of continuing my work here, by awarding me a graduate research assistantship and a tuition scholarship; and in particular I wish to express my thanks to Professor Bohnenblust and the Mathematics Department, on whose recommendation such awards are made.

Abstract

A system, whose state may be described by a point t in a bounded set in Euclidean space, is considered. At every unit interval of time, attractions A_i towards certain points y_i are applied with probabilities $\phi_i(t)$, ($\sum \phi_i(t) = 1$), where t is the state of the system. Given the initial probability distribution $\mu(t)$ for the state of the system, the problem is to obtain limiting theorems for the distribution at the n^{th} unit of time as $n \to \infty$. Subject to certain conditions on $\{A_i\}$ and $\{\phi_i(t)\}$ such convergence theorems are obtained. Some particular properties for the case, where the attractions are toward the vertices of a simplex, are discussed. Finally the one-dimensional learning model is considered.

TABLE OF CONTENTS

1. Introduction.

Certain learning models have been introduced by Bush and Mosteller [1]. These models give rise to a class of transition operators which were studied by Karlin [2] for the case of one dimension. The mathematical description of the simplest process of this kind can be formulated as follows:

A particle on the unit interval executes a random walk subject to two impulses, one of which is applied at the end of every unit of time. If it is located at the point t, then $t \to A_0 t = \lambda_0 t$ with probability $\phi_0(t)$, and $t \to A_1 t = \lambda_1 t + 1 - \lambda_1$ with probability $\phi_1(t)$, where $\phi_0(t) + \phi_1(t) = 1$. If $\mu(t)$ is the initial probability measure for the position of the particle, the probability measure $T\mu$ at the end of the first time interval, is clearly given by

$$T\mu(E) = \int_{A_0^{-1}E} \phi_0(t)\, d\mu(t) + \int_{A_1^{-1}E} \phi_1(t)\, d\mu(t).$$

If $\phi_i(t)$ $(i = 0,1)$ are continuous, then T represents a continuous transformation of the Banach space of Borel measures on the unit interval, into itself. Instead of working with the transformation T on this space, Karlin considers the transformation U, on the space of continuous functions, defined by

$$Ux(t) = \sum_i \phi_i(t)\, x(A_i t) .$$

He shows that T is the conjugate transformation to U. Thus by proving strong convergence for the iterates U^n, subject to certain conditions on $\phi_i(t)$, weak-star convergence is obtained for $T^n\mu$. For probability measures, weak-star convergence is of course convergence in distribution.

It is the object of this thesis to generalize these results. A countable set of transformations $\{A_i\}$ of a compact metric space Ω into itself is introduced in section 3. These have the property that $\rho(A_i t, A_i s) \leq \lambda \rho(t,s)$ $(0 < \lambda < 1)$. Each A_i represents an attraction towards a fixed point y_i $(i = 1,2, \ldots)$.

In section 4 we consider a system whose state may be described by a point t in Ω. Our Markoff process consists in applying at every unit interval of time, one of the transformations $\{A_i\}$, A_i being chosen with probability $\phi_i(t)$, where $\{\phi_i(t)\}$ is a family of continuous functions each satisfying a uniform Lipschitz condition, and the sum of whose Lipschitz constants is finite. If μ is a probability measure defined on the Borel sets of Ω, giving the probability distribution of the initial state of the system, then the probability distribution $T\mu$, for the state of the system at the end of the first unit time interval, is

$$T\mu(E) = \sum \int_{A_i^{-1}E} \phi_i(t)\, d\mu(t).$$

This is a continuous transformation of the space $\mathfrak{M}(\Omega)$ of finite Borel measures into itself. The object is to obtain weak-star convergence theorems for $T^n\mu$. Proceeding as in Karlin's paper we define the transformation U of $C(\Omega)$ (the space of continuous functions on Ω) into itself by

$$Ux(t) = \sum_1^{\infty} \phi_i(t)\, x(A_i t).$$

U is a continuous transformation and T is its conjugate, i.e.,

$$(Ux, \mu) = (x, T\mu) \qquad x \in C(\Omega), \quad \mu \in \mathfrak{M}(\Omega)$$

where

$$(x,\mu) = \int x(t)\, d\mu(t).$$

It follows, since $(U^n x, \mu) = (x, T^n \mu)$ that $(U^n x, \mu)$ represents the expectation of $x(t)$ at the end of the n^{th} time interval.

There is a direct probability interpretation of the transformation U. It is clear from the definition of U that $U\phi_i(t)$ represents the probability that if the state of the systems is initially at t, then at the first unit of time the attraction A_i is applied; and in general, $U^n\phi_i(t)$ represents the probability that at the n^{th} unit of time, the attraction A_i is applied, given that initially the state of the system is t.

At this stage reference should be made to the work of other authors - Ocinescu, Mihoc, Doeblin, Fortet, Ionescu Tulcea and Marinescu [3-9]. We shall refer in the main to the papers [8,9] of the latter two, for they generalize that part of the work of the others which is relevant here. Our transformation U is a particular case of a transformation considered by them, and it follows from their work that if we consider U as a transformation of the space $C_L(\Omega)$ into itself, where $C_L(\Omega)$ is the Banach space of continuous functions $x(t)$, each satisfying a uniform Lipschitz condition on Ω, with the norm

$$||x||_L = \max|x(t)| + \sup_{t \neq s} \left| \frac{x(t) - x(s)}{\rho(t,s)} \right|$$

then U is a quasi-completely continuous transformation. This important result enables us in section 5, to establish uniform $C - 1$ convergence on $C_L(\Omega)$ for the iterates U^n, and in particular to prove $C - 1$ convergence for $U^n\phi_i$.

In order to obtain results for $C(\Omega)$ and $\mathfrak{M}(\Omega)$ we must now specialize Ω to be a compact set in finite-dimensional Euclidean space. We do this in order to have the property that $C_L(\Omega)$ be dense in $C(\Omega)$. In section 6 we obtain by the Banach-Steinhaus theorem $C - 1$ strong convergence for U^n on $C(\Omega)$. In section 7 these results are translated into the conjugate space and $C - 1$ weak-star convergence is obtained for $T^n\mu$.

In section 8, we consider the case where for each i, $(i = 1,2, \ldots)$ the fixed point of A_i is an absorbing point, i.e., y_i is an absorbing point $(i = 1,2, \ldots)$. This condition is expressed by $\phi_i(y_i) = 1$ $(i = 1,2, \ldots)$. In addition we introduce the additional and natural assumption that for each i, as y_i is approached, the probability of choosing the next attraction to be A_i, does not decrease. This is expressed mathematically by assuming $\phi_i(A_it) \geq \phi_i(t)$ for all t $(i = 1,2, \ldots)$. Under these assumptions, we obtain the uniform convergence of U^n on $C_L(\Omega)$, their strong convergence on $C(\Omega)$, and the weak-star convergence of $T^n\mu$ in $\mathfrak{M}(\Omega)$. Moreover the limiting transformations are explicitly determined.

In section 9, we consider certain properties of U and T for a certain class of probabilities $\{\phi_i(t)\}$. In section 10 a slightly different class (though strongly overlapping) is considered. Here the assumption that $\sum_i \phi_i(t)\,\phi_i(s) \neq 0$ for all t,s in Ω is made, and again convergence theorems are obtained. The method used is exactly that of Karlin [2] and indeed the result is indicated in his paper. Here however it is difficult to obtain information about the limiting transformation.

Section 11 introduces the N-dimensional analogue of the model considered by Karlin. It consists of attractions towards the vertices of a simplex, each attraction being directly proportional to the distance from the corresponding vertex. This model is a particular example of the problem treated in earlier sections. Certain properties concerning the convergence of derivatives of $U^n x(t)$ are obtained under further restrictions of $\phi_i(t)$.

The thesis concludes with section 12 which is devoted to establishing certain additional properties for the one-dimensional learning model.

At the start in section 2, those properties of Banach spaces which will be required, are set out. The main theorem of Ionescu Tulcea and Marinescu (theorem 2.2) which provides the quasi-complete continuity of U on $C_L(\Omega)$ (theorem 4.2) is stated. However it is not always necessary to appeal to this theorem to establish this result, apart from the verification that the iterates U^n are uniformly bounded in $C_L(\Omega)$, which is easily proved, and is stated as a separate lemma (lemma 4.8, which is itself derived from lemma 2.1). For example in section 9 once the convergence in $C(\Omega)$ is obtained, it is easily deduced from theorem 2.1 (a simpler theorem than that of Ionescu Tulcea and Marinescu) that U is quasi-completely continuous on $C_L(\Omega)$. For the case of a finite number of attractions this is also true under the hypothesis of section 8.

It is clear that similar theorems may be worked out where instead of $\{\phi_i(t)\}$ satisfying a Lipschitz condition of order 1, they satisfy a Lipschitz condition of order d, where $0 < d \leq 1$ [9, page 146].

Finally, we note that only the case, where the number of attractions is countable, has been treated. However, it is clear from the work of Ionescu Tulcea and Marinescu [8,9] that subject to certain assumptions corresponding to those imposed on $\{\phi_i(t)\}$ here, the transformation U obtained for the non-countable case is quasi-completely continuous on $C_L(\Omega)$. There seems no reason to doubt that convergence theorems similar to those of sections 8 and 10 can be obtained.

2. Theorems on Banach spaces.

We first consider those properties of Banach spaces which shall be required in subsequent sections.

Let E, B be Banach spaces with norms $||x||_E$, $||x||_B$, respectively and with the property that as vector spaces $B \subset E$.

Let U be a linear transformation of B into itself, which is continuous with respect to both norms, i.e.,

$$(2.1) \qquad ||U||_B = \sup_{\substack{x \in B \\ ||x||_B \le 1}} ||Ux||_B < \infty$$

and

$$(2.2) \qquad ||U||_E = \sup_{\substack{x \in B \\ ||x||_E \le 1}} ||Ux||_E < \infty.$$

In addition we assume

(2.3) there exist two constants R,r $(0 < r < 1)$ such that

$$||Ux||_B \le r||x||_B + R||x||_E, \qquad x \in B.$$

The following lemma is due to Ionescu Tulcea and Marinescu [9]. The proof is simple and is given for completeness.

LEMMA 2.1. If $||U^n||_E \le H$, $(n = 1,2, \dots)$ then

$$(2.4) \qquad ||U^n||_B \le C \qquad n = 1,2, \dots$$

where C is some positive constant.

PROOF: For $x \in B$ we have from (2.3)

$$||U^2x||_B \leq r||Ux||_B + R||Ux||_E$$

$$\leq r^2||x||_B + Rr||x||_E + RH||x||_E$$

since $||Ux||_E \leq H||x||_E$ by hypothesis.

Again

$$||U^3x||_B \leq r^3||x||_B + (r^2R + rRH + RH)\ ||x||_E$$

$$\leq r^3||x||_B + A(r^2 + r + 1)\ ||x||_E$$

where $A = \text{Max}(R, R\cdot H)$.

(2.5) $$||U^nx||_B \leq r^n||x||_B + L||x||_E$$

where $L = \frac{A}{1 - r}$.

Hence for each $x \in B$, $||U^nx||_B < \infty$. Therefore by the principle of uniform boundedness,

$$||U^n||_B \leq C$$

where C is some positive constant, and the lemma is established.

We shall denote by S_r, the sphere of radius r in B, i.e.,

(2.6) $$S_r = \left\{ x \in B \mid\ ||x||_B \leq r \right\}.$$

<u>THEOREM 2.1</u>. If

(2.7) S_1 is compact in E (i.e. with respect to $||x||_E$).

(2.8) $||U^n||_E \leq H, \qquad ||U^n||_B \leq K, \qquad n = 1,2, \ldots$

(2.9) For each $x \in B$, there exists an element $Vx \in E$ such that

$$||U^nx - Vx||_E \to 0 \qquad \text{as} \qquad n \to \infty$$

then V is a linear <u>transformation of B into itself</u> continuous with respect to both norms; in fact

(2.10) $\quad ||V||_E \leq H, \qquad ||V||_B \leq K$

and the sequence of transformations U^n converge uniformly in the B-norm to V, i.e.,

(2.11) $\quad ||U^n - V||_B \to 0 \quad \text{as} \quad n \to \infty.$

<u>PROOF</u>: The first step is to establish that for each $x \varepsilon B$, $Vx \varepsilon B$. V is clearly a linear transformation of B into E. For $x \varepsilon S_1$,

$$||U^n x||_B \leq K||x||_B \qquad \text{by (2.8)}$$

$$\leq K \qquad \text{(since } x \varepsilon S_1\text{)}.$$

Therefore for $x \varepsilon S_1$, $\{U^n x\} \varepsilon S_K$ which by (2.7) is clearly a compact set in E and thus is closed in E. From (2.9) it follows that $Vx \varepsilon S_K$, and hence is in B. Since V is linear it is clear that for any $x \varepsilon B$, $Vx \varepsilon B$. Moreover,

$$||V||_B = \sup_{x \varepsilon S_1} ||Vx||_B$$

$$\leq \sup_{x \varepsilon S_K} ||x||_B$$

$$= K$$

$\therefore$ $||V||_B \leq K$ and we have established that V is a continuous linear transformation of B into itself.

From (2.8),

$$||U^n x||_E \leq H||x||_E \qquad x \varepsilon B.$$

$$\therefore \quad ||Vx||_E \leq H||x||_E \qquad \text{(by 2.9).}$$

Hence V is continuous in E-norm and $||V||_E \leq H$. It is also clear from (2.9), that Vx is a fixed point of U.

It remains to prove (2.11). From (2.8) we see that $\{U^n x\}$ form an equicontinuous family of functions defined on B, with respect to either norms. By (2.9), $U^n x$ converges in E-norm and therefore the convergence is uniform on every compact set in E; in particular

$$||U^n x - Vx||_E \to 0$$

uniformly with respect to $x \varepsilon S_1$ a compact set (by 2.7). The hypothesis of lemma 2.1 being satisfied, we may use equation (2.5) as follows

$$\begin{aligned} ||U^n x - Vx||_B &= ||U^{n-k}(U^k x - Vx)||_B \\ &\leq r^{n-k} \, ||U^k x - Vx||_B + L||U^k x - Vx||_E \, . \end{aligned}$$

Choose $k \ni ||U^k x - Vx||_E \leq \frac{\varepsilon}{2L}$ for all $x \varepsilon S_1$. Now choose $N \ni n \geq N, |r^{n-k}| \leq \frac{\varepsilon}{4K}$. For $x \varepsilon S_1$

$$\begin{aligned} ||U^k x - Vx||_B &\leq ||U^k x||_B + ||Vx||_B \\ &\leq ||U^k||_B + ||V||_B \\ &\leq 2K. \end{aligned}$$

Hence for $x \varepsilon S_1$, $n \geq N$

$$||U^n x - Vx||_B \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

i.e., $||U^n x - Vx||_B \to 0$ uniformly with respect to $x \varepsilon S_1$, or

$$||U^n - V||_B \to 0$$

which completes the proof of the theorem.

<u>THEOREM 2.2</u>. Theorem of Ionescu Tulcea and Marinescu.

If

(2.12) $\quad x_n \varepsilon B, \quad ||x_n||_B \leq K, \quad \lim||x_n - x||_E = 0$

implies that

$$x \varepsilon B \quad \text{and} \quad ||x||_B \leq K.$$

(2.13) $\quad ||U^n||_E \leq H, \qquad n = 1,2, \ldots$

(2.14) $\quad U$ transforms every bounded set in B into a compact set in E,

then the norms $||U^n||_B$ are uniformly bounded and U is quasi-completely continuous as a transformation of B into itself, i.e., $\exists \; n \; \ni$

$$||U^n - V||_B < 1$$

where V is a completely continuous transformation of B into itself.

Corollary: The theorem is valid if in (2.3), (2.14), U is replaced by U^m for some positive integer m.

The theorem and corollary are proved in reference [9]. The authors also observe that if the unit sphere S_1 in B is compact in E, that (2.12) and (2.14) are satisfied. We prove this in the following lemma.

<u>LEMMA 2.2.</u> If the unit sphere S_1 in B is compact with respect to $||x||_E$ then the hypotheses (2.12) and (2.14) of the above theorem are automatically satisfied.

<u>PROOF</u>: Let $x_n \varepsilon B, \quad ||x_n||_B \leq K, \quad \lim_{n \to \infty} ||x_n - x||_E = 0.$

$x_n \varepsilon S_K = \{x \varepsilon B | \ ||x||_B \leq K\}$ which is clearly a compact set in E from the hypothesis. Hence there exists a subsequence x_{n_i} converging in E-norm to an element in S_K. This element must be x. Hence $x \varepsilon S_K$, i.e., $x \varepsilon B$ and $||x||_B \leq K$. Thus hypothesis (2.12) is established.

Now let P be a bounded set in B, i.e., $x \varepsilon P$ implies that $||x||_B \leq K$ (a constant).

Let $\{x_v\} \varepsilon P$

$$||Ux_v||_B \leq ||U||_B \ ||x_v||_B$$

$$\leq K||U||_B = R \text{ (say).}$$

Now $S_R = \{x \varepsilon B | \ ||x||_B \leq R\}$ is compact in E. Hence there exists a subsequence $\{Ux_{v_i}\}$ which converges in E-norm to an element $y \varepsilon S_R$, i.e., to an element in B. Hence UP is a conditionally compact set with respect to the E-norm, but the closure of a conditionally compact set in a complete metric space is compact. Hence UP is contained in a compact set in E. Hence the hypothesis (2.14) is established and the lemma is proved.

We conclude this section by stating the general ergodic theorem for Banach spaces as proved by Yosida and Kakutani [10] and the

<u>THEOREM 2.3</u>. Let B be a Banach space and U a completely continuous or quasi-completely continuous transformation of B into itself such that $||U^n|| \leq C$. Then there exists at most a finite number of eigenvalues of U of modulus 1. Denote these values by $\lambda_1, \lambda_2, \ldots, \lambda_k$. Then there exists a system of completely continuous transformations

$U_{\lambda_1}, \dots, U_{\lambda_k}$ (each of them $\neq 0$) and a completely continuous or quasi-completely continuous transformation S (which may vanish) such that

(2.15) $$U^n = \sum_{i=1}^{k} \lambda_i^n U_{\lambda_i} + S^n, \qquad n = 1,2, \dots$$

with

(2.16) $$\begin{cases} UU_{\lambda_i} = U_{\lambda_i} U = \lambda_i U_{\lambda_i}, \quad U^2_{\lambda_i} = U_{\lambda_i}, \quad U_{\lambda_i} U_{\lambda_j} = 0 \ (i \neq j) \\ U_{\lambda_i} S = S U_{\lambda_i} = 0, \quad \|U_{\lambda_i}\| \leq C, \quad i = 1, \dots, k \end{cases}$$

and

(2.17) $$\|S^n\| \leq \frac{M}{(1 + \varepsilon)^n}, \qquad M, \varepsilon \text{ being positive constants}$$

(2.18) For any complex λ, $|\lambda| = 1$, there exists a completely continuous transformation U_λ, which maps B into itself, such that

$$\left\| \frac{1}{n} \left(\frac{U}{\lambda} + \frac{U^2}{\lambda^2} + \dots + \frac{U^n}{\lambda^n} \right) - U_\lambda \right\| \leq \frac{M}{n} \qquad n = 1,2, \dots$$

where M is a constant independent of n, and $U_\lambda \neq 0$ if and only if λ is an eigen value of U

(2.19) $\|U^n\| \to 0$ if and only if U has no eigen-value of modulus 1. If $\|U^n\| \to 0$ then convergence is geometric, i.e.,

$$\|U^n\| \leq \frac{M}{(1 + h)^n}, \qquad M, h \text{ being positive constants.}$$

(2.20) $\|U^n - U_1\| \to 0$, $U_1 \neq 0$ if and only if 1 is an eigen-value of U and there are no other eigen-values of modulus 1. Again if this condition is satisfied, the convergence is geometric, i.e.,

(2.21) $$\|U^n - U_1\| \leq \frac{H}{(1 + h)^n}, \qquad H, h \text{ being constants.}$$

3. Description of a Certain Class of Transformations on a Compact Metric Space.

Let Ω be a compact metric space. Let $\{A_i\}$ be a countable set of 1-1 transformations of Ω into itself with the property that

$$(3.1) \qquad \rho(A_i t, A_i s) \leq \lambda \rho(t,s) \qquad i = 1,2, \ldots$$

for all pairs of points $t,s \in \Omega$ where ρ is the metric on Ω and λ is a constant such that $0 < \lambda < 1$.

LEMMA 3.1. $\{A_i\}$ form a set of 1-1 bicontinuous transformations of Ω into itself. Each A_i represents an attraction towards a fixed point y_i, which is the unique fixed point of A_i, and moreover the rate of attraction to the fixed points is uniform with respect to i, i.e.

$$(3.2) \qquad \lim_{n \to \infty} A_i^n t = y_i$$

for any $t \in \Omega$, the convergence being uniform with respect to i and $t \in \Omega$. For each i, y_i is the unique solution of

$$(3.3) \qquad A_i y_i = y_i.$$

PROOF: It is clear that if $t_j \to t$, we have by (3.1)

$$A_i t_j \to A_i t \qquad \text{each } i.$$

Hence A_i is a continuous transformation. From (3.1)

$$\rho(A_i^n t, A_i^n s) \leq \lambda^n \rho(t,s)$$
$$\leq M\lambda^n \qquad \text{where } M \text{ is a constant, since } \rho(t,s) \text{ is bounded for a compact set.}$$

Choose $s = A_i^p t$. Hence, by above inequality

$$\rho(A_i^n t,\ A_i^{n+p} t) \leq M\lambda^n$$

for every positive integer p. Hence, since a compact metric space is complete,

$$A_i^n t \to y_i.$$

Clearly $A_i y_i = y_i$. Moreover y_i is the unique point satisfying this equation, for let $A_i y_i' = y_i'$, then $A_i^n y_i' = y_i'$ so that since

$$\rho(A_i^n y_i,\ A_i^n y_i') \leq \lambda^n M$$

$$\rho(A_i^n y_i,\ y_i') \leq \lambda^n M.$$

Let $n \to \infty$, then

$$\rho(y_i, y_i') = 0.$$

$$\therefore\ y_i = y_i'.$$

Finally the convergence in (3.2) is uniform with respect to $t \in \Omega$ and all i, for

$$\rho(A_i^n t, y_i) = \rho(A_i^n t,\ A_i^n y_i) \leq \lambda^n M.$$

Hence the convergence is uniform.

The bicontinuity of A_i follows from the fact that a continuous 1-1 transformation of a compact space into a Hausdorff space is bicontinuous. Hence the lemma is proved. For convenience in notation we now introduce a definition.

Definition: $A^{[n]}S = \bigcup_{i_1,\ldots,i_n} A_{i_1}, \ldots, A_{i_n} S$ for any set S

$A^{[n]}\Omega$ represents the range of all products of n transformations.

Since $A_i \Omega \subset \Omega$, it is clear that

$$A^{[1]}\Omega \supset A^{[2]}\Omega \supset A^{[3]}\Omega \quad \ldots\ldots\ldots$$

This is a monotone decreasing sequence of sets, which therefore converges to F, where

$$(3.4) \qquad F = \bigcap_{1}^{\infty} A^{[n]}\Omega \; .$$

F represents the points of Ω which are reached by an infinite number of transformations. We define G to be the complementary set, i.e.,

$$(3.5) \qquad G = F^c .$$

<u>LEMMA 3.2.</u> If the number of transformations $\{A_i\}$ is N (finite) then F is a closed set.

<u>PROOF</u>: A_i being a continuous transformation maps any compact set onto a compact set. Hence $A^{[n]}\Omega$, being a finite union of compact sets and therefore closed sets, is closed. Thus F is a countable intersection of closed sets and therefore is closed.

An alternative way of writing F is provided by the following lemma. The proof is that of Hausdorff [11, page 131].

<u>LEMMA 3.3.</u>

$$(3.6) \qquad F = \bigcup_{\sigma} \bigcap_{n=1}^{\infty} A_{i_1}, \ldots , A_{i_n} \Omega,$$

where the union is taken over all infinite sequences σ ,

$$\sigma = (i_1, i_2, i_3 , \ldots)$$

where the i's range from 1 to N.

<u>PROOF</u>: Denote the right-hand side of F by F'. Any element of F' is given as the limit of a sequence of sets

$$(3.7)\qquad A_{i_1}\Omega \supset A_{i_1}A_{i_2}\Omega \supset A_{i_1}A_{i_2}A_{i_3}\Omega \quad \dots .$$

It is obvious that the element determined by this sequence is in F (see (3.4)).

On the other hand let $x \in F$ (3.4). Then

$$(3.8)\qquad x \in A_{\ell_1^1}\Omega, \quad A_{\ell_1^2}A_{\ell_2^2}\Omega, \quad A_{\ell_1^3}A_{\ell_2^3}A_{\ell_3^3}\Omega, \quad \dots .$$

Consider the first suffix in each of these terms. They form an infinite sequence $\{\ell_1^k\}$. Since the range of suffixes is 1 to N, there must be one integer which occurs infinitely often in the sequence. Let it be denoted by i_1. Remove those terms in (3.8) which do not have i_1 as the first suffix. Consider the second suffixes of the remainder. Apply the same argument to obtain suffixes $i_2, i_3, \dots$, such that

$$x \in A_{i_1}\Omega, \quad A_{i_1}A_{i_2}\Omega, \quad A_{i_1}A_{i_2}A_{i_3}\Omega, \dots .$$

Hence by (3.7) $x \in F'$. Hence $F = F'$ and the lemma is proved.

4. Description of the Markoff Process.

First we state a well-known theorem.

THEOREM 4.1. Let Ω be a compact Hausdorff space, and let $C_R(\Omega)$ be the real Banach space of real continuous functions $x(t)$ defined on Ω with

$$||x|| = \max_{t \in \Omega} |x(t)|$$

Let $C^*(\Omega)$ be the conjugate Banach space of $C_R(\Omega)$. Define partial orderings of both spaces as follows:

$$x \geq y, \text{ if and only if } x(t) \geq y(t) \text{ all } t \in \Omega \text{ ; } x,y \in C(\Omega).$$

$$\mu \geq \nu, \text{ if and only if } \mu(x) \geq \nu(x) \text{ all } x \geq 0, \quad \mu, \nu \in C^*(\Omega).$$

Denote by $\mathfrak{M}(\Omega)$ the space of all real-valued completely-additive regular set-functions $\mu(E)$, defined for all Borel sets E of Ω. If we put $||\mu|| = \sup_{E \subset \Omega} \mu(E) - \inf_{E \subset \Omega} \mu(E)$, and $\mu \geq \nu$ if and only if $\mu(E) \geq \nu(E)$ for any Borel set $E \subset \Omega$, then $\mathfrak{M}(\Omega)$ is isometric and lattice isomorphic to the conjugate space $C^*(\Omega)$ of $C_R(\Omega)$. The correspondence is given by

$$(4.1) \qquad (x, \mu) = \int_\Omega x(t) \, d\mu(t).$$

PROOF: See Kakutani [12].

We shall denote by $C(\Omega)$ the Banach space of complex-valued continuous functions $x(t)$ defined on Ω.

We now return to consideration of the compact metric space Ω and the transformations $\{A_i\}$ introduced in section 3.

Consider a system whose state may be described by a point in Ω. Let $\phi_i(t)$ be a countable family of continuous functions defined on Ω with the property that $0 \leq \phi_i(t) \leq 1$, and $\sum_1^\infty \phi_i(t) = 1$. Our Markoff process consists in applying at every unit interval of time one of the transformations $\{A_i\}$, A_i being applied with probability $\phi_i(t)$, where $t \in \Omega$ represents the state of the system.

Let $\mu(E)$ be a probability measure defined on the Borel sets of Ω, giving the probability distribution of the initial state of the system. Let

$$(4.2) \qquad T\mu(E) = \sum_{i=1}^{\infty} \int_{A_i^{-1}E} \phi_i(t)\, d\mu(t).$$

Since $\{A_i\}$ is a set of continuous transformations, it is clear that $T\mu$ defines a measure defined on all Borel sets of Ω, and represents the probability measure for the state of the system after unit time.

Since Ω is a compact metric space, it is separable and therefore the concepts of Baire measures and Borel measures coincide. Since Baire measures on Ω are regular, this means that Borel measures are regular. Hence $\mathfrak{M}(\Omega)$, defined in theorem 4.1, consists of the set of all finite Borel measures on Ω. Equation (4.2) defines $T\mu$ for any $\mu \in \mathfrak{M}(\Omega)$. Clearly T is linear and $T\mu$ is a Borel measure, since $\{A_i\}$ are continuous. Moreover, T is a positive linear transformation. If $\mu > 0$ then $||T\mu|| = T\mu(\Omega) = \mu(\Omega) = ||\mu||$ by (4.2). An arbitrary measure μ may be written as $\mu = \mu^+ - \mu^-$, with norm given by $||\mu|| = ||\mu^+|| + ||\mu^-||$. Hence

$$||T\mu|| = ||T\mu^+ - T\mu^-|| \leq ||T\mu^+|| + ||T\mu^-||$$

$$= ||\mu^+|| + ||\mu^-|| = ||\mu||$$

$$\therefore \quad ||T\mu|| \leq ||\mu||.$$

Thus we obtain

<u>LEMMA 4.1.</u> T is a positive linear transformation of $\mathfrak{M}(\Omega)$ into itself of norm 1.

Now consider the space $C(\Omega)$ of continuous complex-valued functions on Ω . We define a linear transformation U of this space into itself by $x \to Ux$, where

$$(4.3) \qquad Ux(t) = \sum_{i=1}^{\infty} \phi_i(t)\, x(A_i t)$$

for every continuous function $x(t) \in C(\Omega)$. It is clear that U is linear and positive.

$$||Ux|| = \sup_{t \in \Omega} |Ux(t)|$$

$$\leq \sup_{t \in \Omega} |x(t)| \qquad \text{since} \quad \Sigma\, \phi_i(t) = 1.$$

$$= ||x||.$$

$\therefore$ $||Ux|| \leq ||x||$, i.e., U is continuous with $||U|| \leq 1$. If $x(t)$ is a constant, then $Ux = x$.

$$\therefore \quad ||U|| = 1.$$

$||U^2|| \leq 1$, but again since U preserves constants, $||U^2|| = 1$, and in general $||U^n|| = 1$. Summing up we have that

LEMMA 4.2. U is a positive linear transformation of $C(\Omega)$ into itself of norm 1, which preserves constants.

By theorem 4.1, $\mathfrak{M}(\Omega)$ is the conjugate space of $C_R(\Omega)$ with identification given by equation (4.1), i.e.,

$$(x, \mu) = \int_{\Omega} x(t)\, d\mu(t).$$

As in Karlin's paper [2], we now connect transformations T and U by means of the following lemma,

LEMMA 4.3. T is the adjoint of U, i.e.,

$$(4.4) \qquad (Ux, \mu) = (x, T\mu) \qquad \text{for each } x \in C_R(\Omega),\ \mu \in \mathfrak{M}(\Omega).$$

PROOF: Let $\mu \geqslant 0$, $x \geqslant 0$.

$$\text{Let } \lambda_n(E) = \sum_{i=1}^{n} \int_{A_i^{-1}E} \phi_i(t)\, d\mu(t).$$

$\lambda_n(E)$ is a positive measure, and

$$\lambda_n(E) \nearrow T\mu(E).$$

Since A_i^{-1} is continuous on its domain, and since by the continuity of A_i, A_i is a Borel-measurable transformation we may write

$$\lambda_n(E) = \sum_{i=1}^{n} \int_{E} \phi_i(A_i^{-1}t)\, d\mu\, (A_i^{-1}t)$$

$$(x, \lambda_n) = \int_\Omega x(t)\, d\lambda_n(t)$$

$$= \sum_{i=1}^{n} \int_\Omega x(t)\, \phi_i(A_i^{-1}t)\, d\mu(A_i^{-1}t)$$

$$= \sum_{i=1}^{n} \int_\Omega x(A_i t)\, \phi_i(t)\, d\mu(t)$$

$$= \int_\Omega \sum_{i=1}^{n} \phi_i(t)\, x(A_i t)\, d\mu(t).$$

Now $(x, \lambda_n) \to (x, T\mu)$ as $n \to \infty$ for

$$(x, T\mu) - (x, \lambda_n) = \int x(t)\, d(T\mu - \lambda_n).$$

By the above $T\mu - \lambda_n$ is a positive measure, and there exists $n_0 \ni n \geq n_0$ $(T\mu - \lambda_n)(\Omega) < \varepsilon/||x||$. Hence for $n \geq n_0$

$$(x, T\mu) - (x, \lambda_n) < \varepsilon,$$

$$\therefore\ (x, \lambda_n) \nearrow (x, T\mu).$$

Hence,

$$(x, T\mu) = \lim(x, \lambda_n) = \int_\Omega \sum_{i=1}^{\infty} \phi_i(t)\, x(A_i t)\, d\mu(t),$$

the limit being taken under the integral sign since all the terms are positive.

$$\therefore\ (x, T\mu) = (Ux, \mu).$$

Since an arbitrary $x \in C_R(\Omega)$ can be written as $x^+ - x^-$ where x^+, x^- are positive and in $C_R(\Omega)$, and similarly for $\mu \in \mathfrak{M}(\Omega)$ the result (4.4) immediately follows, and the lemma is proved.

We now consider a subspace $C_L(\Omega)$ of $C(\Omega)$. We define $C_L(\Omega)$ to be the vector subspace of $C(\Omega)$ consisting of those complex valued functions in $C(\Omega)$, which satisfy a Lipschitz condition, i.e.,

$$(4.5) \qquad M(x) = \sup_{t \neq s} \frac{|x(t) - x(s)|}{\rho(t,s)} < \infty$$

where the sup. is taken over all $t,s \in C(\Omega)$, $t \neq s$. This set of functions $C_L(\Omega)$ <u>forms a Banach space</u> under the norm $||x||_L$, where

$$(4.6) \qquad ||x||_L = ||x|| + M(x).$$

We now verify a very important property of $C_L(\Omega)$.

<u>LEMMA 4.4.</u> The unit sphere in $C_L(\Omega)$ is compact in $C(\Omega)$, i.e.,

$$\Sigma_1 = \{x \in C_L(\Omega) \mid ||x||_L \leq 1\}$$

is compact with respect to norm $||x||$.

<u>PROOF</u>: Let $\{x_\nu\}$ be a sequence in Σ_1. Hence

$$(4.7) \qquad ||x_{\nu_i}|| + \sup_{t \neq s} \frac{|x_{\nu_i}(t) - x_{\nu_i}(s)|}{\rho(t,s)} \leq 1,$$

which implies that

$$|x_\nu(t)| \leq 1,$$

and

$$|x_\nu(t) - x_\nu(s)| \leq \rho(t,s).$$

Thus $\{x_\nu(t)\}$ form a bounded sequence of equicontinuous functions defined on a compact metric space. Hence by Arzela's theorem, there exists a subsequence $\{x_{\nu_i}\}$ converging uniformly to a continuous function $x(t)$, i.e.,

$$||x_{\nu_i} - x|| \to 0.$$

It remains to show that $x \in \Sigma_1$. From (4.7)

$$\frac{|x_{\nu_i}(t) - x_{\nu_i}(s)|}{\rho(t,s)} \leq 1 - ||x_{\nu_i}|| \qquad t \neq s.$$

Let $i \to \infty$ to obtain

$$\sup_{t \neq s} \frac{|x(t) - x(s)|}{\rho(t,s)} \leq 1 - ||x||$$

which proves that $x \in C_L(\Omega)$ and $||x||_L \leq 1$, i.e., $x \in \Sigma_1$. This establishes the lemma.

It is now necessary to impose two further restrictions on $\phi_i(t)$. We assume

(4.8) $\underline{\phi_i(t) \in C_L(\Omega)}$ $\qquad i = 1,2, \ldots$

and also

(4.9) $\underline{M = \sum_1^\infty M(\phi_i) < \infty.}$

This latter condition is automatically satisfied if the number of transformations is finite.

LEMMA 4.5. $\sum_1^\infty \phi_i = 1$, where $\{\phi_i\}$ are considered as elements of $C(\Omega)$, i.e., $\sum_1^\infty \phi_i(t) = 1$, the convergence being uniform.

PROOF: By definition $\sum_1^\infty \phi_i(t) = 1$, so it only remains to prove the uniform convergence of the series. However the partial sums are positive, continuous and converge to a continuous function on a compact set and so by Dini's theorem converge uniformly. As is pointed out by Ionescu Tulcea and Marinescu [8,9], the following is true.

LEMMA 4.6. The transformations U defined by (4.3) applied to the elements of $C_L(\Omega)$ is a continuous transformation of $C_L(\Omega)$ into itself, and

$$(4.10) \qquad ||Ux||_L \leq \lambda ||x||_L + R||x||,$$

where λ, R are constants and $0 < \lambda < 1$, λ being the constant introduced in (3.1).

PROOF: Doeblin and Fortet [4] proved the continuity for a finite number of transformations, but proof goes forward in exactly the same way under our assumptions.

$$Ux(t) - Ux(s) = \Sigma\, \phi_i(t)\, x(A_i t) - \Sigma\, \phi_i(s)\, x(A_i s)$$

$$= \Sigma\, \phi_i(t)(x(A_i t) - x(A_i s)) + \Sigma\, (\phi_i(t) - \phi_i(s))\, x(A_i s)$$

$$\frac{Ux(t) - Ux(s)}{\rho(t,s)} = \Sigma\, \phi_i(t) \left(\frac{x(A_i t) - x(A_i s)}{\rho(A_i t, A_i s)} \right) \cdot \frac{\rho(A_i t, A_i s)}{\rho(t,s)}$$

$$+ \Sigma\, \frac{(\phi_i(t) - \phi_i(s))}{\rho(t,s)}\, x(A_i s).$$

Hence by means of (3.1), (4.5) and (4.9) we obtain

$$M(Ux) \leq \lambda M(x) + M||x||.$$

Hence if $x \in C_L(\Omega)$, $M(Ux) < \infty$, i.e., $Ux \in C_L(\Omega)$.

$$||Ux||_L = ||Ux|| + M(Ux)$$

$$\leq ||x|| + \lambda M(x) + M||x||$$

$$= \lambda M(x) + \lambda ||x|| + (1 + M - \lambda)\, ||x||$$

$$\therefore\ ||Ux||_L \leq \lambda ||x||_L + R||x||$$

where $R = 1 + M - \lambda$.

It follows that

$$||Ux||_L \leq (\lambda + R) \; ||x||_L.$$

Hence U is a continuous linear transformation of $C_L(\Omega)$ into itself satisfying (4.10), and thus the lemma is proved. It is clear from the foregoing that we have

<u>LEMMA 4.7.</u> $C(\Omega)$, $C_L(\Omega)$ satisfy the conditions for the Banach spaces E and B respectively, introduced in section 2. Moreover the transformation U introduced in this section satisfies the conditions for the transformation U introduced in section 2 by means of equations (2.1), (2.2) and (2.3).

The following lemma is a direct consequence of lemma 4.7 and lemma 2.1, using the fact that $||U^n|| = 1$.

<u>LEMMA 4.8.</u> $||U^n||_L \leq C$ (a constant) $\qquad n = 1, 2, \ldots$.

We conclude the section with the following theorem.

<u>THEOREM 4.2.</u> U is a quasi-completely continuous transformation of $C_L(\Omega)$ into itself.

This is proved by Ionescu Tulcea and Marinescu. It is a direct consequence of applying lemmas 4.7, 4.4, 2.2 and the fact that $||U^n|| = 1$ to their theorem, i.e., theorem 2.2.

5. Uniform C - 1 Convergence for U^n on $C_L(\Omega)$.

THEOREM 5.1. There exists a non-zero positive continuous transformation U_1 of $C_L(\Omega)$ into itself such that

$$(5.1) \qquad \left|\left|\frac{1}{n} \sum_{1}^{n} U^k - U_1\right|\right|_L \leq \frac{C}{n} \qquad n = 1,2, \ldots$$

where C is some constant independent of n. Moreover, U_1 preserves constant functions.

PROOF: Since $Ux = x$ if $x(t) = \text{constant}$, it follows that $\lambda = 1$ is an eigen-value of U, which by theorem 4.2 is a quasi-completely continuous transformation of $C_L(\Omega)$ into itself. Using lemma 4.8, i.e., $||U^n||_L \leq C$, we obtain (5.1) from theorem 2.3. That U_1 preserves constant functions is obvious from the fact that (5.1) implies

$$(5.2) \qquad \left|\left|\frac{1}{n} \sum_{1}^{n} U^k x - U_1 x\right|\right|_L \to 0 \qquad x \varepsilon C_L(\Omega),$$

which implies that

$$(5.3) \qquad \left|\left|\frac{1}{n} \sum_{1}^{n} U^k x - U_1 x\right|\right| \to 0 \qquad x \varepsilon C_L(\Omega).$$

The positivity of U_1 follows also from (5.3). Thus the theorem is established.

It is easy to see from the definition of U, that $U^n \phi_i(t)$ represents the probability, that given that the initial state of our system is t, that at the end of the n^{th} time interval, the transformation A_i is applied. Since $\phi_i(t) \varepsilon C_L(\Omega)$ we obtain from (5.1) a mean convergence theorem for the limiting probabilities of applying the transformation A_i, as $n \to \infty$ given that the initial state is t. This has been obtained by Doeblin and Fortet [4], and Ionescu Tulcea and Marinescu [8].

LEMMA 5.1. The following inequalities hold for $n = 1, 2, \ldots$,

$$(5.4) \qquad \left\|\frac{1}{n} \sum_{1}^{n} U^k \phi_i - U_1 \phi_i\right\|_L \leq \frac{C'}{n}$$

and consequently,

$$(5.5) \qquad \left\|\frac{1}{n} \sum_{1}^{n} U^k \phi_i - U_1 \phi_i\right\| \leq \frac{C'}{n}$$

where C' is a constant independent of n and i.

PROOF: This follows from (5.1) and the fact that

$$\|\phi_i\|_L \leq \|\phi_i\| + M(\phi_i)$$

$$\leq 1 + M \qquad \text{by (4.9).}$$

Hence the lemma is established.

6. Strong C - 1 Convergence for U^n on $C(\Omega)$.

From now on we restrict consideration to Ω a compact subset of Euclidean space. Since Ω is compact it follows that, since polynomials are dense in $C(\Omega)$, $C_L(\Omega)$ is dense in $C(\Omega)$. It is for this property that the restriction on Ω has been made.

THEOREM 6.1.

$$(6.1) \qquad \left|\left|\frac{1}{n} \sum_1^n U^k x - U_1 x\right|\right| \to 0 \qquad \text{as} \quad n \to \infty$$

for each $x \varepsilon C(\Omega)$, where U_1 is a continuous transformation of $C(\Omega)$ into itself, and $||U_1|| = 1$. Moreover U_1 as a transformation on $C_L(\Omega)$, is the same as that of the last section.

PROOF: Let $V_n = \frac{1}{n} \sum_1^n U^k$

$$||V_n|| \leq 1$$

but since V_n clearly preserves constants $||V_n|| = 1$. From (5.1) we easily obtain that for $x \varepsilon C_L(\Omega)$

$$||V_n x - U_1 x|| \leq ||V_n x - U_1 x||_L \leq \frac{C}{n} ||x||_L,$$

i.e.,

$$\lim_{n \to \infty} ||V_n x - U_1 x|| = 0 \qquad \text{for} \quad x \varepsilon C_L(\Omega).$$

Since $||V_n|| =1$, we have by the Banach-Steinhaus theorem (since $C_L(\Omega)$ is dense in $C(\Omega)$) that

$$\lim_{n \to \infty} ||V_n x - U_1 x|| = 0$$

where U_1 is a continuous linear transformation of $C(\Omega)$ into itself, and is in fact the unique extension of U_1 defined on $C_L(\Omega)$ to the whole space $C(\Omega)$. Also $||U_1|| \leq 1$, but since U_1 preserves constants, $||U_1|| = 1$. Hence the theorem is established.

Corollary: The following relation holds

(6.2) $$UU_1 = U_1U = U_1 = U_1^2.$$

This follows by standard arguments. See Yosida and Kakutani [10].

Definition: $\Psi_i(t) = U_1\phi_i(t), \quad i = 1,2, \ldots .$

<u>LEMMA 6.1.</u> $\Psi_i \in C_L(\Omega)$, $i = 1,2, \ldots$, and $\sum_{i=1}^{\infty} \Psi_i = 1$, convergence being in the sense of $C(\Omega)$, i.e.,

(6.3) $$\sum_1^{\infty} \Psi_i(t) = 1$$

where the convergence is uniform on Ω .

<u>PROOF:</u> Since U_1 transforms $C_L(\Omega)$ into itself, $\Psi_i \in C_L(\Omega)$. By lemma 4.5, $\sum_1^{\infty} \phi_i = 1$, convergence being in the sense of $C(\Omega)$. Since U_1 is continuous on $C(\Omega)$ by theorem 6.1, we have that

$$\sum_1^{\infty} \Psi_i = 1$$

in the sense of $C(\Omega)$, which proves the lemma.

7. Weak-Star C - 1 Convergence for $T^n\mu$.

THEOREM 7.1. For any measure $\mu \in \mathfrak{M}(\Omega)$, we have the following:

$$(7.1) \qquad \frac{1}{n} \sum_{1}^{n} T^k\mu \rightharpoonup T_1\mu ,$$

where the half-arrow denotes weak-star convergence, i.e.,

$$(x, T_1\mu) = \lim_{n \to \infty} (x, \frac{1}{n} \sum_{1}^{n} T^k\mu),$$

and where T_1 is a positive continuous linear transformation of norm 1 of $\mathfrak{M}(\Omega)$ into itself.

PROOF: From (4.4), we have for $x \in C_R(\Omega)$

$$(Ux, \mu) = (x, T\mu)$$

and in general

$$(x, T^n\mu) = (U^n x, \mu).$$

Thus

$$(7.2) \quad (x, \frac{1}{n} \sum_{1}^{n} T^k\mu) = \frac{1}{n} \sum_{1}^{n} (x, T^k\mu) = \frac{1}{n} \sum_{1}^{n} (U^k x, \mu) = (\frac{1}{n} \sum_{1}^{n} U^k x, \mu).$$

From (6.1) we have

$$\lim_{n \to \infty} \frac{1}{n} \sum_{1}^{n} U^k x = U_1 x \qquad \text{(convergence in sense of } C(\Omega)\text{)}.$$

Let T_1 be the conjugate transformation of U_1, i.e.,

$$(7.3) \quad (x, T_1\mu) = (U_1 x, \mu)$$

$$(U_1 x, \mu) = (\lim_{n \to \infty} \frac{1}{n} \sum_{1}^{n} U^k x, \mu) = \lim_{n \to \infty} (\frac{1}{n} \sum_{1}^{n} U^k x, \mu).$$

Hence by (7.2) and (7.3)

$$(x, T_1\mu) = \lim_{n \to \infty} (x, \frac{1}{n} \sum_{1}^{n} T^k\mu)$$

for all $x \varepsilon C_R(\Omega)$. T_1 being the adjoint of U_1, T_1 is a continuous linear transformation of $\mathfrak{M}(\Omega)$ into itself of norm 1. Let $\mu \geq 0$, then $T\mu \geq 0$. If $x \geq 0$, it follows from (7.1) that $(x, T_1\mu) \geq 0$. Hence $T_1\mu \geq 0$. Thus T_1 is positive. The theorem is therefore established.

Corollary: The following relation holds

(7.4) $$TT_1 = T_1T = T_1 = T_1^2 .$$

This is an immediate consequence of (4.4) and (6.2).

<u>LEMMA 7.1.</u> For any open set $H \subset G$ (defined by (3.5)), $T_1\mu(H) = 0$. If there are only a finite number of transformations A_i, then $T_1\mu(G) = 0$.

<u>PROOF</u>: Let $x(t) \varepsilon C(\Omega)$ where

(7.5) $$\begin{cases} x(t) > 0 & t \varepsilon H \\ x(t) = 0 & t \varepsilon H^c. \end{cases}$$

$H \subset G$, $H^c \supset F$. F, G were defined in section 3. It is clear by the method of definition that there exists N such that for $n \geq N$, $A^{[n]}\Omega \subset H^c$. Thus for $n \geq N$

$$U^n x(t) = \sum \phi_{i_1}(t)\phi_{i_2}(A_{i_1}t) \ldots\ldots \phi_{i_n}(A_{i_{n-1}} \ldots A_{i_1}t)\, x(A_{i_n} \ldots A_{i_1}t)$$

$$= 0 \qquad \text{by (7.5).}$$

Hence $\lim_{n \to \infty} U^n x(t) = 0$. Let μ be any positive measure

$$(x, T_1\mu) = (U_1 x, \mu) = (\lim \frac{1}{n} \sum_1^k U^k x, \mu)$$

$$= (\lim U^n x, \mu)$$

$$= 0.$$

$$\therefore \int x(t)\, d(T_1\mu(t)) = 0$$

$$\therefore \int_H x(t)\, d(T_1\mu(t)) = 0 \qquad \text{by } (7.5).$$

Now T_1 is a positive transformation. Thus $T_1\mu(H) \geq 0$, but since $x(t) > 0$ on H, it follows that $T_1\mu(H) = 0$. Since T_1 is positive and every measure μ may be written as the difference of two positive measures, it follows that $T_1\mu(H) = 0$ for all μ. If the number of transformations is finite then by lemma 3.1, G is an open set. Let μ be a positive measure. There exists a sequence of open sets H_i

$$H_1 \subset H_2 \cdots \subset H_n \subset \cdots \to G$$

$$T_1\mu(H_n) \to T_1\mu(G).$$

Hence

$$T_1\mu(G) = 0.$$

Hence in the case of a finite number of transformations the whole measure is concentrated on F.

Hence the lemma is established.

8. Convergence Theorems for U^n, T^n for a Class of Absorption Problems.

Throughout this section the following assumptions are made

(8.1) $\quad \phi_i(y_i) = 1 \qquad\qquad i = 1,2, \ldots$

i.e., there is absorption at each of the points y_i, and

(8.2) $\quad \phi_i(A_i t) \geq \phi_i(t) \qquad\qquad$ each t,i,

i.e., the probability of applying A_i does not decrease as y_i is approached.

Under these assumptions convergence theorems will be obtained for U^n and T^n. No use will be made of theorem 4.2 (due to Ionescu Tulcea and Marinescu), which gives the quasi-complete continuity of U relative to $C_L(\Omega)$ until lemma 8.4 is reached. Here for the case where the number of transformations $\{A_i\}$ is finite, the quasi-complete continuity of U will be a consequence of our convergence theorems. However in the general case, it has been found necessary to appeal to theorem 4.2 to establish the quasi-complete continuity of U relative to $C_L(\Omega)$, although one suspects that this is unnecessary. Furthermore, until lemma 8.4 is reached no use will be made of sections 5, 6, and 7 which depend on theorem 4.2.

LEMMA 8.1. U preserves the values at the points y_j $(j = 1,2, \ldots)$, i.e.,

(8.3) $\quad Ux(y_i) = x(y_i) \qquad\qquad i = 1,2, \ldots .$

<u>PROOF</u>: $$Ux(t) = \sum_{i=1}^{\infty} \phi_i(t)\, x(A_i t)$$

$$Ux(y_j) = \sum_{i=1}^{\infty} \phi_i(y_j)\, x(A_i y_j).$$

Since $\sum \phi_i(t) = 1$ and $\phi_i(t) \geq 0$ it follows from (8.1) that

$$\phi_i(y_j) = \delta_{ij} \qquad \text{(Kronecker delta)}$$

Hence

$$Ux(y_j) = x(A_j y_j) = x(y_j) \qquad \text{since } A_j y_j = y_j.$$

<u>THEOREM 8.1</u>. If $x(t)$ is a fixed point of U in $C(\Omega)$ having the value zero at each of the vertices y_j, $j = 1,2, \ldots$, then $x(t) \equiv 0$. Two continuous fixed points of U which are equal at y_j, $j = 1,2, \ldots$, are identical.

<u>PROOF:</u> This is an extension of a theorem due to Bellman. Let $x(t)$ be a real-valued continuous fixed point of U such that $x(y_j) = 0$, $(j = 1,2, \ldots)$.

Let t_o be a point where $\max x(t)$ is reached.

$$x(t_o) = Ux(t_o) = \sum \phi_i(t_o)\, x(A_i t_o)$$

$$\therefore \sum_{1}^{\infty} \phi_i(t_o)[x(t_o) - x(A_i t_o)] = 0.$$

This is a series of positive or zero terms. Therefore

$$\phi_i(t_o)[x(t_o) - x(A_i t_o)] = 0, \qquad i = 1,2, \ldots .$$

Since $\sum \phi_i(t_o) = 1$, $\phi_i(t_o) \geq 0$, there exists at least one i, say $i = k$, such that $\phi_k(t_o) > 0$. Hence

$$x(t_o) = x(A_k t_o).$$

Repeating the above argument with t_o replaced by $A_k t_o$, we obtain

$$\sum_1^\infty \phi_i(A_k t_o)[x(A_k t_o) - x(A_i A_k t_o)] = 0.$$

By (8.2), $\phi_k(A_k t_o) > \phi_k(t_o)$, which is greater than zero. Hence

$$x(A_k t_o) = x(A_k^2 t_o).$$

Repeating this process we obtain

$$x(t_o) = x(A_k t_o) = \qquad = x(A_k^n t_o) = x(\lim A_k^n t_o) = x(y_k) = 0.$$

Hence

$$\max x(t) = 0.$$

Applying the same argument to $y(t) = -x(t)$ we obtain $\min x(t) = 0.$ Hence

$$x(t) \equiv 0.$$

The proof for a complex-valued fixed point follows by separating it into its real and imaginary parts. The second part of the theorem follows from the first part.

<u>THEOREM 8.2.</u> U^n converges strongly on $C(\Omega)$, i.e., there exists a continuous transformation U_1 of $C(\Omega)$ into itself such that

(8.4) $$\lim_{n \to \infty} ||U^n x - U_1 x|| = 0$$

for each $x \in C(\Omega)$.

<u>PROOF</u>: The proof is in several stages.

(I) For each $x \in C_L(\Omega)$, $\{U^n x(t)\}$ form a uniformly bounded family of uniformly equicontinuous functions on Ω .

PROOF: Since $||U^n||_L \leq C$ (by lemma 4.8)

$$||U^n x||_L \leq C||x||_L$$

and hence in particular

$$M(U^n x) \leq C||x||_L$$

or

(8.5) $$|U^n x(t) - U^n x(s)| \leq C||x||_L \ \rho(t,s).$$

Hence $\{ U^n x(t) \}$ form an equicontinuous family of functions on Ω for each $x \in C_L(\Omega)$. Since $||U^n|| = 1$ the family is uniformly bounded. Thus the first part of the proof is established.

(II) There exists a number $\delta > 0$, such that at each point $t \in \Omega$ there exists an index i (depending on t) for which $\phi_i(t) > \delta$. (δ is independent of t.)

PROOF: We set $\theta(t) = \sup_i \phi_i(t)$. $\theta(t)$ is a lower semi-continuous function on Ω and therefore, on the compact set Ω , reaches its minimum. Since $\sum \phi_i(t) = 1$, and each $\phi_i(t) \geq 0$, it follows that at each point i there exists at least one index i such that $\phi_i(t) > 0$. Hence $\theta(t) > 0$ for each $t \in \Omega$, but since it reaches its minimum δ' (say), $\delta' > 0$ and $\theta(t) \geq \delta'$. Let δ be any positive number $< \delta'$. It is clear that the assertion is fulfilled for this choice of δ, and the second part of the theorem is established.

(III) Let $x(t) \in C_L(\Omega)$, $x(y_j) = 0$ ($j = 1,2, \dots$), then given any positive number ε, there exist spheres $S_i(\varepsilon)$ with centers y_i ($i = 1,2, \dots$) and each of the radius η , such that if t is in any one of the spheres, $|U^n x(t)| < \varepsilon/2$, all n. Moreover, there exists a fixed integer n_o independent of t and i such that $A_i^{n_o} t \varepsilon S_i(\varepsilon)$ for all $t \varepsilon \Omega$,$(i = 1,2, \ldots)$.

PROOF: We observe that from (1) by equicontinuity we may choose η such that $\rho(t,s) < \eta$ implies that $|U^n x(t) - U^n x(s)| < \varepsilon/2$ for all n. Choose $S_i(\varepsilon)$ to be the sphere of radius η about y_i $(i = 1,2, \ldots)$. Then for $t \varepsilon S_i(\varepsilon)$

$$|U^n x(t) - U^n x(y_i)| < \varepsilon/2,$$

but $x(y_i) = 0$ and therefore by lemma 8.1, $U^n x(y_i) = 0$. Hence for $t \varepsilon S_i(\varepsilon)$ and all n,

$$|U^n x(t)| < \varepsilon/2.$$

It is an immediate consequence of lemma 3.1, that there exists an integer n_o, such that $A_i^{n_o} t \varepsilon S_i$ for each i and all $t \varepsilon \Omega$. Hence (III) is established.

(IV) Let $x(t) \varepsilon C_L(\Omega)$, $x(y_j) = 0$. Then given any positive number ε, there exists an integer m such that

$$||U^m x|| \leq \varepsilon.$$

PROOF: Suppose this not true, i.e., $||U^n x|| > \varepsilon$ all n. Let $||x|| = H$. We assert that under this assumption

$$||U^{kn_o} x|| \leq \alpha^k H \qquad k = 1,2, \ldots ,$$

where n_o is the integer introduced in (III) and

$$\alpha = 1 - \frac{1}{2}\delta^{n_o} < 1$$

where δ is the positive number introduced in (II). We prove this assertion by induction. It is clearly true for $k = 0$. Let it be true for k,

$$(8.6)\quad U^{(k+1)n_o}x(t) = U^{n_o}U^{kn_o}x(t)$$

$$= \sum \phi_{i_1}(t)\phi_{i_2}(A_{i_1}t)\ \cdots\ \phi_{i_{n_o}}(A_{i_{n_o-1}}\cdots A_{i_1}t)U^{kn_o}x(A_{i_{n_o}}\cdots A_{i_1}t).$$

Consider t fixed. In accordance with (II), we can choose a positive integer ℓ so that $\phi_\ell(t) > \delta$. But then by assumption (8.2) $\phi_\ell(A_\ell t) > \delta$, $\phi_\ell(A_\ell^2 t) > \delta$ etc., so that

$$(8.7)\qquad \phi_\ell(t)\ \phi_\ell(A_\ell t)\ \cdots\ \phi_\ell(A_\ell^{n_o-1}t) > \delta^{n_o}.$$

From (8.6)

$$|U^{(k+1)n_o}x(t)| \le \phi_\ell(t)\ \phi_\ell(A_\ell t)\ \cdots\ \phi_\ell(A_\ell^{n_o-1}t)|\,U^{kn_o}X(A_\ell^{n_o}t)|$$

$$+ \sum{}'\phi_{i_1}(t)\phi_{i_2}(A_{i_1}t)\ \cdots\ \phi_{i_{n_o}}(A_{i_{n_o-1}}\cdots A_{i_1}t)|U^{kn_o}x(A_{i_{n_o}}\cdots A_{i_1}t)$$

where $\sum{}'$ denotes that one term has been omitted in the summation, i.e., $i_1 = i_2 = \cdots = i_{n_o} = \ell$. Replace $|U^{kn_o}x(A_\ell^{n_o}t)|$ by $(|U^{kn_o}x(A_\ell^{n_o}t)| - \alpha^k H) + \alpha^k H$ and use $||U^{kn_o}x|| \le \alpha^k H$ to obtain

$$|U^{(k+1)n_o}x(t)| \le \phi_\ell(t)\ \phi_\ell(A_\ell t)\ \cdots\ \phi_\ell(A_\ell^{n_o-1}t)(|U^{kn_o}x(A_\ell^{n_o}t)| - \alpha^k H)$$

$$+ \sum \phi_{i_1}(t)\phi_{i_2}(A_{i_1}t)\ \cdots\ \phi_{i_{n_o}}(A_{i_{n_o-1}}\cdots A_{i_1}t)\ \alpha^k H$$

by the induction assumption. Thus

$$(8.8)\quad |U^{(k+1)n_o}x(t)| \le -\phi_\ell(t)\phi_\ell(A_\ell t)\ \cdots\ \phi_\ell(A_\ell^{n_o-1}t)(\alpha^k H - |U^{kn_o}x(A_\ell^{n_o}t)|)$$

$$+ \alpha^k H \qquad (\text{since } \sum \phi_i(t) = 1).$$

By (III) $A_\ell^{n_o}t \in S_\ell(\varepsilon)$ and hence

$$(8.9)\quad \alpha^k H - |U^{kn_o}x(A_\ell^{n_o}t)| > \alpha^k H - \varepsilon/2.$$

By the induction hypothesis, $||U^{kn_o}x|| \leq \alpha^k H$ and then it follows that $\alpha^k H > \varepsilon$, for otherwise the assumption that $||U^n x|| > \varepsilon$, all n, would be violated. Hence in particular, $\varepsilon/2 < \frac{\alpha^k H}{2}$, and from (8.9) we have

$$\alpha^k H - |U^{kn_o}x(A_\ell^{n_o}t)| > \alpha^k H - \frac{\alpha^k H}{2} = \frac{\alpha^k H}{2}.$$

Therefore by (8.7), (8.8)

$$|U^{(k+1)n_o}x(t)| \leq - \delta^{n_o}\frac{\alpha^k H}{2} + \alpha^k H = \alpha^k H \left(1 - \frac{\delta^{n_o}}{2}\right) = \alpha^{k+1} H.$$

The right-hand side of this relation is independent of t, so that

$$||U^{(k+1)n_o}x|| \leq \alpha^{k+1} H.$$

Hence our assertion that $||U^{kn_o}x|| \leq \alpha^k H$, $k = 1,2, \ldots$, has been established. However this contradicts our assumption that $||U^n x|| > \varepsilon$, all n. Hence there exists an integer m such that

$$||U^m x|| \leq \varepsilon.$$

Hence we have established (IV).

(V) If $x \varepsilon C_L(\Omega)$, $x(y_j) = 0$, then $||U^n x|| \to 0$.

PROOF: This is an immediate consequence of (IV) and the fact that $||U|| = 1$.

(VI) For each $x \varepsilon C_L(\Omega)$, there exists an element $y \varepsilon C(\Omega)$ such that

$$\lim_{n \to \infty} ||U^n x - y|| = 0.$$

PROOF: By Arzela's theorem we have, since by (1) $\{U^n x(t)\}$ form a uniformly bounded set of uniformly equicontinuous functions defined on a compact metric space, that there exists a subsequence $\{U^{n_i} x(t)\}$ such that

$$U^{n_i} x(t) \to y(t)$$

the convergence being uniform. Hence $y(t)$ is in $C(\Omega)$ and we have

$$(8.10) \qquad \lim_{i \to \infty} \| U^{n_i} x - y \| = 0.$$

Since U is continuous

$$(8.11) \qquad \lim_{i \to \infty} \| U^{n_i + 1} x - Uy \| = 0.$$

Hence

$$(8.12) \qquad \lim_{i \to \infty} U^{n_i}(x - Ux) = y - Uy,$$

the convergence being in the sense of $C(\Omega)$. Now since by lemma 8.1 Ux preserves values at the points y_j, and also since $Ux \in C_L(\Omega)$, we have that $x - Ux \in C_L(\Omega)$, with value zero at each of points y_j. Hence by (IV)

$$\| U^n (x - Ux) \| \to 0.$$

Therefore (8.12) gives

$$y = Uy.$$

Thus y is a fixed point of U and by (8.10) and lemma 8.1, it is clear that it has values $x(y_j)$ at the points y_j. However by theorem 8.1, such a fixed point is unique. Thus y is independent of the subsequence $\{U^{n_i} x\}$. Hence

$$\lim_{n \to \infty} ||U^n x - y|| = 0.$$

Thus (VI) is established.

(VII) Theorem 8.2 now follows, for if $x \varepsilon C(\Omega)$, we have by the Banach-Steinhaus theorem, since $C_L(\Omega)$ is dense in $C(\Omega)$ and $||U^n|| = 1$, that

$$\lim_{n \to \infty} ||U^n x - U_1 x|| = 0$$

where U_1 is a continuous transformation of $C(\Omega)$ into itself and $||U_1|| \leq 1$. However since U preserves constants, so does U_1 and therefore $||U_1|| = 1$. Hence we have established Theorem 8.2.

<u>THEOREM 8.3.</u> U_1 is a continuous linear transformation of $C_L(\Omega)$ into itself and

(8.13) $$\lim_{n \to \infty} ||U^n - U_1||_L = 0.$$

<u>PROOF:</u> Since $||U^n|| = 1$, and $||U^n||_L \leq C$ (by lemma 4.8) we have by applying lemmas 4.4, 4.7 and theorem 8.2 to theorem 2.1, the fact that U_1 is a continuous transformation of $C_L(\Omega)$ into itself, with

(8.14) $$||U_1||_L \leq C$$

together with (8.13). Hence the theorem is established.

<u>LEMMA 8.2.</u> U_1 preserves values at the points y_j, i.e., for $x \varepsilon C(\Omega)$

$$U_1 x(y_j) = x(y_j) \qquad j = 1,2, \dots .$$

<u>PROOF:</u> This is a trivial deduction from theorem 8.2 and the fact that U preserves the values of a function at the points y_j $(j = 1,2, \dots)$.

<u>LEMMA 8.3.</u> $\Psi_i(t) = U_1\phi_i(t)$ is the unique continuous fixed point of U having the values 1 at y_i and zeros at all y_j $(j \neq i)$. Moreover

(8.15) $\sum_{1}^{\infty} \Psi_i = 1$ the convergence being that of $C(\Omega)$.

<u>PROOF:</u> $U\Psi_i(t) = UU_1\phi_i(t) = U_1\phi_i(t) = \Psi_i(t)$, since $UU_1 = U$ by (8.4).

Thus Ψ_i is a fixed point of U. Since U_1 preserves values at the points y_j (lemma 8.2) we have from the assumption (8.1) that

(8.16) $\Psi_i(y_j) = U_1\phi_i(y_j) = \phi_i(y_j) = \delta_{ij}$ (Kronecker delta).

Thus $\Psi_i(t)$ is a fixed point of U having the value 1 at y_i and value zero at all y_j $(j \neq i)$, and thus by theorem 8.1 it is unique. From theorem 8.3, $\Psi_i = U_1\phi_i$ is in $C_L(\Omega)$. By lemma 4.5 $\Sigma\, \phi_i = 1$, convergence being in the sense of $C(\Omega)$, but since U_1 is continuous on $C(\Omega)$, it follows that

$$\sum_{1}^{\infty} \Psi_i = 1,$$

the convergence being in the sense of $C(\Omega)$. Hence the lemma is established.

<u>THEOREM 8.4.</u> The following is an explicit form for $U_1 x$ where $x \in C(\Omega)$,

(8.17) $U_1 x = \sum_{i=1}^{\infty} x(y_i)\, \Psi_i$

the convergence being in the sense of $C(\Omega)$.

PROOF: Let

$$q(t) = \sum_{i=1}^{\infty} x(y_i)\ \Psi_i(t)$$

$$\leq \max |x(t)| \sum_{i=1}^{\infty} \Psi_i(t).$$

Since $\sum_{i=1}^{\infty} \Psi_i(t) = 1$, the convergence being uniform (by 8.15) it follows that the series for $q(t)$ is uniformly convergent. Hence $q(t) \varepsilon C(\Omega)$. We thus have $q = \sum_{i=1}^{\infty} x(y_i)\ \Psi_i$, the convergence being in the sense of $C(\Omega)$. Hence using lemma 8.3

$$Uq = \sum_{i=1}^{\infty} x(y_i)\ U\Psi_i = \sum_{i=1}^{\infty} x(y_i)\ \Psi_i = q.$$

Thus by lemma 8.3 and theorem 8.1 we have that q is the unique fixed point having the values $x(y_j)$ at the points y_j. Since $UU_1 = U_1$ (by 8.4) we have by lemma 8.2, that $U_1 x$ is also a fixed point of U_1 having the values $x(y_j)$ at the points y_j. Hence by the uniqueness theorem 8.1, $U_1 x = q$, and thus the theorem is proved.

LEMMA 8.4. U is a quasi-completely continuous transformation of $C_L(\Omega)$ into itself with $\lambda = 1$ as the only eigen-value of modulus 1. Also

(8.18) $\quad U^n \phi_i(t) \to \Psi_i(t)$

the convergence being uniform geometric with respect to t and i.

PROOF: If the number of transformations A_i is finite, we see from (8.17) that U_1 maps $C_L(\Omega)$ into a finite-dimensional space and so is completely continuous. Hence from (8.13) we see that U is a quasi-completely continuous transformation of $C_L(\Omega)$ into itself. Observe that we have not appealed to theorem 4.2 (theorem of Ionescu Tulcea and Marinescu). However for the general case, we appeal to theorem 4.2 to obtain the quasi-complete continuity of U with respect to $C_L(\Omega)$.

Using lemmas 4.7, 4.8 and equation (8.13) we may apply theorem 2.3 to U relative to $C_L(\Omega)$, to obtain from (2.18) that U_1 is completely continuous relative to $C_L(\Omega)$, and from (2.20) that $\lambda = 1$ is the only eigen-value of modulus 1. Equation (2.21) gives

$$||U^n x - U_1 x||_L \leq \frac{H}{(1+h)^n} \; ||x||_L.$$

In particular,

$$||U^n \phi_i - U_1 \phi_i|| \leq \frac{H}{(1+h)^n} \; ||\phi_i||_L$$

$$\leq \frac{H'}{(1+h)^n} \qquad H', h \text{ being constants (by 4.9).}$$

Since $U_1 \phi_i = \Psi_i$, (8.18) is established and the lemma is proved.

THEOREM 8.5. For any $\mu \in \mathfrak{M}(\Omega)$

(8.19) $\quad T^n \mu \rightharpoonup T_1 \mu \qquad$ (weak-star)

where T_1 is the adjoint of U_1. The explicit form of T_1 is given by

(8.20) $\quad T_1 \mu = \sum_1^\infty I_j \left(\int_\Omega \Psi_j(t) \, d\mu(t) \right)$

where I_j is a probability measure with all its measure concentrated at the point y_j.

PROOF: (8.19) is merely a translation of theorem 8.2 from $C_R(\Omega)$ into its conjugate space. The proof is exactly the same as that of theorem 7.1. We now obtain the explicit form of $T_1\mu$. From (8.17)

$$U_1 x = \sum_1^\infty x(y_i)\, \Psi_i.$$

Hence $(U_1x, \mu) = \sum_1^\infty x(y_i)\, (\Psi_i, \mu)$. Let $T'\mu = \sum_1^\infty I_j(\Psi_j, \mu)$. This is clearly an element in $\mathfrak{M}(\Omega)$. It is also clear that

$$(U_1x, \mu) = (x, T'\mu) \qquad \text{all } x \in C_R(\Omega).$$

Hence

$$T' = U_1^* = T_1.$$

Thus

$$T_1\mu = \sum_1^\infty I_j(\Psi_j, \mu)$$

and the theorem is established.

Probability Interpretation of $\Psi_i(t)$.

In section 5, it was pointed out that $U^n\phi_i(t)$ represents the probability that given the initial state of our systems was t, that at the end of the n^{th} time interval, the transformation A_i is applied. $\Psi_i(t) = \lim_{n \to \infty} U^n\phi_i(t)$ thus represents the limiting probability of applying A_i given that initially the state of the system was t.

Another point of view is obtained from (8.20). If $\mu(t) = I_{t_o}$ i.e, the probability measure concentrated on point y_j, then

$$T_1\mu = \sum_1^\infty \Psi_j(t_o)\, I_j$$

so that $\Psi_j(t_o)$ gives the probability that if the initial state is t_o, the limiting state is y_j.

To sum up, we have two probability interpretations for $\Psi_j(t)$:

(1) Limiting probability as $n \to \infty$ that at the n^{th} step A_i is applied, where the initial state is t.

(2) Probability that the limiting state is y_j, given that the initial state is t.

9. Properties of U and T for the case when there exists an index k such that $\phi_k(t) > 0$ for all $t \in \Omega$ with the possible exception of y_k.

In this section we assume the existence of an index k such that

$$(9.1) \qquad \phi_k(t) > 0 \quad \text{for} \quad t \neq y_k \quad (0 \leq \phi_k(y_k) \leq 1).$$

We generalize a method of Bellman to obtain a fixed point theorem similar to that of theorem 8.1

THEOREM 9.1. Under the above condition, the only fixed points of U in $C(\Omega)$ are constants.

PROOF: Let $x(t)$ be a fixed point. Let t_o be a point where the maximum is achieved.

Assume that $\max x(t) \neq x(y_k)$.

$$x(t_o) = Ux(t_o) = \sum_1^\infty \phi_i(t_o)\, x(A_i t_o)$$

$$\sum_1^\infty \phi_i(t_o)[x(t_o) - x(A_i t_o)] = 0.$$

This is a series of positive terms. Hence

$$\phi_k(t_o)[x(t_o) - x(A_k t_o)] = 0.$$

If $\phi_k(t_o) = 0$ then by (9.1) $t_o = y_k$ and $\max x(t) = x(y_k)$ contrary to the assumption. Hence

$$x(t_o) = x(A_k t_o).$$

Thus $A_k t_o$ is a maximum point for $x(t)$. Apply the above argument with t_o replaced by $A_k t_o$, to obtain

$$\phi_k(A_k t_o)[x(A_k t_o) - x(A_k^2 t_o)] = 0.$$

Again, $\phi_k(A_k t_o) \neq 0$, otherwise the assumption is contradicted. Hence

$$x(A_k t_o) = x(A_k^2 t_o)$$

and proceeding in this way we obtain

$$x(t_o) = x(A_k t_o) = \qquad = x(A_k^n t_o) = \lim_{n \to \infty} x(A_k^n t_o) = x(y_k)$$

since $x(t)$ is continuous, but this violates the assumption that

$$x(t_o) \neq x(y_k).$$

Hence

$$x(t_o) = x(y_k)$$

$$\therefore \quad \max x(t) = x(y_k)$$

Consider the function $y(t) = - x(t)$. It is a continuous fixed point of U. Hence

$$\max y(t) = y(y_k)$$

or

$$\min x(t) = x(y_k)$$

$$\therefore \qquad \max x(t) = \min y(t) = x(y_k)$$

or

$$x(t) = \text{constant}.$$

The proof for a complex-valued fixed point is immediate by splitting into real and imaginary parts.

Corollary: $U_1 x(t) = k$ (a constant dependent on x).

<u>PROOF:</u> From (6.2) we have that $U_1 x$ is a fixed point of U, for each $x \varepsilon C(\Omega)$. Hence the corollary follows from the theorem.

<u>THEOREM 9.2.</u> If μ is any probability measure, i.e., $\mu(E) \geq 0$ for every Borel set and $\mu(\Omega) = 1$, then

$$T_1 \mu (E) = \vee(E)$$

where $\vee(E)$ is a probability measure independent of $\mu(E)$, so that we may say $\frac{1}{n} \sum_1^n T^k \mu$ converges in distribution to $\vee$, where $\vee$ is a probability measure independent of the initial distribution μ. Moreover $\vee$ is the unique probability distribution which is a fixed point of T.

<u>PROOF:</u> Let μ, μ' be two probability measures. Let $T_1 \mu = \vee$, so that by (7.1), $\frac{1}{n} \sum_1^n T^k \mu \rightharpoonup \vee$ (weak-star convergence). Let $T_1 \mu' = \vee'$, so that $\frac{1}{n} \sum_1^n T^k \mu' \rightharpoonup \vee$. For any $x \varepsilon C_R(\Omega)$

$$
\begin{aligned}
(x, \vee - \vee') &= \lim_{n \to \infty} (x, \frac{1}{n} \sum_1^n T^k \mu - \frac{1}{n} \sum_1^n T^k \mu') \\
&= \lim_{n \to \infty} (x, \frac{1}{n} \sum_1^n T^k(\mu - \mu')) \\
&= \lim_{n \to \infty} (\frac{1}{n} \sum_1^n U^k x, \mu - \mu') \\
&= (U_1 x, \mu - \mu') \quad \text{by (6.1)} \\
&= (k_x, \mu - \mu') \quad \text{by corollary to last theorem} \\
&= \int k_x \, d(\mu - \mu') \\
&= k_x(\int \mu - \int \mu') \\
&= 0.
\end{aligned}
$$

Hence $\nu = \nu'$. From (7.4), $T_1\mu = \nu$ is a fixed point of T. Let ν_o be another fixed point of T which is a probability measure, then $T_1 \nu_o = \nu_o$ but from above, $T_1 \nu_o = \nu$. Hence $\nu = \nu_o$. Thus the theorem is proved.

10. Convergence theorems for U^n and T^n for a class of problems similar to those considered in section 9.

In this section we make the assumption that

$$\sum \phi_i(t)\, \phi_i(s) \neq 0 \qquad t,s \in \Omega . \tag{10.1}$$

This is almost equivalent to the class of problems considered in section 9, for if (9.1) were strengthened to $\phi_k(t) > 0$ all t, for some k, then (10.1) would be consequence.

Under this assumption a convergence theorem for U^n will be obtained by straightforward application of the method of Karlin [2]. As is pointed out by Ionescu Tulcea and Marinescu, this could be obtained by a method of Ocinescu and Mihoc [3], but according to their theorem as stated, this would apply to the case where $\phi_k(t) > 0$ all t, for some k.

We shall not make use in this section of the theorem 2.2 or its consequence, theorem 4.2, to derive the quasi-complete continuity of U with respect to $C_L(\Omega)$. This will be a consequence of our convergence theorem.

We first prove the following lemma.

LEMMA 10.1. Let two particles start from t, s respectively. We say that a success occurs at a given unit of time if the same transformation A_i is applied to both particles at the same time, otherwise we say that failure occurs. Then a success run of length r (r any positive integer) is certain to occur in finite time.

PROOF: The probability of success at initial step is $\sum_{1}^{\infty} \phi_i(t)\phi_i(s)$. This being the sum of a uniformly convergent series of continuous functions (c.f. lemma 4.5) on the product spaces $\Omega \times \Omega$ is continuous. Hence it achieves its minimum δ which by (10.1) is > 0. Thus the probability of success at the initial step is $\geq \delta$. The conditional probability of success at the second step is $\geq \delta$,and in general the conditional probability of success at the k^{th} step is $\geq \delta$. Consequently, by the theory of recurrent events, a success run of length r is certain to happen in finite time. Thus the lemma is established.

THEOREM 10.1. U^n converges strongly on $C(\Omega)$, i.e.,

(10.2) $$\lim_{n \to \infty} ||U^n x - k_x|| = 0$$

for each $x \varepsilon C(\Omega)$, where k_x is a constant function dependent on x.

PROOF: Let $x(t) \varepsilon C_L(\Omega)$.

$$U^n x(t) - U^n x(s) = \sum_{i_1, j_1} \phi_{i_1}(t)\phi_{j_1}(s)[U^{n-1}x(A_{i_1} t) - U^{n-1}x(A_{j_1} s)]$$

$$= \sum \phi_{i_1}(t)\phi_{j_1}(s)\phi_{i_2}(A_{i_1} t)\phi_{j_2}(A_{j_1} s)[U^{n-2}x(A_{i_2} A_{i_1} t) - U^{n-2}x(A_{j_2} A_{j_1} s)]$$

$$= \sum \phi_{i_1}(t)\phi_{j_1}(s)\phi_{i_2}(A_{i_1} t)\phi_{j_2}(A_{j_1} s) \ldots \phi_{i_r}(A_{i_{r-1}} \ldots A_{i_1} t)\phi_{j_r}(A_{j_{r-1}} \ldots A_{j_1} s)$$

$$[U^{n-r}x(A_{i_r} \ldots A_{i_1} t) - U^{n-r}x(A_{j_r} \ldots A_{j_1} s)].$$

We now divide the right-hand side into two parts S_1, S_2. We take all the terms for which $i_k = j_k$ $(k = 1, \ldots, r)$, and place these in S_1.

The sum of the coefficients of these terms clearly represents the probability of a success run of length r occurring in the first r trials of model discussed in lemma 10.1. The remaining terms are placed in S_2, and clearly the sum of the coefficient of these terms represents the probability of not obtaining a success run of length r ending on the r^{th} trial. We now make a further reduction of terms in S_2 as above. We transfer to S_1 terms of the type

$$\phi_{i_1}(t)\phi_{j_1}(s)\phi_{i_2}(A_{i_1}t)\phi_{i_2}(A_{j_1}s)\dots\ \phi_{i_{r+1}}(A_{i_r}A_{i_{r-1}}\dots A_{i_2}A_{i_1}t)$$

$$\phi_{i_{r+1}}(A_{i_r}\dots A_{i_2}A_{j_1}s)[U^{n-r-1}x(A_{i_{r+1}}A_{i_r}\dots A_{i_1}t)-U^{n-r-1}x(A_{i_{r+1}}\dots A_{i_2}A_{j_1}s)]$$

Observe that $i_1 \neq j_1$, otherwise this term would have been included in S_2. The sum of the coefficients of this set of terms is clearly the probability of a success run of length r ending at the $(r+1)^{st}$ trial in the model discussed in lemma 10.1, where the first trial consisted of failure. Thus the sum of the coefficients in S_1 represents the probability of a first success run of length r ending at the r^{th} or the $(r+1)^{st}$ trial. Hence the sum of the coefficient of the terms in S_2 represents the probability of no success run of length r in $r+1$ trials. Proceeding in this manner

$$(10.3) \qquad U^n x(t) - U^n x(s) = S_1 + S_2$$

where

$$|S_2| \leq 2\ (\text{Probability of no success run of length } r \text{ in } n \text{ trials}) \max |x(t)|.$$

Let $n \to \infty$, by lemma 10.1

$$(10.4) \qquad S_2 \to 0 \quad \text{uniformly with respect to } t,s \in \Omega.$$

Consider S_1. It is of the form

$$\Sigma\, p[U^k x(A_{i_r} \dots A_{i_1}\tau) - U^k x(A_{i_r} \dots A_{i_1}\sigma)]$$

where $\Sigma\, p \leq 1$.

$$|U^k x(A_{i_r} \dots A_{i_1}\tau) - U^k x(A_{i_r} \dots A_{i_1}\sigma)| \leq M(U^k x) \,.\, \rho(A_{i_r} \dots A_{i_1}\tau, A_{i_r} \dots A_{i_1}\sigma).$$

Apply lemma 4.8, i.e., $||U^n||_L \leq C$ and (3.1) to obtain

$$|U^k x(A_{i_r} \dots A_{i_1}\tau) - U^k x(A_{i_r} \dots A_{i_1}\sigma)| \leq C\, \lambda^r\, ||x||_L\, \rho(\tau, \sigma)$$

$$\leq D\, \lambda^r\, ||x||_L$$ since $\rho(\tau,\sigma)$ is bounded on a compact set.

$$\therefore \quad |S_1| \leq D\, \lambda^r\, ||x||_L\, \Sigma\, p$$

$$\leq D\, \lambda^r\, ||x||_L.$$

Hence for $n \geq r$

$$|U^n x(t) - U^n x(s)| = |S_1 + S_2|$$

$$\leq |S_1| + |S_2|$$

$$\leq D\, \lambda^r\, ||x||_L + |S_2|.$$

Choose $r \ni D\, \lambda^r\, ||x||_L \leq \varepsilon/2$. Having chosen r, by (10.4) we can choose N so that for $n \geq N$, $|S_2| \leq \varepsilon$. Hence $n \geq N$

$$|U^n x(t) - U^n x(s)| \leq \varepsilon.$$

Hence

(10.5) $$\lim_{n \to \infty} |U^n x(t) - U^n x(s)| = 0,$$

the convergence being uniform on $\Omega \times \Omega$. Consider the set $\{U^n x(s)\}$

where s is some fixed point in Ω . This is a bounded set of numbers (bounded by $\max |x(t)|$). Hence there is a convergent subsequence such that

$$U^{n_i}x(s) \longrightarrow k_x.$$

For any $t \varepsilon \Omega$,

$$U^{n_i}x(t) - k_x = (U^{n_i}x(t) - U^{n_i}x(s)) + (U^{n_i}x(s) - k_x).$$

It follows by (10.5) that

$$\lim_{i \to \infty} |U^{n_i}x(t) - k_x| = 0,$$

k_x being a constant dependent on x, and the convergence being uniform with respect to t. Choose i such that

$$|U^{n_i}x(t) - k_x| \leq \varepsilon \qquad \text{all} \quad t \varepsilon \Omega$$

$$|U^{n_i+1}x(t) - k_x| = \left| \sum_{j=1}^{\infty} \phi_j(t)(U^{n_i}x(A_jt) - k_x) \right|$$

$$\leq \sum_{j=1}^{\infty} \phi_j(t)|U^{n_i}x(A_jt) - k_x|$$

$$\leq \varepsilon,$$

and in general,

$$|U^{n_i+p}x(t) - k_x| \leq \varepsilon.$$

Hence $|U^n x(t) - k_x|$ converges uniformly to zero, i.e.,

$$\lim_{n \to \infty} ||U^n x - k_x|| = 0 \qquad \text{for} \quad x \varepsilon C_L(\Omega).$$

Let $x(t) \varepsilon C(\Omega)$. Since $||U^n|| = 1$ and $C_L(\Omega)$ is dense in $C(\Omega)$, it follows by the Banach-Steinhaus theorem that

$$\lim_{n \to \infty} ||U^n x - U_1 x|| = 0$$

where U_1 is a continuous transformation of $C(\Omega)$ into itself such that $||U_1|| \leq 1$. Since U_1 obviously preserves constants, $||U_1|| = 1$. Since $U_1 x = k_x$ for $x \in C_L(\Omega)$,

(10.6) $$U_1 x = k_x$$

a constant dependent on x, for all $x \in C(\Omega)$. Hence for $x \in C(\Omega)$

$$\lim_{n \to \infty} ||U^n x - k_x|| = 0,$$

and the theorem is established.

<u>THEOREM 10.2.</u> U_1 defined in the last section by

$$U_1 x = k_x$$

is a continuous linear transformation of $C(\Omega)$ into $C(\Omega)$ and $C_L(\Omega)$ into $C_L(\Omega)$. Moreover

(10.7) $$\lim_{n \to \infty} ||U^n - U_1||_L = 0$$

and U is a quasi-completely continuous transformation of $C_L(\Omega)$ into itself.

<u>PROOF:</u> We have $||U^n|| = 1$. Also $||U^n||_L \leq C$ by lemma 4.8. Hence by means of lemmas 4.4, 4.7, and theorem 10.1 we see that the hypothesis of theorem 2.1 is satisfied and hence U_1 is a continuous linear transformation of $C_L(\Omega)$ into itself and

$$\lim_{n \to \infty} ||U^n - U_1||_L = 0.$$

Since U_1 maps $C_L(\Omega)$ into the one-dimensional space of constants, it is completely continuous. Hence U is quasi-completely continuous as a transformation of $C_L(\Omega)$ into $C_L(\Omega)$. This proves the theorem.

Observe again as remarked at the beginning of the section, that no appeal was made to theorem 2.2 of Ionescu Tulcea and Marinescu [8,9] in order to obtain this result except through the easily proved fact that $||U^n||_L \leq C$, which we gave as a separate lemma (lemma 4.8). Now that we have obtained the quasi-complete continuity of U, all the results of sections 5, 6, and 7 are available. It is clear that the transformation U_1, that we have obtained in this section, is the same as that introduced in those sections.

By means of (10.7) some strong results can be obtained by appealing to theorem 2.3.

<u>LEMMA 10.1.</u> $\lambda = 1$ is the only eigen-value of modulus 1 of U with respect to $C_L(\Omega)$. Also

(10.8) $\quad U^n\phi_i(t) \to U_1\phi_i(t) = \Psi_i$ (constant)

the convergence being uniform geometric with respect to t and i.

<u>PROOF</u>: The first part follows from theorem 2.3. Equation (2.21) gives

$$||U^n x - U_1 x||_L \leq \frac{H}{(1+h)^n} ||x||_L \qquad H,h \text{ constants.}$$

In particular

$$||U^n\phi_i - U_1\phi_i|| \leq \frac{H}{(1+h)^n} ||\phi_i||_L$$

$$\leq \frac{H'}{(1+h)^n} \qquad H',h \text{ constants(by (4.9)).}$$

By theorem 10.1, $U_1\phi_i = \Psi_i$ is a constant. Hence (10.8) is established and the lemma is proved.

As has been pointed out, Ψ_i which is a constant represents the limiting probability of applying the transformation A_i at n^{th} step as $n \to \infty$, given that the initial state is t. Then since Ψ_i is a constant, this limiting probability is independent of the initial state. The above lemma was established by Ocinescu and Mihoc for a finite number of transformations where there exists one k such that $\phi_k(t) > 0$ for all t.

THEOREM 10.3. For any $\mu \in \mathfrak{M}(\Omega)$

(10.9) $\quad T^n\mu \longrightarrow T_1\mu \qquad$ (weak-star)

where T_1 is the adjoint of U_1. Moreover if μ is a probability measure, then

(10.10) $\quad T_1\mu = \nu$

where ν is a probability measure independent of μ. It is the unique probability measure which is a fixed point of T.

PROOF: (10.9) is merely a translation into the conjugate space of the result of theorem 10.1. The proof is exactly the same as that of theorem 7.1. The proof of the second part of the theorem is the same as that of theorem 9.2.

It would be desirable to have some information concerning the limiting measure. The following result holds for constant non-zero probabilities $\phi_i(t)$.

THEOREM 10.4. Let $\phi_i(t) = \phi_i$ be a non-zero constant for each i. Then the limiting probability measure ν (10.10) is continuous, i.e., the measure of each point in Ω is zero. In fact $T_1\mu$ is continuous for any measure $\mu \varepsilon \mathfrak{M}(\Omega)$.

<u>PROOF:</u> Let $\nu(t_o) > 0$ where t_o is a point in Ω. Choose $y_k \neq t_o$. There exists an n such that $A_k^n\Omega$ does not include t_o. Now since $\nu = T\nu$ we have

$$\nu(t_o) = \Sigma\, \phi_{j_1} \ldots \phi_{j_n}\ \nu\,(A_{j_n}^{-1} \ldots A_{j_1}^{-1}t_o).$$

Each coefficient on the right-hand side is greater than zero and the sum of the coefficients is 1. Moreover $\nu(A_k^{-n}t_o) = 0$ since $A_k^{-n}t_o = \phi$. Hence

$$\nu(t_o) < \max_{j_1,\ldots,j_n} \nu\,(A_{j_n}^{-1} \ldots A_{j_1}^{-1}t_o).$$

Hence $\nu(t_o) < \nu\,(A_{j_n}^{-1} \ldots A_{j_1}^{-1}t_o)$ for some $j_1, \ldots, j_n$. Apply the same argument to $t_1 = A_{j_n}^{-1} \ldots A_{j_1}^{-1}t_o$; we obtain a sequence of points $t_o, t_1, t_2, \ldots$ such that

$$\nu(t_o) < \nu(t_1) < \nu(t_2) < \nu(t_3) < \ldots .$$

Since there is strict inequality no two of the points coincide. Hence there is a countable set each of which has measure $> \nu(t_o) > 0$. Hence the measure is infinite. This being a contradiction we have $\nu(t_o) = 0$. Thus since the same proof may be applied to $T_1\mu$ for any $\mu \varepsilon \mathfrak{M}(\Omega)$, the theorem is established.

We conclude this section with the following theorem.

<u>THEOREM 10.5.</u> If $UA_i\Omega = \Omega$, and if $\phi_i(t) > 0$ all i, all $t \varepsilon \Omega$ then there exists an n such that U^n is a strictly positive transformation, i.e.,

(10.11) if $x(t) \geq 0$ and $\not\equiv 0$ then $U^n x(t) > 0$ $t \varepsilon \Omega$.

<u>PROOF:</u> By hypothesis F (defined (3.4)) is the full space Ω. Let t_o be such that $x(t_o) > 0$. By lemma 3.2 there exists a sequence of integers $i_1, i_2, \ldots, i_n, \ldots$ such that

$$A_{i_1}\Omega \supset A_{i_1}A_{i_2}\Omega \ldots \ldots A_{i_1}A_{i_2} \ldots A_{i_n}\Omega \quad \ldots \quad \to t_o.$$

Since $x(t_o) > 0$ there exists a neighborhood U of t_o for which

$$x(t) > 0 \qquad t \varepsilon U.$$

Choose n so that $A_{i_1} \ldots A_{i_n}\Omega \ \varepsilon\ U$.

$$U^n x(t) = \sum \phi_{j_1}(t)\phi_{j_2}(A_{j_1}t) \ldots \phi_{j_n}(A_{j_{n-1}} \ldots A_{j_1}t)\, x(A_{j_n} \ldots A_{j_1}t).$$

Since $A_{i_1} \ldots A_{i_n} t \ \varepsilon\ U$, and the coefficients are > 0, it follows from the last equation that

$$U^n x(t) > 0,$$

and the theorem is proved.

11. A Special problem involving attractions towards the vertices of a simplex.

We now introduce a special model which is a generalization from 1 to N dimensions of the learning model considered by Karlin.

Let Ω be a simplex in E_N (Euclidean space of N dimensions). Let $V_1, V_2, \ldots, V_{N+1}$ be the vertices of the simplex. Any point of Ω is given by its barycentric coordinates

$$t = (t_1, t_2, \ldots t_{N+1})$$

where

$$t_i \geq 0, \qquad \sum_{1}^{N+1} t_i = 1.$$

Let B_i denote the $(N+1) \times (N+1)$ projection matrix

$$\begin{pmatrix} 0 & 0 & 0 & & 0 \\ 0 & 0 & 0 & & 0 \\ & & & & \\ 1 & 1 & 1 & & 1 \\ & & & & \\ 0 & 0 & 0 & & 0 \\ 0 & 0 & 0 & & 0 \end{pmatrix} \qquad i^{th} \text{ row.}$$

Observe that $B_i t = y_i$. Let I denote the identity $(N+1) \times (N+1)$ matrix. We define our transformations on the simplex Ω as follows

(11.1) $$A_i = \lambda_i I + (1 - \lambda_i) B_i \qquad 0 < \lambda_i < 1 \quad i = 1, \ldots, N+1$$

Clearly A_i represents a transformation which carries a point R into a point R' on the line RV_i where

(11.2) $$(R'V_i) = \lambda_i (RV_i),$$

i.e, A_i represents an attraction towards the vertices V_i.

We have by (11.1) that if $t = (t_1, t_2, \dots, t_{n+1})$ is any point of Ω, the coordinates of the transformed point are given by

$$(11.3) \qquad A_i t = (\lambda_i t_1, \lambda_i t_2, \dots, \lambda_i t_i + 1 - \lambda_i, \lambda_i t_{i+1}, \dots, \lambda_i t_{N+1})$$

It is obvious that

$$(11.4) \qquad A_i \Omega = \{ t \mid t_i \geq 1 - \lambda_i \}$$

A_i^{-1} has as domain $A_i \Omega$ and is defined by

$$A_i^{-1} = \lambda_i^{-1} I + (1 - \lambda_i^{-1}) B_i.$$

It is clear that this set of transformations satisfy the conditions imposed in section 3. In particular

$$\rho(A_i t, A_i s) \leq \lambda \rho(t,s) \qquad \text{all } i,t,$$

where $\lambda = \max \lambda_i$.

LEMMA 11.1. $\bigcup_1^{N+1} A_i \Omega = \Omega$ if and only if $\sum_1^{N+1} \lambda_i \geq N$.

PROOF: $A_i \Omega = \{ t \mid t_i \geq 1 - \lambda_i \}$

$$(A_i \Omega)^c = \{ t \mid t_i < 1 - \lambda_i \}$$

$t \in \bigcap_i (A_i \Omega)^c$ if and only if $0 \leq t_i < 1 - \lambda_i$, $i = 1, \dots, N+1$.

In order to obtain a point t satisfying these conditions, we must select t_i $(i = 1, \dots, N+1)$ to satisfy the above inequalities and in addition $\sum t_i = 1$. This can be done if and only if

$$\sum_1^{N+1} (1 - \lambda_i) > 1$$

or

$$\sum_1^{N} \lambda_i < N.$$

Thus $\bigcap_1^{N+1} (A_i \Omega)^c$ is non-void if and only if $\sum_1^{N+1} \lambda_i < N$. By taking complements the result follows.

Let $|S|$ denote the Lebesgue measure of any set.

LEMMA 11.2. If $\sum_1^{N+1} \lambda_i^N < 1$ then $|F| = 0$ where F is defined by (3.4).

PROOF: $|A_i \Omega| = \lambda_i^N |\Omega| = \lambda_i^N$ taking $|\Omega| = 1$

$$\therefore \quad |\bigcup_i A_i \Omega| \leq \sum_1^{N+1} \lambda_i^N$$

$$|A_j \bigcup_i A_i \Omega| \leq \lambda_j^N \sum_1^{N+1} \lambda_i^N$$

$$|\bigcup_{i,j} A_i A_j \Omega| \leq \left(\sum_1^{N+1} \lambda_i^N\right)^2$$

and proceeding in this manner we obtain

$$|A^{[n]} \Omega| \leq \left(\sum_1^{N+1} \lambda_i^N\right)^n.$$

using the notation introduced in section 3. Hence $|\bigcap A^{[n]} \Omega| = 0$, i.e., $|F| = 0$, and the lemma is proved.

Since the model under consideration is a special case of those considered in previous sections, all the limiting theorems proved hold for this model. However we may obtain some properties concerning the convergence of derivatives of $U^n x$ for this model.

We begin by proving a theorem which for this model gives more information than lemma 4.8 (i.e., $||U^n||_L \leq C$), when we have differentiable probabilities $\phi_i(t)$.

<u>Consider</u> $t_1, t_2, \ldots, t_N$ <u>as independent variables</u>,

$$t_{N+1} = 1 - t_1 - t_2 - \cdots - t_N.$$

We write

$$x(t) = x(t_1, t_2, \ldots, t_N).$$

Let $C^m(\Omega)$ be the set of m-times continuously differentiable functions.

<u>THEOREM 11.1.</u> Let $\phi_i(t) \varepsilon C^m(\Omega)$, If $x(t) \varepsilon C^m(\Omega)$, then there exist constants K_ℓ depending on $x(t)$ ($\ell = 1, \ldots, m$) such that

$$\left| \frac{\partial^\ell U^n x(t)}{\partial t_1^{i_1} \cdots \partial t_N^{i_N}} \right| \leq K_\ell \qquad n = 1,2, \ldots$$

for all $i_1, i_2, \ldots, i_N$ such that $i_1 + i_2 + \cdots + i_N = \ell \leq m$, i.e., all first m derivatives are uniformly bounded, for each $x(t) \varepsilon C^m(\Omega)$.

<u>PROOF:</u> If $||x|| \leq K_0$ then $||U^n x|| \leq K_0$ so that the theorem is true for $\ell = 0$. Suppose the theorem is true for $\ell - 1$.

$$Ux(t) = \sum_{i=1}^{N} \phi_i(t)\, x(\lambda_i t_1, \ldots, \lambda_i t_i + 1 - \lambda_i, \ldots, \lambda_i t_N)$$

$$+ \phi_{N+1}(t)\, x(\lambda_{N+1} t_1, \ldots, \lambda_{N+1} t_N)$$

$$\left|\frac{\partial^{\ell} Ux(t)}{\partial t_1^{i_1} \cdots \partial t_N^{i_N}}\right| \leq \sum_{i=1}^{N+1} \phi_i(t)\, \lambda_i^{\ell} \max_{\substack{t \\ i_1,\ldots,i_N}} \left|\frac{\partial^{\ell} x(t)}{\partial t_1^{i_1} \partial t_2^{i_2} \cdots \partial t_N^{i_N}}\right|$$

$$+ C \max_{\substack{t \\ j \leq \ell-1}} \left|\frac{\partial^{j} x(t)}{\partial t_1^{i_1'} \cdots \partial t_N^{i_N'}}\right|$$

where C is a constant bounding sums of absolute values of derivatives of $\phi_i(t)$ with binomial coefficients. Observe that

$$\sum_{i=1}^{N+1} \phi_i(t)\ \lambda_i^{\ell} \leq \max_i \lambda_i^{\ell} = p(\text{say})$$

where $p < 1$.

$$\max_{t,i_j} \left|\frac{\partial^{\ell} U^n x(t)}{\Pi \partial t_j^{i_j}}\right| \leq p \max_{t,i_j} \left|\frac{\partial^{\ell} U^{n-1} x(t)}{\Pi\, \partial t_j^{i_j}}\right| + C \max_{j \leq \ell-1} \left|\frac{\partial^{j}\ U^{n-1} x(t)}{\Pi \partial t_r^{i_r}}\right|$$

$$\leq p \max_{t,i_j} \left|\frac{\partial^{\ell} U^{n-1} x(t)}{\Pi \partial t_j^{i_j}}\right| + C' \quad \text{(by induction hypothesis)}$$

$$\leq p^2 \max_{t,i_j} \left|\frac{\partial^{\ell} U^{n-2} x(t)}{\Pi \partial t_j^{i_j}}\right| + pC' + C'$$

$$\leq p^n \max_{t,i_j} \left|\frac{\partial^{\ell} x(t)}{\Pi \partial t_j^{i_j}}\right| + \frac{C'}{1-p}$$

$$\leq \text{constant.}$$

Hence the theorem is established by induction.

Corollary: Under the conditions of the theorem, the sequence of functions

$$W_n(t) = \frac{\partial^r}{\prod_1^N \partial t_j^{i_j}} [U^n x(t)]$$

form an equicontinuous system of functions on Ω for each set of indices $i_1, \dots, i_N$ such that $\sum_1^N i_j = r \leq m - 1$.

We use the notation $D^r x(t)$ to stand for any one of the r^{th} derivatives of $x(t)$.

THEOREM 11.2. Let $x(t) \varepsilon C^m(\Omega)$. If $U^n x(t)$ converges uniformly to $U_1 x(t)$ then

$$D^r \left(U^n x(t)\right) \to D^r \left(U_1 x(t)\right) \qquad 0 \leq r \leq m - 1 \tag{11.5}$$

the convergence being uniform on Ω .

PROOF: We prove the theorem by induction. By hypothesis the theorem is true for $r = 0$. Suppose the theorem is true for $r - 1$. By the corollary to theorem 11.1

$$W_k(t) = D^r(U^k x(t))$$

form an equicontinuous family of functions on Ω . Moreover, by theorem 11.1 they are uniformly bounded. Hence by Arzela's theorem there exists a subsequence $W_{k_i}(t)$ such that

$$\lim_{i \to \infty} W_{k_i}(t) = \Psi(t), \tag{11.6}$$

the convergence being uniform on Ω. Suppose

$$D^r = \frac{\partial^r}{\prod_{j=1}^{N} \partial t_j^{i_j}}$$

At least one of the indices $i_j \neq 0$. Let $i_1 \neq 0$. Let

$$D^{r-1} = \frac{\partial^{r-1}}{\left(\prod_{j=2}^{N} \partial t_j^{i_j}\right) \partial t_1^{i_1-1}} .$$

By the induction hypothesis

$$\lim_{k \to \infty} D^{r-1} \left(U^k x(t)\right) = D^{r-1} \left(U_1 x(t)\right)$$

the convergence being uniform. In particular

$$\lim_{i \to \infty} D^{r-1} \left(U^{k_i} x(t)\right) = D^{r-1} \left(U_1 x(t)\right)$$

but by (11.6) this sequence differentiated with respect to t_1 converges uniformly and hence

$$\Psi(t) = D^r \left(U_1 x(t)\right) .$$

Hence $\Psi(t)$ is independent of the subsequence, and the theorem is proved.

This theorem means that if $\phi_i(t) \in C^m(\Omega)$, and satisfy either the conditions of section 9 or those of section 10, and if $x(t)$ is any function in $C^m(\Omega)$ then the first (m-1) derivatives converge uniformly over Ω.

LEMMA 11.3. Every point t in the set F may be represented in the form

$$(11.7) \qquad t = (1-\lambda_{i_1})\, B_{i_1} + \lambda_{i_1}(1-\lambda_{i_2})\, B_{i_2}$$

$$+ \cdots + \lambda_{i_1} \cdots \lambda_{i_{n-1}}(1-\lambda_{i_n})\, B_{i_n} + \cdots$$

where $i_1, i_2, \ldots, i_n, \ldots$ is a sequence of integers ranging from 1 to N+1. This representation is not necessarily unique.

<u>PROOF</u>: By lemma 3.2 there exists a sequence $i_1, i_2, \ldots$ such that

$$(11.8) \quad A_{i_1}\Omega \supset A_{i_1}A_{i_2}\Omega \supset A_{i_1}A_{i_2}A_{i_3}\Omega \supset \ldots. \quad \to t$$

$$\therefore \quad \lim t_n = t,$$

where

$$t_n = (\lambda_{i_1} I + (1-\lambda_{i_1})B_{i_1})(\lambda_{i_2} I + (1-\lambda_{i_2})B_{i_2}) \cdots (\lambda_{i_n} I + (1-\lambda_{i_n})B_{i_n})s$$

and s is an arbitrary point in Ω. Now $B_k B_j = B_k$. Thus we have

$$t_n = \lambda_{i_1}\lambda_{i_2}\cdots\lambda_{i_n} I + (1-\lambda_{i_1})B_{i_1} + \lambda_{i_1}(1-\lambda_{i_2})B_{i_2} + \cdots$$

$$+ \lambda_{i_1} \cdots \lambda_{i_{n-1}}(1-\lambda_{i_n})B_{i_n}.$$

Letting $n \to \infty$, the result (11.7) follows.

12. Properties of the One-dimensional Learning Model.

In this section we specialize the model of the last section to the case where $N = 1$. This is the model treated by Karlin [2]. Some additional properties are obtained in this section.

Here Ω is the unit interval. Write t for t_1 so that $t_2 = 1 - t$. We change the notation of the last section in that we replace the index 2 by the index 0. Thus our attractions consist of A_0 towards the origin and A_1 towards the point 1 where

(12.1) $$A_0 t = \lambda_0 t; \qquad A_1 t = \lambda_1 t + 1 - \lambda_1.$$

By lemma 3.2, the set F is closed. It is shown by Karlin [2, page 753] that if $\lambda_1 + \lambda_2 < 1$, then F consists of a Cantor-like set. If $\lambda_1 + \lambda_2 = 1$, then it is clear that $F = \Omega$.

Any sequence $\alpha = (a_1, a_2, \dots)$ where a_k ranges over the pair of numbers 0,1 gives rise to a point $t(\alpha)$ in Ω, according to lemma 11.3, by the equation

(12.2) $$t(\alpha) = (1 - \lambda_1) \sum_{n=1}^{\infty} a_n \lambda_1^{\sum_1^{n-1} a_i} \lambda_0^{n-1-\sum_1^{n-1} a_i}.$$

If $\lambda_0 + \lambda_1 < 1$, then each point in F is described by a unique sequence α. This follows from the fact that no point can be in the range of both A_0 and A_1.

If $\lambda_0 + \lambda_1 = 1$, there is exactly one point in the range of both A_0 and A_1, namely, the point $\lambda_0 (= 1 - \lambda_1)$. At most two sequences α_1, α_2 give rise to the same point, and such sequences considered as the coefficients of the dyadic expansion of a number between zero and one correspond to the same number. Again, this statement is easily verified.

Let

$$\alpha = (a_1, a_2, \ldots)$$
$$\alpha' = (a_1', a_2', \ldots)$$

Define $\alpha < \alpha'$ if $a_k < a_k'$ for the first inequality of corresponding elements (i.e., $a_1 = a_1', \ldots, a_{k-1} = a_{k-1}'$). With this convention we have the following lemma:

<u>LEMMA 12.1.</u> If $\lambda_o + \lambda_1 \leq 1$, then $\alpha \leq \alpha'$ implies that $t(\alpha) \leq t(\alpha')$

<u>PROOF</u>: Let $a_1 = a_1', a_2 = a_2', \ldots, a_{k-1} = a_{k-1}', \; a_k = 0, a_k' = 1.$

$$t(\alpha') - t(\alpha) \geq (1-\lambda_1)\, \lambda_1^{\sum_1^{k-1} a_i} \lambda_o^{k-1-\sum_1^{k-1} a_i} \Big[1 - \sum_{k+1}^{\infty} \lambda_1^{\sum_k^{n-1} a_i} \lambda_o^{n-k-\sum_k^{n-1} a_i}\Big]$$

$$\geq (1-\lambda_1)\, \lambda_1^{\sum_1^{k-1} a_i} \lambda_o^{k-1-\sum_1^{k-1} a_i} \Big[1 - \sum_{k+1}^{\infty} \lambda_1^{n-k-1} \lambda_o\Big]$$

since $a_k = 0$

$$\geq (1-\lambda_1)\, \lambda_1^{\sum_1^{k-1} a_i} \lambda_o^{k-1-\sum_1^{k-1} a_i} \Big[1 - \frac{\lambda_o}{1-\lambda_1}\Big]$$

$$\geq 0 \qquad \text{if and only if} \quad \lambda_o + \lambda_1 \leq 1.$$

We now define transformations A_o, A_1 on the sequences $\alpha = (a_1, a_2, \ldots)$ by

(12.3) $\quad A_o\alpha = (0, a_1, a_2, \ldots), \qquad A_1\alpha = (1, a_1, a_2, \ldots)$

It is clear that

(12.4) $\quad t(A_o\alpha) = A_o t(\alpha), \qquad t(A_1\alpha) = A_1 t(\alpha)$

THEOREM 12.1. Let $\lambda_o + \lambda_1 \leq 1$. If $\phi_o(t), \phi_1(t)$ are constants ϕ_o, ϕ_1 respectively, and if $F(t)$ is the cumulative distribution corresponding to the limiting measure $V(E)$, defined in section 10, then t, $F(t)$ can be expressed in the following parametric form

(12.5)
$$\begin{cases} t(\alpha) = (1-\lambda_1) \sum_{n=1}^{\infty} a_n \lambda_1^{\sum_1^{n-1} a_i} \lambda_o^{n-1-\sum_o^{n-1} a_i} \\ G(\alpha) = (1-\phi_1) \sum_{n=1}^{\infty} a_n \phi_1^{\sum_1^{n-1} a_i} \phi_o^{n-1-\sum_1^{n-1} a_i} \\ G(t) = G(\alpha(t)) \qquad t \in F. \end{cases}$$

PROOF: It is only necessary to define G on F for $V(F^c) = 0$ by lemma 7.1. It is clear from lemma 12.1 that the first two formulas in (12.5) are monotonic functions of α, and thus $G(t)$ is well-defined and is monotonic on F. (If $\lambda_o + \lambda_1 = 1$, then two α's give rise to the same $t(\alpha)$ but these two α's give the same $G(\alpha)$, so that there is no ambiguity.) $G(0) = 0$, $G(1) = 1$. Hence $G(t)$ represents a distribution on F. Clearly it is continuous on F. By extending it in an obvious way we have a distribution on Ω.

Our object is to obtain the unique fixed point of T, i.e., ν such that

(12.6) $\nu(E) = \phi_0 \, \nu(A_0^{-1}E) + \phi_1 \nu(A_1^{-1}E).$

If E is in F^c, so is $A_0^{-1}E$, $A_1^{-1}E$. Thus, since $\nu(F^c) = 0$, it is sufficient to obtain a continuous probability distribution H(t) (see theorem 10.4) on F such that (12.6) is satisfied.

If $t \in A_0\Omega$, but not in $A_1\Omega$, then (12.6) gives

(12.7) $$H(t) = \phi_0 H(A_0^{-1}t).$$

If $t \in A_1\Omega$, but not in $A_0\Omega$, then (12.6) gives

(12.8) $$H(t) = \phi_1 H(A_1^{-1}t) + \phi_0.$$

Now $t \in A_0\Omega \cap A_1\Omega$ if and only if $\lambda_0 + \lambda_1 = 1$ and $t = \lambda_0$. For this point, $H(\lambda_0) = \phi_0$, which is consistent with both the above equations. Thus the continuous distribution H(t) which satisfies

(12.9) $$H(A_0 t) = \phi_0 H(t) \qquad \text{(from (12.7))}$$

(12.10) $$H(A_1 t) = \phi_1 H(t) + 1 - \phi_1 \qquad \text{(from (12.8))}$$

is the required distribution.

Now we may obtain for $G(\alpha)$ equations similar to (12.4), so that we find G(t) satisfies (12.9) and (12.10) for all $t \in F$ and therefore all $t \in F$, and thus is the required distribution. Hence the theorem is proved.

An obvious question to be asked in this type of problem is whether or not the limiting distribution is absolutely continuous or singular. The following theorem provides an answer under very special conditions.

THEOREM 12.2. Let $\lambda_o = \lambda_1 = 1/2$. Under the conditions of section 10, the fixed points of T are singular except for the case $\phi_o(t) = \phi_1(t) = 1/2$. In this case the fixed point is absolutely continuous, and the unique fixed point which is a probability distribution is given by $F(t) = t$.

PROOF: $\nu(E) = T\nu(E) = \int_{A_1^{-1}E} \phi_1(t)\, d\nu(t) + \int_{A_o^{-1}E} \phi_o(t)\, d\nu(t).$

Let $F(t)$ be the cumulative distribution corresponding to ν. Consider first $\lambda_1 + \lambda_o \geq 1$. The above equation yields the following relations for $F(t)$:

$$0 \leq \tau < 1-\lambda_1; \qquad F(\tau) = \int_0^{\tau/\lambda_o} \phi_2(t)\, dF(t)$$

$$1-\lambda_1 \leq \tau \leq \lambda_o; \qquad F(\tau) = \int_0^{\frac{\tau-(1-\lambda_1)}{\lambda_1}} \phi_1(t)\, dF(t) + \int_0^{\tau/\lambda_o} \phi_o(t) dF(t)$$

$$\tau > \lambda_o; \qquad F(\tau) = \int_0^{\frac{\tau-(1-\lambda_1)}{\lambda_1}} \phi_1(t)\, dF(t) + \int_0^1 \phi_o(t) dF(t).$$

Suppose $F(\tau) = \int_0^\tau f(t)\,dt.$

$$\therefore \quad F'(\tau) = f(\tau) \qquad \text{a.e.}$$

Then the above relations give

$$0 \le \tau < 1-\lambda_1; \qquad f(\tau) = \frac{1}{\lambda_0}\,\phi_0(\tau/\lambda_0)\, f(\tau/\lambda_0) \quad \text{a.e.}$$

$$1-\lambda_1 \le \tau \le \lambda_0; \qquad f(\tau) = \frac{1}{\lambda_1}\,\phi_1\left(\frac{\tau-(1-\lambda_1)}{\lambda_1}\right) f\left(\frac{\tau-(1-\lambda_1)}{\lambda_1}\right)$$

(12.11)

$$+ \frac{1}{\lambda_0}\,\phi_0(\tau/\lambda_0)\, f(\tau/\lambda_0) \quad \text{a.e.}$$

$$\lambda_0 < \tau \le 1; \qquad f(\tau) = \frac{1}{\lambda_1}\phi_1\left(\frac{\tau-(1-\lambda_1)}{\lambda_1}\right) f\left(\frac{\tau-(1-\lambda_1)}{\lambda_1}\right)$$

a.e.

Now assume the given hypothesis $\lambda_1 = \lambda_2 = 1/2$. Consider the Fourier coefficients of f

$$\phi(n) = \int_0^1 f(t)\, e^{2\pi int}\,dt.$$

Using the above relations we obtain

$$\phi(n) = \int_0^1 \phi_2(x)\, f(x)\, e^{\pi inx}\,dx + \int_0^1 \phi_1(x)\, f(x)\, e^{\pi inx}\, e^{\pi in}\,dx.$$

$$\therefore \quad \phi(2n) = \int_0^1 f(x)\, e^{2\pi inx}\,dx$$

$$= \phi(n).$$

Thus

$$\phi(n) = \phi(2n) = \dots\dots\dots = 0 \quad \text{by the Riemann-Lebesgue lemma.}$$

Hence $\phi(n) = 0$, all n. Therefore, $f(x) = \text{const.}$

Since F is a probability distribution, it follows that $f(x) = 1$. From (12.11) we have

$$0 \le \tau < 1/2 \qquad f(\tau) = 2\phi_o(2\tau)\, f(2\tau)$$

$$\therefore \qquad 1 = 2\phi_o(2\tau).$$

Hence

$$\phi_o(t) = 1/2 \qquad \text{all } t.$$

$$\therefore \quad \phi_1(t) = 1/2.$$

The transformation T preserves absolute continuity, and singularity. Since any measure $\nu = \nu_1 + \nu_2$ where ν_1 is absolutely continuous and ν_2 is singular, the decomposition being unique, we have from the uniqueness of the fixed point ν of T (under the conditions of section 10) that the fixed point is either absolutely continuous or singular. Hence the theorem is proved.

Finally we note that if $\phi_1(t) = \lambda_1$, $\phi_2(t) = \lambda_2$ then the unique fixed point of T is given by $F(t) = t$. This may be verified directly or it is easily seen from 12.5.

REFERENCES

[1] R.R. Bush and C.F. Mosteller, A mathematical model for simple learning, Psych. Rev. 58 (1951), 313-323.

[2] S. Karlin, Some random walks arising in learning models 1, Pacific Journal of Mathematics 3 (1953), 725-756.

[3] O. Onicescu et G. Mihoc, Sur les chaînes des variables statistiques, Bull. Sci. Math. 59 (1935), 174-192.

[4] W. Doeblin et R. Fortet, Sur des chaînes à liaisons complètes, Bull. Soc. Math. France 65 (1937), 132-148.

[5] R. Fortet, Sur des operations lineàires non-complètement continues, Comptes Rendus, Acad. Sci. Paris 204 (1937), 1543-5.

[6] R. Fortet, Sur l'iteration des substitutions algebriques lineaires a une infinite de variables et ses applications a la theorie des probabilités en chaîne, Revista de Ciencias 40 (1938), 185-262, 337-448, 481-536.

[7] R. Fortet, Sur une suite également repartie, Stud. Math. 9 (1940), 54-70.

[8] C.T. Ionescu Tulcea et G. Marinescu, Sur certaines chaînes à liaisons complètes, Comptes Rendus, Acad. Sci. Paris 227 (1948), 667-669.

[9] C.T. Ionescu Tulcea et G. Marinescu, Theorie ergodique pour des classes d'operations non-completement continues, Annals Of Mathematics 52 (1950), 140-147.

[10] K. Yosida and S. Kakutani, Operator-theoretical treatment of Markoff's Process and mean ergodic theorem, Annals of Mathematics 42, (1941), 188-228.

[11] F. Hausdorff, Mengenlehre, third revised edition, Dover, New York.

[12] S. Kakutani, Concrete Representation of Abstract (M)-Spaces, Annals of Mathematics 42 (1941), 994-1023.