Invariant Distributions and the Limiting Behavior of Markovian Economic Models
Author(s): Carl A. Futia
Source: *Econometrica,* Vol. 50, No. 2 (Mar., 1982), pp. 377-408
Published by: The Econometric Society
Stable URL: http://www.jstor.org/stable/1912634
Accessed: 09/12/2010 05:13

INVARIANT DISTRIBUTIONS AND THE LIMITING BEHAVIOR OF MARKOVIAN ECONOMIC MODELS

By Carl A. Futia

Equilibria in stochastic economic models are often time series which fluctuate in complex ways. But it is sometimes possible to summarize the long run, average characteristics of these fluctuations. For example, if the law of motion determined by economic interactions is Markovian and if the equilibrium time series converges in a specific probabilistic sense then the long run behavior is completely determined by an invariant probability distribution. This paper develops and unifies a number of results found in the probability literature which enable one to prove, under very general conditions, the existence of an invariant distribution and the convergence of the corresponding Markov process.

VIRTUALLY ALL OF ECONOMIC THEORY focuses upon the study of economic equilibrium. This concept has recently undergone several subtle elaborations. No longer must a system of markets in equilibrium be thought of as one at rest in a static steady state. Instead there is a growing body of literature (e.g., [4, 5, 12, 16, 20, 21]) which defines equilibrium as a stochastic process of market clearing prices and quantities which is consistent with the self-interested behavior of economic agents.

Needless to say equilibrium stochastic processes can be very complex time series which fluctuate in irregular ways. For theoretical and econometric purposes it is useful to have a convenient way of summarizing the "average" behavior of such processes over time. This paper draws together and unifies a number of fundamental results from the probability literature which enable one to do this for discrete time, Markov processes on general state spaces.

The starting point of the analysis is a set S of economic states (e.g., prices and/or quantities). The only technical restriction placed upon S is that it be a Borel subset of a complete, separable metric space. The second datum is a transition probability $P(s, \cdot)$ on S. The number $P(s, A)$ records the probability that the economic system moves from the state s to some state in the Borel subset A of S during one unit of elapsed time. In economic applications the transition probability is usually derived from hypotheses about market clearing and the maximizing behavior of economic agents.

The transition probability (together with an initial probability measure on S) defines a discrete time Markov process. One way of summarizing the dynamic behavior implied by P is to look for an invariant probability. A probability measure λ on S is invariant for P if for all Borel subsets A of S one has the equality $\int P(s, A)\lambda(ds) = \lambda(A)$. An invariant probability is a kind of probabilistic steady state for the dynamics defined by P. Of course there may be no invariant probability for P at all; and even if one exists it may convey no information about the average behavior of the process over time except under very special initial conditions.

There is a second way of summarizing the behavior of Markov processes defined by the transition probability P. Let $P^n(s, A)$ denote the n step transition

probability derived from P. If the sequence of probability measures $(P^n(s, \cdot)$, $n = 1, 2, \ldots)$ is convergent in the weak topology on probabilities, then two conclusions follow. First, the limit of the sequence is an invariant probability λ_s for P. Secondly, if the Markov process starts from the state s then its average behavior over time is determined completely by λ_s. Thus one would like to prove that the sequence $(P^n(s, \cdot))$ converges for each s; if possible, one would also like to deduce that the invariant probabilities λ_s do not depend on s so that P has only one invariant probability.

It is the thesis of this paper that for most Markovian stochastic models one can prove the existence of invariant distributions and the convergence of the process to them by appealing to some very general theorems. One usually verifies the hypotheses of the theorems by verifying certain properties of the transition probability.

The mathematical theory I shall discuss is well known (in various guises) to experts in stochastic processes. Some of it is also familiar to a handful of economic theorists. The role of this paper is, therefore, largely a pedagogical one. It presents a unified mathematical framework and a collection of techniques in a form useful for direct economic applications.

I would next like to outline this paper's contents and suggest how the main results might be applied to stochastic models which have appeared in the literature.

Section 1 establishes the unifying mathematical theme of this paper by defining the concept of a Markov operator. Such operators are linear transformations which are defined naturally by transition probabilities and which mirror all the essential features of the associated Markov process.

In Section 2 we establish conditions sufficient to imply the existence of an invariant probability for a Markov process. These conditions require that the process not wander too often from compact sets and that the associated Markov operator have the so-called Feller property of mapping continuous functions into continuous functions. These conditions easily lead to the main implication of Theorem 2 in [4] and also allow a quick proof of the existence of invariant distributions in the Grandmont-Hildenbrand temporary equilibrium model [12].

Section 2 also develops an easy-to-use uniqueness criterion for invariant distributions. The criterion can be simply stated in terms of the transition probability and allows one (for example) to easily prove the uniqueness of the stochastic steady state in the Brock-Majumdar growth model [4] for an open, dense set of exogenous shock distributions.

A powerful and frequently employed tool (see for example [16, 22]) for proving the convergence of Markov processes to invariant distributions is the famous condition of Doeblin. Sections 3 and 4 establish the equivalence between Doeblin's condition and the quasi-compactness of the associated Markov operator. More importantly, they present results which enable one to recognize quasi-compact operators by inspecting the associated transition probability.

In many applications (e.g., [5, 13, 22]) one must derive the transition probability from more basic economic hypotheses. This often makes it difficult to show that the intuitively correct definition of the transition probability has the neces-

sary measurability properties. Section 5 shows how these problems can be overcome for economic models which can be represented as "random dynamical systems." We also establish conditions on such systems sufficient to imply the quasi-compactness of the corresponding Markov operator.

Section 6 studies random contractions; these are random dynamical systems having additional contraction mapping properties. Random contractions arise naturally when studying the dynamics defined by optimal policies for concave, stochastic, dynamic programming problems [10]. They also arise from stochastic learning models [6, 25, 32] and from certain kinds of stochastic games [31].

One important topic has been omitted from this paper. In economic applications (e.g., [11, 23]) one often wishes to determine how the invariant distributions of a Markov process change when the transition probability is altered. Such comparative statics questions are difficult to answer in any generality. The only known results along these lines require strong restrictions upon the transition probability. The reader is referred to [7] for the basic results on this problem.

The Markov processes studied in this paper are all discrete time processes. The reader interested in studying continuous time Markov models should consult [2] for a discussion of diffusion processes applied to economic problems.

A final note: the reader is assumed to be familiar with the definitions of several standard mathematical terms used in this paper. For the convenience of those unfamiliar with these terms the Appendix contains references to appropriate definitions.

1. TRANSITION PROBABILITIES AND MARKOV OPERATORS

In this section we shall show how a transition probability that defines a Markov process also gives rise to a continuous, linear operator on a certain Banach space. From such an operator one can actually reconstruct the original transition probability. In fact, this procedure establishes a one to one correspondence between transition probabilities and Markov operators (m-operators). We shall see in later sections that this correspondence allows one to bring powerful Banach space techniques to bear upon the study of the asymptotic properties of Markov processes.

1.1 Transition Probabilities

To begin the discussion let $(S, \mathbf{S})$ be a measurable space (references to undefined mathematical terms can be found in the Appendix to this paper). The set S should be interpreted as the set of possible states of the economic system.

DEFINITION 1.1: A *transition probability* on $(S, \mathbf{S})$ is a function $P : S \times \mathbf{S} \to [0, 1]$ with two properties: (a) for each measurable set $A \in \mathbf{S}$ the real valued function $P(\cdot, A)$ is $\mathbf{S}$ measurable; (b) for each point s in S the set function $P(s, \cdot)$ with domain $\mathbf{S}$ is a probability measure on $(S, \mathbf{S})$. The number $P(s, A)$ should be interpreted as the probability that the economic system will move from the state s to some state in the set A during one unit of time.

Given a transition probability P on $(S, \mathbf{S})$, one can define, for each nonnegative integer, a transition probability P^i on $(S, \mathbf{S})$ by the formulae:

$$P^0(s, A) = \begin{cases} 1 & \text{if } s \in A, \\ 0 & \text{otherwise,} \end{cases}$$

$$P^i(s, A) = \int P^{i-1}(s, dt) P(t, A).$$

Clearly $P^i(s, A)$ can be interpreted as the probability of reaching a state in A from the state s after precisely i units of time elapse. Furthermore, $P^1 = P$.

It is well known that once an initial probability is specified, a transition probability can be used to define a Markov process whose random variables take their values in $(S, \mathbf{S})$ (see, for example, Breiman [3, p. 130] or Neveu [24, pp. 169–170]).

1.2 *Markov Operators*

Let $B(S)$ denote the set of all bounded, $\mathbf{S}$ measurable, real valued functions on S. This set is a Banach space under the sup norm $|f| \equiv \sup_{s \in S} |f(s)|$. A transition probability defines a continuous linear transformation T from $B(S)$ to itself via the formula

$$(Tf)(s) \equiv \int f(t) P(s, dt).$$

The measurability of Tf follows from two observations. First, f can be expressed as the difference of two nonnegative, measurable functions (Bartle [1, p. 12]). Secondly, every nonnegative function in $B(S)$ is the pointwise limit of functions taking on only finitely many values (Bartle [1, p. 13]).

The operator T defined above is called the *Markov operator* associated with P. T is a continuous mapping of operator norm equal to 1 (equivalently, $\sup_{|f| \le 1} |Tf| = 1$). Note that $Tf(s)$ is the mathematical expectation of the random variable f with respect to the probability measure $P(s, \cdot)$. Exactly the same formula can be used to define Markov operators T^i associated with the i-step transition probabilities P^i. Note that T^i is just T composed with T^{i-1}. Thus T^i is in fact the ith power of T.

The adjoint T^* of the operator T is defined by the formula

$$(T^*\lambda)(A) = \int P(t, A) \lambda(dt).$$

T^* maps the Banach space of bounded finitely additive set functions defined on $(S, \mathbf{S})$ to itself. This is the Banach space dual to $B(S)$ and is denoted by $ba(S)$; it has as its norm the total variation norm defined by

$$|\lambda| = \sup \sum_{i=1}^{n} |\lambda(A_i)|$$

where the supremum is taken over all finite partitions of S into disjoint subsets $\{A_i\}$. One can easily show, using the monotone convergence theorem, that $T^*\lambda$ is

countably additive whenever λ is. Thus T^* maps probabilities into probabilities. If $\lambda(A)$ is the probability that the economic system is in the set A at date τ, then $(T^*\lambda)(A)$ is the probability that it is in A at date $\tau + 1$.

2. EQUICONTINUOUS MARKOV OPERATORS

In this section we study a class of Markov operators with three very important properties. First of all, if T is such an operator there is always at least one probability λ such that $T^*\lambda = \lambda$. Such probabilities are called *invariant* and are the equilibria or steady states of the stochastic, dynamic process defined by T. Secondly, if we define the continuous linear operator $A_n(T)$ by the formula $A_n(T) \equiv (1/n)\sum_{i=0}^{n-1} T^i$ (for $n \geqq 1$) then, for any probability λ on $(S, \mathbf{S})$, the sequence of probabilities $\{A_n(T^*)\lambda\}$ converges in the weak topology to an invariant probability. (Note that $A_n(T)^* = A_n(T^*)$.) Thus the statistical, long run average behavior of the dynamic process is completely determined by the set of invariant probabilities. Finally, a simple criterion is available that allows one to determine whether the invariant probability is unique. This criterion is the direct generalization of that used when S is a finite set. In section three we will see that this uniqueness criterion can be generalized in such a way as to provide a method for proving the convergence of the sequence $\{(T^*)^i\lambda\}$ rather than just convergence of $\{A_n(T^*)\lambda\}$.

The reader should note that the results presented in this section are just minor variations (for the case of a noncompact state space S) of results obtained by Sine [33] and Jamison [17, 18].

2.1 *Convergence Properties of Equicontinuous Operators*

We begin the discussion by defining equicontinuous Markov operators.

ASSUMPTION: The state space S is a separable metric space. (Recall that a topological space is separable whenever it has a countable subset whose closure is the whole space.) The σ-field $\mathbf{S}$ is just the Borel σ-field of S. Finally, S is a Borel subset of a complete, separable metric space W, and its topology and metric are inherited from W.

This assumption will be maintained throughout this paper. For economic applications it seems quite unrestrictive. For example, any Borel subset of a finite dimensional Euclidean space or, more generally, of any separable Banach space has the required properties.

Define $C(S)$ to be the closed subspace of $B(S)$ consisting of the continuous functions.

DEFINITION 2.1: An m-operator T is said to be *stable* if Tf is continuous and bounded whenever f is. A stable m-operator is said to be *uniformly mean stable* (u.m.s) if the sequence of continuous functions $\{A_n(T)f\}$ is uniformly convergent for each f in $C(S)$.

The property of being a stable operator is often referred to in the probability literature as the Feller property. Please note that the term "stable" used in this sense does *not* necessarily connote any kind of dynamic stability as it often would in economic contexts.

DEFINITION 2.2: Let B be a set of probability measures on $(S, \mathbf{S})$. Then B is said to be ϵ-*tight* if there is a compact set $F_\epsilon \subseteq S$ such that $\lambda(F_\epsilon) \geqq 1 - \epsilon$ for all λ in B. The set B is said to be *tight* if it is ϵ-tight for all $\epsilon > 0$.

DEFINITION 2.3: An m-operator T is said to be *tight* if, for each $\epsilon > 0$, there is an integer n such that the family of probability measures $\{(T^*)^n \delta_s, s \in S\}$ is ϵ-tight.

In other words, an m-operator is tight provided that for each $\epsilon > 0$, there is an n and a compact set F_ϵ such that $P^n(s, F_\epsilon) \geqq 1 - \epsilon$ for all s (where P is the transition probability corresponding to the operator).

DEFINITION 2.4: An m-operator T is *equicontinuous* if it is uniformly mean stable and tight.

Now let $rca(S)$ denote the closed subspace of $ba(S)$ consisting of all regular, countably additive measures. In order to develop the properties of equicontinuous operators we introduce a topology on $rca(S)$ which is coarser than that defined by the total variation norm.

DEFINITION 2.5: The *weak topology* on $rca(S)$ is the coarsest topology making each linear functional in the set $\{\lambda \to \int f d\lambda, f \in C(S)\}$ continuous.

In other words, a sequence $\{\lambda_n\} \subseteq rca(S)$ converges weakly to an element λ_0 if and only if, for each f in $C(S)$, the sequence of real numbers $\{\int f d\lambda_n\}$ converges to $\int f d\lambda_0$.

Since S is a metric space every countably additive measure is regular (Parthasarathy [27, p. 27, Theorem 1.2]). Thus T^* maps $rca(S)$ to itself; that T^* is weakly continuous then follows easily from the fact that T is a stable operator.

For compact S it is well known that the set of probabilities on $(S, \mathbf{S})$ is a convex, weakly compact subset of $rca(S)$ (Parthasarathy [27, p. 45, Theorem 6.4]). The standard fixed point argument then shows that T^* has an invariant probability, i.e., that there exists a probability λ such that $T^*\lambda = \lambda$. But in some applications one needs tools which can be applied to the case where S is not compact. The following results show that the assumption of tightness can replace the compactness of S when proving the existence of invariant distributions.

LEMMA 2.6: *Suppose T is a tight m-operator. If λ is a probability, then the set of probabilities $\{(T^*)^i \lambda, i = 0, 1, \dots \}$ is tight.*

PROOF: Fix $\epsilon > 0$. There is by hypothesis an integer n and a compact set F_ϵ^n such that $P^n(s, F_\epsilon^n) \geqq 1 - \epsilon$ for all s. But then

$$P^{n+m}(s, F_\epsilon^n) = \int P^n(t, F_\epsilon^n) P^m(s, dt)$$

$$\geqq 1 - \epsilon.$$

Thus, for any probability λ,

$$(T^{*n+m}\lambda)(F_\epsilon^n) \equiv \int P^{n+m}(t, F_\epsilon^n)\lambda(dt)$$

$$\geqq (1 - \epsilon)\lambda(S) = 1 - \epsilon.$$

Therefore, the family $\{T^{*n+m}\lambda, m \geqq 0\}$ is ϵ-tight. Now by Parthasarathy [27, p. 29, Theorem 3.2] there is, for each i between zero and $n - 1$, a compact set F_ϵ^i such that $(T^{*i}\lambda)(F_\epsilon^i) \geqq 1 - \epsilon$. Define $F_\epsilon \equiv \bigcup_{i=0}^n F_\epsilon^i$. Then $(T^{*i}\lambda)(F_\epsilon) \geqq 1 - \epsilon$ for all $i \geqq 0$. Hence $\{T^{*i}\lambda\}$ is ϵ-tight. Since ϵ was chosen arbitrarily, the proof of the lemma is complete. $\hspace{2cm}$ Q.E.D.

COROLLARY 2.7: *If T is a tight m-operator, then for any probability λ, the family* $\{A_n(T^*)\lambda, n \geqq 1\}$ *is tight.*

Tight families of probabilities have an important property.

PROPOSITION 2.8: *Let B be a tight set of probabilities. Then the closure of B in the weak topology is compact* (cf. Parthasarathy [27, p. 47, Theorem 6.7]).

THEOREM 2.9: *Suppose T is a stable, tight m-operator. Then there is a probability λ such that $T^*\lambda = \lambda$.*

PROOF: Let μ be an arbitrary probability, and consider the sequence $\{A_n(T^*)\mu\}$. Applying Corollary 2.7 and Proposition 2.8, we see that this sequence has weakly compact closure in the set of all probabilities. Since S is separable metric the weak topology on the set of all probabilities on S is metrizable. Therefore, there must be a probability λ and a subsequence $\{A_{n_i}(T^*)\mu\} \equiv \alpha$ that converges weakly to λ. Since T^* is weakly continuous, the sequence $\{T^* \cdot A_{n_i}(T^*)\mu\} \equiv \beta$ converges weakly to $T^*\lambda$. The sequence $\beta - \alpha$ is just $\{(1/n)(T^{*(n_i+1)}\mu - \mu)\}$. But this converges weakly to zero. Since $rca(S)$ is a topological vector space in its weak topology, it follows that the sequences β and α have the same weak limit. Thus $T^*\lambda = \lambda$, completing the proof. $\hspace{1cm}$ Q.E.D.

Theorem 2.9 shows that stochastic, dynamic processes giving rise to stable, tight Markov operators have stochastic steady states, i.e., invariant probabilities. But this fact in itself is not very interesting unless one knows that these invariant probabilities accurately mirror the long run average behavior of the process. We

next show that the additional hypothesis of uniform mean stability implies the desired convergence result.

THEOREM 2.10: *Suppose T is an equicontinuous m-operator. Then, for any probability μ, the sequence $\{A_n(T^*)\mu\}$ converges weakly to only one invariant probability λ.*

PROOF: The proof of Theorem 2.8 shows that there is an invariant probability λ that is the weak limit of a subsequence $\{A_{n_i}(T^*)\mu\}$. In other words, for each bounded continuous f,

$$(f, A_{n_i}(T^*)\mu) \to (f, \lambda) \equiv \int f \, d\lambda.$$

But $(f, A_{n_i}(T^*)\mu) = (A_{n_i}(T)f, \mu)$. Since T is uniformly mean stable, both sequences of continuous functions $\{A_n(T)f\}$ and $\{A_{n_i}(T)f\}$ converge uniformly to the same limit. Hence

$$\lim_i (A_{n_i}(T)f, \mu) = \lim_n (A_n(T)f, \mu)$$

$$= \lim_n (f, A_n(T^*)\mu).$$

Thus $\{A_n(T^*)\mu\}$ converges weakly to λ. Since $rca(S)$ is Hausdorff in the weak topology, this is the only limit point of the sequence.   Q.E.D.

The reader should be careful to note that the limiting invariant probability λ appearing in Theorem 2.10 will generally depend upon the choice of the initial probability μ.

2.2 *A Uniqueness Criterion*

In economic problems it is often important to have a criterion that will imply the uniqueness of the invariant probability. We develop just such a criterion in this subsection. It is applicable to processes giving rise to equicontinuous *m*-operators.

One crucial hypothesis that we shall use time and again is the hypothesis of stability. In other words, the operator in question must preserve continuous functions (i.e., it must "respect the topology of S"). In Section 3 we shall present an example which shows that if the stability hypothesis is dropped, our uniqueness criterion is no longer valid; this is despite the fact that the operator in question exhibits very nice convergence properties.

The results in this subsection are those of Sine [33] modified by replacing the assumption of compact S by the assumption of tightness on T. For this reason we shall omit many technical details and refer the interested reader to Sine [33] for the necessary proofs.

UNIQUENESS CRITERION 2.11: There is a point s_0 in S with the following property. For any neighborhood U of s_0, and any point s in S, one can find an integer n such that $P^n(s, U) > 0$.

The reader should note that when S is a finite set in the discrete topology, this criterion directly generalizes the usual uniqueness criterion for finite Markov chains.

We wish to prove the following theorem.

THEOREM 2.12: *Let T be an equicontinuous m-operator with transition probability P. Then the following conditions are equivalent:* (a) *there is a unique probability λ satisfying $T^*\lambda = \lambda$;* (b) *P satisfies the uniqueness criterion 2.11.*

Theorem 2.12 has a very useful corollary.

COROLLARY 2.13: *Let T_1, T_2 be stable m-operators, and let a be a real number $0 < a \leqq 1$. Suppose that the m-operators $T = aT_1 + (1 - a)T_2$ and T_1 are both equicontinuous, and further that T_1 has a unique invariant probability. Then T has a unique invariant probability.*

PROOF: Let λ be the unique probability such that $T_1^*\lambda = \lambda_1$ and let P_1 be the transition probability for T_1. By Theorem 2.12, P_1 satisfies the uniqueness criterion with respect to some point s_0. Let a neighborhood U of s_0 be given, and let s be any point in S. We can then find an integer n such that $P_1^n(s, U) > 0$. But $P^n(s, U) \geqq a^n P_1(s, U) > 0$. Thus P also satisfies the uniqueness criterion; another application of 2.12 then gives the result.   *Q.E.D.*

The proof of Theorem 2.12 will take up the rest of this section. The reader uninterested in the technical details should move on to Section 3.

If T is an m-operator, define K_T to be the subset of $rca(S)$ consisting of the probabilities invariant under T. Our first task is to show that K_T is completely determined by its extreme points provided T is equicontinuous.

LEMMA 2.14: *Suppose T is equicontinuous. Then K_T is convex and weakly compact.*

PROOF: Because T^* is linear, K_T is obviously convex; K_T is weakly closed because T^* is weakly continuous. To show that K_T is weakly compact, it will, in view of Proposition 2.8 and the fact that K_T is weakly closed, suffice to show that K_T is a tight family of probabilities.

By hypothesis, T is a tight operator. Let P be the corresponding transition probability, and fix $\epsilon > 0$. We can then choose an n and a compact set F_ϵ such that $P^n(s, F_\epsilon) \geqq 1 - \epsilon$. Now suppose λ is in K_T. Since λ is invariant under T^*, we

conclude that

$$\lambda(F_\epsilon) = ((T^*)^n \lambda)(F_\epsilon) = \int P^n(s, F_\epsilon) \lambda(ds)$$

$$\geqq (1 - \epsilon)\lambda(S) = 1 - \epsilon.$$

Thus K_T is ϵ-tight for every ϵ. This completes the proof.   *Q.E.D.*

In order to state the principal implication of Lemma 2.14, we need a definition. Suppose A is a convex subset of a linear space.

DEFINITION 2.15: An element x in A is said to be an *extreme point* of A if the following condition holds. Whenever y, z are in A and $x = ty + (1 - t)z$ with $0 < t < 1$, then $x = y = z$.

COROLLARY 2.16: *If* T *is equicontinuous, then* K_T *is the closed* (*in the weak topology*) *convex hull of its extreme points* (cf. Robertson [**29**, p. 138]).

The idea behind the proof of Theorem 2.12 is to show that the uniqueness criterion is equivalent to the assertion that K_T has only one extreme point. Corollary 2.16 then implies that K_T must consist of a single point. In order to carry through this program, we must first characterize the extreme points of K_T. The remarkable property of equicontinuous operators is that such a characterization can be developed using only the supports of invariant probabilities.

DEFINITION 2.17: Let μ be a probability. The support of μ, denoted $\sigma(\mu)$, is the smallest, closed subset of S with the property $\mu(\sigma(\mu)) = 1$.

Since S has been assumed to be a separable metric space, Parthasarathy, [**27**, p. 27, Theorem 2.1] tells us that every probability in S has a unique support.

DEFINITION 2.18: Let F be a non-empty, closed subset of S. F is said to be *self-supporting* if $\sigma(T^*\delta_s) \subseteq F$ for every point s in F (recall that δ_s is the unit point mass at s).

Thus, a closed set is self-supporting provided that no probability can ever escape from it.

PROPOSITION 2.19: *Let* λ *be in* K_T. *Then* $\sigma(\lambda)$ *is a self-supporting set* (cf. Sine [**33**, Theorem 1.3]).

Since the support of every invariant probability is a self-supporting set, one might hope to show that the extreme, invariant probabilities are supported on self-supporting sets with special properties. When T is equicontinuous, this turns out to be the case.

DEFINITION 2.20: A self-supporting set F is called *minimal* if $\phi \neq F_0 \subseteq F$ and F_0 self-supporting implies $F_0 = F$.

The main step in the proof of Theorem 2.12 is the following result.

THEOREM 2.21: *Let T be an equicontinuous m-operator. Then a probability λ is an extreme point of K_T if and only if $\sigma(\lambda)$ is a minimal, self-supporting set* (cf. Sine [33, Corollary 2.3]).

PROOF OF THEOREM 2.12: First suppose that the uniqueness criterion holds with respect to the point s_0. Let μ be an extreme invariant probability. Since $\sigma(\mu)$ is a closed set and is self-supporting, $s_0 \in \sigma(\mu)$. This shows that the supports of any two extreme invariant probabilities must intersect. But in view of Theorem 2.21, the support of any invariant probability is a minimal, self-supporting set. Since the intersection of two self-supporting sets is again self-supporting, it must then be that all extreme invariant probabilities have identical supports. Let λ, μ be two such probabilities with $\sigma(\lambda) = \sigma(\mu)$. As in Sine [33, Theorem 2.2], if $\lambda \neq \mu$ we can find a μ-invariant function g such that $\int g \, d\lambda \neq \int g \, d\mu$. But then g cannot be constant on $\sigma(\mu)$; this contradicts the minimality of $\sigma(\mu)$ implied by Theorem 2.21 (cf. Sine [33, Theorem 1.2]). Thus, $\lambda = \mu$, and there must therefore be only one extreme, invariant probability. But then Corollary 2.16 implies that K_T consists of only one probability.

To prove the other direction, suppose μ is the unique invariant probability in K_T. Let $s_0 \in \sigma(\mu)$ and let U be a neighborhood of s_0. Choose a nonnegative continuous function f that vanishes on the complement of U but such that $1 = f(s_0)$. We may assume that $0 \leqq f \leqq 1$. Define

$$h_U(s) = \begin{cases} 1 & \text{if } s \in U, \\ 0 & \text{otherwise.} \end{cases}$$

Now let s be any point in S. Since T is equicontinuous and μ is the unique invariant probability, $\{A_n(T^*)\delta_s\}$ converges weakly to μ. Thus $(f, A_n(T^*)\sigma_s) \rightarrow (f, \mu) > 0$. In particular, there is a finite n such that $(f, A_n(T^*)\delta_s) > 0$. Since $f \leqq h_U$, it follows that $(h_U, A_n(T^*)\delta_s) > 0$. But then there must be an $n_0 \leqq n - 1$ such that

$$P^{n_0}(s, U) = (T^{*n_0}\delta_s)(U) > 0.$$

Thus P satisfies the uniqueness criterion. This completes the proof of Theorem 2.12.   *Q.E.D.*

3. QUASI-COMPACT MARKOV OPERATORS

The equicontinuous operators studied in the previous section enjoyed several very useful properties. Unfortunately, it is in practice often difficult to determine directly whether a given operator is equicontinuous. In this section we introduce

the class of quasi-compact Markov operators. If a quasi-compact operator is also a stable operator it is then equicontinuous. In the following section we shall see that, unlike the general equicontinuous operator, quasi-compact operators are easily recognized; in fact, we shall find that "most" operators are quasi-compact.

If a given m-operator T can be shown to be quasi-compact, one can immediately deduce the existence of at least one invariant probability and also the convergence of the sequence $\{A_n(T^*)\lambda\}$, for any probability λ, to some invariant probability. This is true even if T *fails* to be stable. In fact even more is true. The set of invariant probabilities is finite dimensional and the convergence occurs at an arithmetic rate with respect to the total variation norm on probability measures; this, of course, is a much stronger convergence result than simply convergence in the weak topology (a much coarser topology than the total variation topology). Unfortunately, as we shall show by example, the uniqueness theorem, 2.12, is not valid, even for a quasi-compact operator, unless the operator is also stable.

The last part of this section is devoted to a generalization of the uniqueness criterion of Section 2 that is applicable to stable, quasi-compact m-operators. If this criterion is satisfied, then one can infer not only the uniqueness of the invariant probability, but the convergence of the sequence $\{T^{*i}\lambda\}$, for any probability λ, to the invariant probability at a geometric rate in the total variation norm.

3.1 *Quasi-Compact Operators*

To begin the discussion, we introduce the following notation. If X is a Banach space write $\lim(X)$ to denote the Banach space of all continuous, linear maps of X to itself. Define the unit ball of the Banach space X to be $bX \equiv \{x \in X \mid |x| \leqq 1\}$. Note that $\lim(X)$ is a Banach space under the operator norm $|T| \equiv \sup_{x \in bX} |Tx|$.

DEFINITION 3.1: Let X be a Banach space and T an operator in $\lim(X)$. Then T is said to be *compact* if the image of bX under T has compact closure in X. The operator T is said to be *quasi-compact* if there is a compact operator L and an integer n such that $|T^n - L| < 1$.

The following theorem reveals the importance of quasi-compact operators.

THEOREM 3.2 (Dunford and Schwartz [9, p. 711, Corollary 4]): *Let T be a quasi-compact operator in $\lim(X)$ with norm not exceeding one. Then the sequence of linear operators $\{A_n(T)\}$ converges in the Banach space $\lim(X)$ to an operator V. V is a projection (i.e., $V^2 = V$); the image of X under V is finite dimensional and consists precisely of the fixed points of T.*

Theorem 3.2 allows us to prove that every stable, quasi-compact m-operator is equicontinuous.

THEOREM 3.3: *Let $T \in \text{lin}(B(S))$ be a stable, quasi-compact m-operator. Then T is equicontinuous.*

PROOF: Since T is stable, the sequence $A_n(T)f$ is a sequence of continuous functions whenever f is continuous. In view of Theorem 3.2, $A_n(T)$ is a convergent sequence of operators in the operator norm. Hence T is uniformly mean stable.

We now show that T is a tight operator. First note that T^* is also a quasi-compact operator for $|T^{*n} - L^*| < 1$. And L^* is compact if and only if L is compact [9, p. 485, Theorem 2]. Note that if $A_n(T) \to V$, then $A_n(T^*) \to V^*$.

Now let K_T denote the set of invariant probabilities, i.e., the probabilities that are fixed points of T^*. It is well known (see, for example, Neveu [24, p. 181]) that when T is quasi-compact, K_T is a closed (in the total variation topology), convex set with a finite number of extreme points. Denote these by $\{\eta_1, \ldots, \eta_I\}$. Since S is assumed to be a metric subspace of a complete, separable metric space, every probability on S is tight [27, p. 29, Theorem 3.2]. Thus given $\epsilon > 0$, there exists compact sets K_ϵ^i, $i = 1, \ldots, I$, such that $\eta_i(K_\epsilon^i) \geq 1 - \epsilon$. Let $K_\epsilon \equiv \bigcup_1^I K_\epsilon^i$. Then if λ is in K_T, $\lambda(K_\epsilon) \geq 1 - \epsilon$.

The fact that T is a tight operator now follows easily. For from Theorem 3.2 we conclude that the sequence of operators $\{T^{*i}, i = 0, 1, \ldots\}$ must be infinitely often in every neighborhood of the operator V^*. Given $\epsilon > 0$, choose an integer m so that $|T^{*m} - V^*| < \epsilon/2$. Then, for any probability measure μ, $|T^{*m}\mu - V^*\mu| < \epsilon/2$. Since K_T has been shown to be a tight family, and since $V^*\mu \in K_T$, we can find a compact set $B_{\epsilon/2}$ such that $V^*\mu(B_{\epsilon/2}) \geq 1 - \epsilon/2$. But then $T^{*m}\mu(B_{\epsilon/2}) \geq 1 - \epsilon$. Defining $K_\epsilon \equiv B_{\epsilon/2}$, and letting $\mu = \delta_s$, we have shown the family of probabilities $\{T^{*m}\delta_s\}$ to be ϵ-tight. Hence T is a tight operator. $\qquad Q.E.D.$

Of course, if a quasi-compact m-operator fails to be stable it cannot be equicontinuous. Nevertheless, Theorem 3.2 asserts that for such operators, invariant probabilities exist (because T^* is quasi-compact and preserves probabilities) and that the Cesaro iterates of any initial probability converge in total variation to an invariant probability.

If an m-operator is quasi-compact, one gets "free of charge" the following rate of convergence result.

THEOREM 3.4 (Yosida and Kakutani [34, p. 204, Corollary]): *Let T, V be as in Theorem 3.2. Then there is a constant M such that*

$$|A_n(T) - V| \leq \frac{M}{n} \qquad for \quad n = 1, 2, \ldots.$$

Putting together Theorems 3.2 and 3.4, we can assert that whenever T is quasi-compact the Cesaro iterates of any initial probability converge at an arithmetic rate in the total variation norm to an invariant probability. This is much stronger than simply asserting convergence in the weak topology on probability measures.

We close this subsection with an example showing that the stability hypothesis in Theorem 2.12 is a necessary one.

Let $S = [0, 1]$ and let η denote lebesgue measure. Define a transition probability by the formula

$$P(s, A) = \begin{cases} \eta(A) & \text{if } s \neq 0, \\ 1 & \text{if } s = 0 \text{ and } 0 \in A, \\ 0 & \text{if } s = 0 \text{ and } 0 \notin A. \end{cases}$$

The m-operator T defined by P is evidently compact. For the range of T is two dimensional; if $f \in B(S)$, then $Tf(0) = f(0)$ and $Tf(s) = \int f \, d\eta$ if $s \neq 0$. Thus T is a quasi-compact operator, and Theorems 3.2 and 3.4 give all the desired existence and convergence results. Furthermore, the uniqueness criterion 2.11 holds with $s_0 = 0$. But clearly there is more than one invariant probability; for the point mass concentrated at zero as well as lebesgue measure on $[0, 1]$ are invariant. Thus Theorem 2.12 does not hold. Of course, the problem is that T is not a stable operator, as can easily be seen by considering Tf where $f(s) = s$.

3.2 A Generalized Uniqueness Criterion

It often happens that one would like to infer convergence of the sequences $\{T^{*n}\nu\}$, for any probability ν, rather than simply convergence of $\{A_n(T^*)\nu\}$. One might also ask for rates of convergence that are geometric rather than arithmetic. In this subsection, we develop for stable, quasi-compact m-operators a generalization of the Uniqueness Criterion 2.11 which, if satisfied, implies these results.

GENERALIZED UNIQUENESS CRITERION 3.5: *There should exist a point s_0 in S with the following property. For any integer $k \geq 1$, any point s in S, and any neighborhood U of s_0, one can find an integer n such that $P^{nk}(s, U) > 0$.*

The main theorem in this subsection is the following.

THEOREM 3.6: *Let T be a stable, quasi-compact m-operator with transition probability P satisfying 3.5. Regard T as an operator on $C(S)$. Then 1 is the only proper value of T of modulus one.*

The significance of Theorem 3.6 arises from the following result.

THEOREM 3.7 (Yosida and Kakutani [34, p. 204, Corollary]): *Let T, V be as in Theorem 3.2. In order that the sequence $\{T^n\}$ be convergent in $\text{lin}(X)$ it is necessary and sufficient that 1 be the only proper value of T of modulus one. In this case, there exist constants M, $\epsilon > 0$ such that*

$$|T^n - V| \leq \frac{M}{(1 + \epsilon)^n} \qquad\qquad (n = 1, 2, \ldots).$$

We apply Theorems 3.6 and 3.7 to give the desired convergence results as follows. Let $X = C(S)$. Theorem 3.6 shows that if T is regarded as an operator on $C(S)$, then T has 1 as its only proper value of modulus one. (Of course, T maps $C(S)$ to itself because T is a stable operator by hypothesis.) Theorem 3.7 then tells us that the sequence of operators $\{T^n\}$ converges geometrically to a limit V. Using Proposition 1.7, we can then infer the same result for the operators $\{T^{*n}\}$, V^* on $C(S)^*$. In general, $C(S)^*$ properly includes the space $rca(S)$ (because S need not be compact). However, the Banach space $rca(S)$ is a closed subspace of $C(S)^*$ and is in fact mapped to itself by T^* (because T is an m-operator). Thus, as operators on $rca(S)$, $\{T^{*n}\}$ converges geometrically to V^*.

The readers should note that the Uniqueness Criterion 2.11 is just the case $k = 1$ of Criterion 3.5. Furthermore, Theorem 3.3 tells us that every stable, quasi-compact m-operator is equicontinuous. Therefore, if the transition probability for such an operator satisfies 3.5, we can apply Theorem 2.12 to infer the uniqueness of the invariant probability.

The reader uninterested in the details behind the proof of 3.6 should now proceed to Section 4.

To prove Theorem 3.6 we need some preliminary definitions.

Let $\hat{C}(S)$ denote the Banach space of all bounded *complex valued* continuous functions with the sup norm (the absolute value of a complex number is just its complex modulus). If T is a stable m-operator, then T defines an operator $\hat{T}$ on $\hat{C}(S)$ in the following way. Suppose g is in $\hat{C}(S)$. Then g can be written uniquely as $g(s) = f_1(s) + if_2(s)$ with f_j in $C(S)$. Define $\hat{T}g \equiv Tf_1 + iTf_2$. Clearly, $\hat{T}$ is a continuous linear operator of norm one on $\hat{C}(S)$.

DEFINITION 3.8: If T is a stable m-operator, a complex number α is a *proper value* of T on $C(S)$ if there is an $x \in C(S)$ such that $\hat{T}x = \alpha x$.

THEOREM 3.9: *Let T be a stable, quasi-compact m-operator. Suppose that for each n the only fixed points of the operator T^n on $C(S)$ are the constant functions. Then 1 is the only proper value of T of modulus 1.*

PROOF: Let α be a proper value of T of modulus 1, and let $g \in \hat{C}(S)$ be a function such that $\hat{T}g = \alpha g$. Now $C(S)$ is isometrically isomorphic with the Banach space of all real valued continuous functions on some compact Hausdorff space [9, p. 274, Theorem 18]. Therefore, we can apply [9, p. 711, Lemma 5] to infer the existence of an integer k such that $\alpha^k = 1$. Therefore $\hat{T}^k g = g$ and g is a fixed point of $\hat{T}^k$. Write g uniquely as $g = g_1 + ig_2$. From the definition of $\hat{T}$ we note that

$$\hat{T}^k g = T^k g_1 + iT^k g_2$$
$$= g_1 + ig_2.$$

Therefore, g_1, g_2 are fixed points of T^k, and by hypothesis must therefore be real constants. Thus, g is a complex *constant* and is therefore a fixed point of $\hat{T}$ (not

just $\hat{T}^k$). This last observation follows from the definition of $\hat{T}$ and the fact that m-operators preserve real constants. But then $g = \hat{T}g = \alpha g$. Therefore $\alpha = 1$.

$$Q.E.D.$$

Theorem 3.6 will now be a direct consequence of Theorem 3.9 and of the following result.

THEOREM 3.10: *Let T be a stable, quasi-compact m-operator. Suppose the corresponding transition probability P satisfies Condition 3.5. Then, for each k, the only fixed points of T^k in $C(S)$ are the constant functions.*

PROOF: First note that if T is quasi-compact, then so is T^k for each k. For suppose L is a compact operator and n an integer such that $|T^n - L| < 1$. We then infer $|T^{kn} - L \cdot T^{(k-1)n}| < 1$. But [9, p. 486, Corollary 5] asserts that $LT^{(k-1)n}$ is a compact operator. Hence T^k is quasi-compact, and stable. Therefore T^k is equicontinuous for each k.

Now note that if P satisfies the conditions in 3.5, then P^k satisfies the uniqueness criterion 2.11. But P^k is just the transition probability for the m-operator T^k. Theorem 2.12 then asserts that there is a unique invariant probability for the operator T^k. But this in turn implies that the only fixed points of T^k in $C(S)$ are the constant functions. To see this, suppose to the contrary that f is not constant and $T^k f = f$. There must then be two points s_0, s_1 such that

$$(f, \delta_{s_0}) = f(s_0) \neq f(s_1) = (f, \delta_{s_1}).$$

But then

$$(f, \delta_{s_0}) = \left(A_n(T^k)f, \delta_{s_0} \right)$$

$$= \left(f, A_n(T^{*k})\delta_{s_0} \right)$$

$$\neq \left(f, A_n(T^{*k})\delta_{s_1} \right)$$

$$= (f, \delta_{s_1}).$$

Hence $\lim_n (A_n(T^{*k})\delta_{s_0}) \neq \lim_n A_n(T^{*k})\delta_{s_1}$. This contradicts the uniqueness of the invariant probability for T^k.

$$Q.E.D.$$

4. RECOGNIZING QUASI-COMPACT OPERATORS

The results of Section 3 make it desirable to develop easy-to-apply criteria that allow one to recognize a quasi-compact operator by identifying characteristics of the corresponding transition probability. This section provides proofs of several well known results along these lines.

The first is that any operator defined by integrating a bounded transition density with respect to a finite measure is quasi-compact. The second is the well known characterization of the class of quasi-compact m-operators as precisely

the class of operators with transition probabilities that satisfy Doeblin's famous condition. Finally, we show that a non-trivial convex combination of a quasi-compact m-operator and any m-operator is again quasi-compact. This last result immediately implies that quasi-compact m-operators are open and dense in the set of all m-operators.

4.1 Weakly Compact Operators

DEFINITION 4.1: The *weak topology* on a Banach space X is defined to be the coarsest topology on X making every element of X^* continuous. Recall that X^* denotes the Banach space dual to X.

WARNING 4.2: If $X = rca(S)$ the weak topology on X defined in 4.1 is *not* the same topology as the weak topology of Definition 2.5. This is because, in general, $rca(S)^* \neq C(S)$. However, no serious confusion should result since we shall never apply Definition 4.1 to $rca(S)$.

DEFINITION 4.3: If T is in $\lin(X)$, T is said to be *weakly compact* if the image under T of the unit ball of X has compact closure in the weak topology on X. An operator T is said to be *quasi-weakly compact* if there is a weakly compact operator L and an integer n such that $|T^n - L| < 1$.

The class of weakly compact m-operators properly includes the class of compact m-operators. However, one has the following result.

PROPOSITION 4.4: *An operator T in $\lin(B(S))$ is quasi-weakly compact if and only if it is quasi-compact.*

PROOF: To prove the non-trivial direction, we first make some preliminary observations. First of all, if K and L are weakly compact operators on $B(S)$, then $K \cdot L$ is a compact operator. To see this, first note that $B(S)$ is isometrically, isomorphic to the Banach space of all continuous functions on some compact Hausdorff space [9, p. 274, Theorem 18]. The assertion then follows from the similar assertion for $C(S)$, S compact Hausdorff [9, p. 494, Corollary 5]. Secondly, the product of a weakly compact operator with any continuous operator is again weakly compact [9, p. 484, Theorem 5].

With these observations in hand, we can now use an argument appearing in Jamison [18].

Suppose T is quasi-weakly compact. Then there is an integer n such that $T^n = L + V$ where L is weakly compact and $|V| < 1$. Let L_m be the *compact* operator defined to be the sum of all terms in $(L + V)^m$ in which L appears as a factor at least twice. Then

$$(L + V)^m = L_m + LV^{m-1} + VLV^{m-2}$$

$$+ \cdots + V^{m-1}L + V^m.$$

Hence

$$|T^{nm} - L_m| \leqq m|L| |V|^{m-1} + |V|^m.$$

Since $|V| < 1$, we may choose m so large as to make the right hand side of this inequality strictly less than one. Thus T is quasi-compact. $Q.E.D.$

We shall next exhibit a large class of easily recognized, weakly compact operators on $B(S)$.

THEOREM 4.6: *Suppose $p(s,t)$ is a real valued function on $S \times S$ that is bounded and measurable with respect to the smallest σ-field on $S \times S$ that contains $\mathbf{S} \times \mathbf{S}$. Let η be a finite, positive measure on $(S, \mathbf{S})$. Define an operator $T \in \text{lin}(B(S))$ by the formula*

$$Tf(s) = \int f(t)p(s,t)\eta(dt).$$

Then T is weakly compact.

PROOF: Since T is weakly compact if and only if T^* is weakly compact [**9**, p. 485, Theorem 8] it suffices to show that $T^* : ba(S) \to ba(S)$, is weakly compact. By Theorem 12, p. 314 of [**9**], it suffices to exhibit a nonnegative element η of $ba(S)$ such that $\lim_{\eta(A) \to 0} T^*\mu(A) = 0$ uniformly for all μ in the unit ball of $ba(S)$.

Let η be as in the hypothesis of the theorem. Define $M \equiv \sup_{s,t} |p(s,t)|$. Then $M < \infty$ by hypothesis. By definition

$$T^*\mu(A) = \int_S \left(\int_A p(s,t)\eta(dt) \right) \mu(ds)$$

$$\leqq \int_S \left| \left(\int_A p(s,t)\eta(dt) \right) \right| \mu(ds)$$

$$\leqq \int_S \left(\int_A |p(s,t)|\eta(dt) \right) \mu(ds)$$

$$\leqq \int_S M\eta(A)\mu(ds)$$

$$\leqq \int_S M\eta(A)|\mu|(ds)$$

(where $|\mu|$ is the total variation (see [**9**, p. 97]) of the set function μ)

$$\leqq M\eta(A) \qquad \text{since} \quad |\mu|(S) = 1.$$

This completes the proof of the theorem. $Q.E.D.$

COROLLARY 4.7: *Maintain the hypothesis of Theorem 4.6, dropping only the requirement that η be a finite measure. Instead assume that the function*

$$q(t) \equiv \sup_{s \in S} p(s,t)$$

is integrable, i.e., that $\int q(t)\eta(dt) < \infty$. Then T is weakly compact.

PROOF: Define a finite measure ν by the formula $\nu(A) = \int_A q(t)\eta(dt)$. Define a bounded, measurable function $\tilde{p}(s,t)$ by

$$\tilde{p}(s,t) = \begin{cases} \dfrac{p(s,t)}{q(t)} & \text{if } q(t) \neq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Clearly

$$\int f(t)p(s,t)\eta(dt) = \int f(t)\tilde{p}(s,t)\nu(dt).$$

The conclusion now follows from 4.6.   Q.E.D.

4.2 Doeblin's Condition and Quasi-Compactness

In this subsection we prove that an m-operator is quasi-compact if and only if its transition probability satisfies Doeblin's condition.

DEFINITION 4.8: Let P be a transition probability. We say that P satisfies *Doeblin's condition* if the following assertion holds. There is a probability η, an integer n, and an ϵ with $0 < \epsilon < 1$ such that if $A \in S$ and $\eta(A) \leqq \epsilon$, then $P^n(s,A) \leqq 1 - \epsilon$ for all s.

Loosely speaking, a transition probability satisfies Doeblin's condition if there is a probability η such that the corresponding Markov process is not concentrated on sets of small η measure.

THEOREM 4.9: *Let T be an m-operator and P the corresponding transition probability. Then T is quasi-compact if and only if P satisfies Doeblin's condition.*

PROOF: Suppose T is quasi-compact. It is well known (see for example Neveu [24, p. 181]) that the set K_T consisting of all invariant probabilities then has a finite number of extreme points $\{\eta_1, \ldots, \eta_K\}$. Define $\eta \equiv (1/K)\sum \eta_i$.

Now choose a positive number δ so that $0 < (2 - 3\delta)/2K < 1$. If $\eta(A) \leqq (2 - 3\delta)/2K$, then $\eta_j(A) \leqq 1 - (3/2)\delta$ for all j and hence $\mu(A) \leqq 1 - (3/2)\delta$ for any μ in K_T. Let V be the limit of the sequence of operators $\{A_n(T^*)\}$. Choose an integer n so that $|T^{*n} - V^*| < \delta/2$. We conclude that

$$|P^n(s,A) - V^*\delta_s(A)| \leqq |T^{*n}\delta_s - V^*\delta_s|$$

$$\leqq |T^{*n} - V^*| < \delta/2.$$

Hence $P^n(s, A) \leqq 1 - \delta$ if $\eta(A) \leqq (2 - 3\delta)/2K$. Define $\epsilon = \min\{(2 - 3\delta)/2K, \delta\}$. Clearly P satisfies Doeblin's condition with respect to η, ϵ, and n.

We next show that if P satisfies Doeblin's condition, then T is quasi-weakly compact, hence quasi-compact by Proposition 4.4.

Let η, ϵ, n be the data with respect to which P satisfies Doeblin's condition. Since S was assumed to be a separable metric space, its Borel σ-field is countably generated. We can therefore apply a well known lemma (see, for example, Revuz [28, p. 32, Lemma 5.3], or Orey [26, p. 5, Proposition 1.1]) to conclude that

$$P^n(s, A) = \int_A q(s, t)\eta(dt) + Q(s, A)$$

where $q(s, t)$ is a positive measurable function and the measures $Q(s, \cdot)$ are all singular with respect to η.

Define an operator L on $B(S)$ by the equation

$$Lf(s) = \int f(t)p(s, t)\eta(dt) \qquad \text{where}$$

$$p(s, t) \equiv \min\left(q(s, t), \frac{2}{\epsilon}\right).$$

By Theorem 4.6, L is a weakly compact operator. We claim that $|T^n - L| < 1$. Since

$$|T^n - L| = 1 - \inf_s \int p(s, t)\eta(dt)$$

it will suffice to show that

$$1 - \int p(s, t)\eta(dt) \leqq 1 - \epsilon \qquad \text{for all } s.$$

For each s, choose a subset $N_s \subseteq S$ to satisfy

$$\eta(N_s) = 0, \qquad Q(s, N_s) = Q(s, S).$$

Define $B_s \equiv \{t \in S - N_s \mid q(s, t) \geqq 2/\epsilon\}$. Since $P(s, S) \equiv 1$, we conclude that $\eta(B_s) \leqq \epsilon/2$. Now

$$1 - \int p(s, t)\eta(dt) = P(s, S) - \int p(s, t)\eta(dt)$$

$$= \int (q(s, t) - p(s, t))\eta(dt) + Q(s, S)$$

$$= \int_{N_s \cap B_s} [q(s, t) - p(s, t)]\eta(dt) + Q(s, N_s \cap B_s)$$

$$\leqq \int_{N_s \cap B_s} q(s, t)\eta(dt) + Q(s, N_s \cap B_s)$$

$$= P(s, N_s \cap B_s) \leqq 1 - \epsilon$$

because $\eta(N_s \cap B_s) \leqq \epsilon/2 \leqq \epsilon$. $Q.E.D.$

4.3 *A Final Recognition Result*

In this subsection we prove a simple recognition theorem that has many important applications.

Let X be a Banach space and suppose T_1, T_2 are in $\lin(X)$. Let α be a real number $0 < \alpha \leqq 1$.

THEOREM 4.10: *Suppose $|T_i| \leqq 1$ and that T_1 is quasi-compact. Then $T \equiv \alpha T_1 + (1 - \alpha) T_2$ is quasi-compact.*

PROOF: Let L be a compact operator and n an integer such that $|T_1^n - L| < 1$. Note that for any real number α, $\alpha^n L$ is a compact operator. Consider the expression

$$|T^n - \alpha^n L| = |(\alpha T_1 + (1 - \alpha) T_2)^n - \alpha^n L|.$$

Since $|T_i| \leqq 1$, the binomial expansion of T^n shows that

$$|(\alpha T_1 + (1 - \alpha) T_2)^n - \alpha^n T_1^n| \leqq 1 - \alpha^n.$$

Thus

$$
\begin{aligned}
|T^n - \alpha^n L| &\leqq \alpha^n |T_1^n - L| + 1 - \alpha^n \\
&< \alpha^n + 1 - \alpha^n = 1. \hspace{3em} Q.E.D.
\end{aligned}
$$

Theorem 4.10 has an easy and interesting corollary. Let T_1, T_2 be m-operators. Suppose α is a measurable, real valued function on S with $0 \leqq \alpha(s) \leqq 1$. Define the operator $T = \alpha(s) T_1 + (1 - \alpha(s)) T_2$ by the formula

$$Tf(s) = \alpha(s)(T_1 f(s)) + (1 - \alpha(s))(T_2 f(s)).$$

The transition probability for T is just $\alpha(s) P_1(s, \cdot) + (1 - \alpha(s)) P_2(s, \cdot)$. Thus the process corresponding to T is just a weighted average of the processes corresponding to T_1 and T_2, with weights depending upon the state s.

COROLLARY 4.11: *Let T_1, T_2 be m-operators and let α be a real valued function defined as above. Suppose that T_1 is quasi-compact, and that $\inf_s \alpha(s) > 0$. Then T is quasi-compact.*

PROOF: Let $B \equiv \inf_s \alpha(s)$. Then note that

$$T = B T_1 + (\alpha(s) - B) T_1 + (1 - \alpha(s)) T_2.$$

Since $B > 0$, Theorem 4.10 applies and the result is proven. $\hspace{2em} Q.E.D.$

As a final remark, the reader should note that Theorem 4.10 asserts that the set of quasi-compact m-operators on $B(S)$ is dense in the set of all m-operators. Since the set of quasi-compact m-operators is clearly relatively open in the set of all m-operators, it is therefore open and dense.

5. RANDOM DYNAMICAL SYSTEMS

In many economic problems a Markov process on a state space S is defined as a "state dependent stochastic average" of deterministic adjustment processes on S. Processes defined in this way are studied in this section where they are termed random dynamical systems. We shall develop conditions under which the results of previous sections can be applied to such processes.

DEFINITION 5.1: A *random dynamical system* (r.d.s.) is defined by the following data: (a) a state space $(S, \mathbf{S})$ satisfying the assumption made in Section 2, i.e., S is a separable metric space, and a Borel subset and metric subspace of a complete, separable metric space; $\mathbf{S}$ is its Borel σ-field; (b) a measurable space $(E, \mathbf{E})$ of "events"; (c) a stochastic kernel $Q : S \times \mathbf{E} \to [0, 1]$; thus $Q(s, A)$ is the probability of realizing the event $A \in \mathbf{E}$, given that the current state is $s \in S$; (d) a mapping $\theta : S \times E \to S$; for each s in S, $\theta(s, \cdot)$ is assumed to be measurable, while for each e in E, $\theta(\cdot, e)$ is assumed to be continuous.

Loosely speaking, the state of the system evolves from one period to the next in the following way. If the state at date t is s_t, an event $e_t \in E$ is realized according to the probability law $Q(s_t, \cdot)$. The state at date $t + 1$ is then $s_{t+1} = \theta(s_t, e_t)$.

Given a random dynamical system, one should be able to define a transition probability on the state space in the following way. If $B \subseteq S \times E$ and s is in S, define $B_s \equiv \{ e \in E \mid (s, e) \in B \}$. Then we define a function $P = S \times \mathbf{S} \to [0, 1]$ by the formula:

$$P(s, A) = Q\left(s, (\theta^{-1}A)_s\right).$$

The first theorem in this section asserts that this definition actually works.

THEOREM 5.2: *Let the function P be defined as above from the data of a random dynamical system. Then P is a transition probability on $(S, \mathbf{S})$.*

To prove 5.2, we need to establish two facts. We must first show that for each s in S, $P(s, \cdot)$ is a probability on $(S, \mathbf{S})$. But it is easy to see that this must be so. For $(\theta^{-1}A)_s \in \mathbf{E}$ for each s in S and A in $\mathbf{S}$ because $\theta(s, \cdot)$ is assumed to be measurable. In other words $(\theta^{-1}A)_s = \{ e \mid \theta(s, e) \in A \}$. This last equality also shows that $P(s, \cdot)$ is a probability; for $Q(s, \cdot)$ is and taking inverse images under $\theta(s, \cdot)$ preserves all the required set theoretic operations.

The proof of 5.2 will be complete if we can show that for each $A \in \mathbf{S}$, $P(\cdot, A)$ is a measurable function. We do this in the following sequence of lemmas.

DEFINITION 5.3: $\mathbf{S} \otimes \mathbf{E}$ is the smallest σ-field of $S \times E$ that contains all sets of the form $A \times B$ with $A \in \mathbf{S}$, $B \in \mathbf{E}$.

LEMMA 5.4: *The mapping θ is $\mathbf{S} \otimes \mathbf{E}$ measurable.*

PROOF: Since $\mathbf{S}$ is the Borel σ-field of S, it will suffice to show that for each closed set $F \subseteq S$, $\theta^{-1}F$ is in $\mathbf{S} \otimes \mathbf{E}$. The proof will depend upon the fact that S is a separable metric space.

Let $C = \{s_1, s_2, \ldots\}$ be a countable dense subset of S. For each pair of integers $i, j > 0$ define a subset A_j^i of S by the equality $A_j^i \equiv \{s \in S \mid d(s_j, s) < 1/i\}$. Clearly $A_j^i \in \mathbf{S}$ and $S = \bigcup_{j=1}^{\infty} A_j^i$ for any fixed i. Next, define a family of subsets of E by the equalities

$$B_j^i \equiv \{e \in E \mid d(\theta(s_j, e), F) < 1/i\}$$

where $d(s, F) \equiv \inf_{t \in F} d(s, t)$. Since B_j^i is the inverse image of an open subset of S under $\theta(s_j, \cdot)$, $B_j^i \in \mathbf{E}$.

We now claim that

$$\theta^{-1}F = \bigcap_i \left[\bigcup_j \left(A_j^i \times B_j^i \right) \right].$$

To prove the claim, suppose first that (s, e) is an element of the right hand side. It follows that for each i, there is an integer j_i such that $(s, e) \in A_{j_i}^i \times B_{j_i}^i$. Thus $d(s_{j_i}, s) < 1/i$ and $d(\theta(s_{j_i}, e), F) < 1/i$. The sequence $\{s_{j_i}, i = 1, 2, \ldots\}$ therefore converges to s and, since $\theta(\cdot, e)$ is continuous, $\{\theta(s_{j_i}, e)\}$ converges to $\theta(s, e)$. But $\theta(s_{j_i}, e)$ must also converge to some point in F. Hence $\theta(s, e) \in F$ and so $(s, e) \in \theta^{-1}F$.

Now suppose that $(s, e) \in \theta^{-1}F$. Let $\{s_{j_i}\}$ be a sequence from C that converges to s. Because $\theta(\cdot, e)$ is continuous, we may choose a subsequence $\{s_{j_i}\}$ such that $d(s_{j_i}, s) < 1/i$ and $d(\theta(s_{j_i}, e), F) < 1/i$. But then $(s, e) \in A_{j_i}^i \times B_{j_i}^i$. This completes the proof of the claim and of the lemma.   Q.E.D.

LEMMA 5.5: *For any $A \in \mathbf{S} \otimes \mathbf{E}$, $Q(s, A_s)$ is a measurable function of s.*

PROOF: Let $\mathbf{A}$ be the subset of $\mathbf{S} \otimes \mathbf{E}$ consisting of those sets A for which $Q(s, A_s)$ is measurable. We are going to show that $\mathbf{A}$ contains the Boolean algebra of rectangles and also that $\mathbf{A}$ is a monotone class. Since the monotone class lemma (e.g., Neveu [24, p. 14]) implies that $\mathbf{S} \otimes \mathbf{E}$ is the smallest monotone class containing the Boolean algebra of rectangles, this will prove the lemma.

We first show that $\mathbf{A}$ contains the Boolean algebra generated by the rectangles, i.e., that it contains all sets of the form $A = B \times C$ with $B \in \mathbf{S}$, $C \in \mathbf{E}$, and is closed under finite unions and intersections, and under complementation. Let $A = B \times C$ be a rectangle. Then

$$Q(s, A_s) = \begin{cases} 0 & \text{if } s \notin B, \\ Q(s, C) & \text{if } s \in B. \end{cases}$$

This is clearly a measurable function. Now let $\{A_1, \ldots, A_n\}$ be a finite collection of n rectangles. We wish to show that $Q(s, (\bigcup_1^n A_i)_s)$ is measurable. We do this by induction, noting that the assertion is true for $n = 1$, and then assuming

that it is true for $n - 1$. But note that

$$Q\left(s, \left(\bigcup_1^n A_i\right)_s\right) = Q\left(s, \left(\bigcup_1^{n-1} A_i\right)_s\right) + Q(s, (A_n)_s)$$

$$- Q\left(s, \left(\bigcup_1^{n-1} (A_i \cap A_n)\right)_s\right).$$

This follows from the observation that the operation of "taking the s-section of set A" commutes with all the usual operations of union, intersection, and complementation. Since the intersection of two rectangles is again a rectangle, this exhibits $Q(s, (\bigcup_1^n A_i)_s)$ as a sum of functions known to be measurable by induction. Finally, note that the complement of any rectangle is a finite union of rectangles, and that the intersection of a finite number of rectangles is the complement of the union of the complements. Thus, **A** contains the Boolean algebra generated by the rectangles.

We next show that **A** is a monotone class. It will suffice to show that if $A_1 \subseteq A_2 \subseteq \ldots$ is a countable, increasing sequence of sets in **A**, then $A \equiv \bigcup_1^\infty A_i$ is in **A**. Define $B_n = \bigcup_1^n A_i$. For any fixed s, we know that

$$\lim_{n \to \infty} Q(s, (B_n)_s) = Q(s, A_s)$$

because taking s-sections commutes with unions and because $Q(s, \cdot)$ is a probability. Thus $Q(s, A_s)$ is the pointwise limit of measurable functions and is therefore measurable. This shows that **A** is a monotone class and thus completes the proof of the lemma. $Q.E.D.$

If $A \in \mathbf{S}$, Lemma 5.4 shows that $\theta^{-1}A \in \mathbf{S} \otimes \mathbf{E}$. Lemma 5.5 then shows that $Q(s, (\theta^{-1}A)_s)$ is measurable, i.e., that $P(s, A)$ is measurable, thus completing the proof of Theorem 5.2.

In view of Theorem 5.2, a random dynamical system defines a transition probability P and thus an m-operator T on $B(S)$. The results of Section 2 make it desirable to find conditions under which T can be shown to be a stable operator. The next proposition is a result in this direction that seems to be sufficient for most applications. Let $ca(E)$ denote the subset of $ba(E)$ which consists of the countably additive set functions.

PROPOSITION 5.6: *Suppose a random dynamical system is given. Assume that the map $S \to ca(E)$ defined by $s \to Q(s, \cdot)$ is continuous when $ca(E)$ is regarded as a Banach space in the total variation norm. If $f \in C(S)$, it then follows that $Tf \in C(S)$.*

PROOF: Let $f \in C(S)$ be given, and let $\{s_n\}$ be a convergent sequence in S with limit s_0. We need to show that $Tf(s_n)$ converges to $Tf(s_0)$.

For each s in S, denote by f_s the measurable function on E defined by

$f_s(e) \equiv f(\theta(s, e))$. Note that

$$Tf(s) \equiv \int f(t) P(s, dt)$$

$$= \int f(\theta(s, e)) Q(s, de)$$

$$\equiv (f_s, Q(s, \cdot)).$$

Because f is a bounded function, the family of measurable functions $\{f_{s_n}\}$ is contained in a bounded subset of $B(E)$. Since $s \to Q(s, \cdot)$ is assumed to be continuous with respect to the total variation norm on $ca(E)$ it follows that

$$\lim_{n \to \infty} |(f_{s_n}, Q(s_n, \cdot) - Q(s_0, \cdot))| = 0.$$

Thus

$$\lim_{n \to \infty} (f_{s_n}, Q(s_n, \cdot)) = \lim_{n \to \infty} (f_{s_n}, Q(s_0, \cdot))$$

if either limit exists. But since f_{s_n} converges pointwise to f_{s_0} (since $\theta(\cdot, e)$ is continuous for each e) the Lebesgue dominated convergence theorem implies that the limit on the right hand side exists and equals $(f_{s_0}, Q(s_0, \cdot)) = Tf(s_0)$. Thus $\lim_n Tf(s_n) = Tf(s_0)$. $Q.E.D.$

Here are three commonly encountered situations where the hypothesis of Proposition 5.6 is satisfied.

(a) The probability $Q(s, \cdot)$ is independent of s. Hence the map $s \to Q(s, \cdot)$ is just a constant map and therefore continuous.

(b) More generally, suppose that there is a positive constant M such that for each $A \in \mathbf{E}$,

$$|Q(s, A) - A(t, A)| \leqq Md(s, t).$$

Then

$$|Q(s, \cdot) - Q(t, \cdot)| \leqq 2Md(s, t)$$

and so the map $s \to Q(s, \cdot)$ is Lipschitz with constant not exceeding $2M$.

(c) Suppose there is a $\mathbf{S} \otimes \mathbf{E}$ measurable function $q(s, e)$ such that $s \to q(s, e) \in B(E)$ is norm continuous. Suppose in addition that there is a positive, finite measure η on $(E, \mathbf{E})$ such that $Q(s, A) = \int_A q(s, e) \eta(de)$. Then it is easy to see that the map $s \to Q(s, \cdot)$ is continuous with respect to the total variation norm on $ca(E)$.

In Sections 3 and 4 we saw that quasi-compact m-operators were very well behaved. In view of these results, it is desirable to find conditions under which the m-operators defined by a random dynamical system is quasi-compact. The difficult part of this problem is to describe conditions under which the operator can be shown to be weakly compact. Our next proposition presents a general condition that ensures weak compactness.

PROPOSITION 5.7: *Let T be the m-operator defined by a random dynamical system. Assume that there are nonnegative measurable functions $\alpha(s)$, $\beta(s)$, and positive, finite measures η on $(E, \mathbf{E})$, μ on $(S, \mathbf{S})$ such that* (a) $Q(s, B) \leqq \alpha(s) \cdot \eta(B)$ *for all $B \in \mathbf{E}$;* (b) $\eta((\theta^{-1}A)_s) \leqq \beta(s)\mu(A)$ *for all $A \in \mathbf{S}$. If $\sup_s (\alpha(s) \cdot \beta(s)) < \infty$, then T is weakly compact.*

PROOF: Let $T^* \in \mathrm{lin}(ba(S))$ be the adjoint of T. Now T is weakly compact if and only if T^* is weakly compact [9, p. 485, Theorem 8]. We shall show that T^* is weakly compact. By [9, p. 314, Theorem 12], it is sufficient to exhibit a positive $\nu \in ba(S)$ such that $\lim_{\nu(A) \to 0} T^*\lambda(A) = 0$ uniformly for λ in the unit ball of $ba(S)$ and $A \in \mathbf{S}$.

Note that, for $A \in \mathbf{S}$,

$$P(s, A) \equiv Q\big(s, (\theta^{-1}A)_s\big) \leqq \alpha(s)\eta\big((\theta^{-1}A)_s\big)$$

$$\leqq \alpha(s)\,\beta(s)\,\mu(A) \leqq M \cdot \mu(A).$$

Let λ be in the unit ball of $ba(S)$. Then

$$|T^*\lambda(A)| \equiv \left| \int P(s, A)\lambda(ds) \right| \leqq \int |P(s, A)|\,|\lambda|(ds)$$

$$\leqq M\mu(A)|\lambda|(S) = M\mu(A).$$

Setting $\nu = \mu$, we conclude that the image of the unit ball in $ba(S)$ under T^* has weakly compact closure. Hence T^* is weakly compact.   Q.E.D.

6. RANDOM CONTRACTIONS

The equicontinuous m-operators we have studied have all been quasi-compact operators on $B(S)$. Although we found in Section 4 that such operators are an open, dense subset of all operators, there is an interesting class of equicontinuous operators which do not arise in this way.

EXAMPLE 6.1: Let $S = [0, 1]$ and let $0 < \alpha < 1$. Define a transition probability by the formula

$$P(s, A) = \begin{cases} 1 & \text{if } \alpha \cdot s + (1 - \alpha) \cdot 1/2 \in A, \\ 0 & \text{otherwise.} \end{cases}$$

Thus P describes the process that sends the point s to the point $\alpha s + (1 - \alpha) \cdot 1/2$ with probability one.

It is easy to check directly that the m-operator T defined by P is equicontinuous. But T is certainly not quasi-compact. For if it were, the sequence $\{A_n(T^*)\mu\}$ would, for any probability μ, converge in total variation to $\delta_{1/2}$. But if $s \neq 1/2$,

$$|A_n(T^*)\delta_s - \delta_{1/2}| = 1 \qquad \text{for all } n.$$

Of course, the operator T, although not quasi-compact, is nonetheless well behaved precisely because the transition probability P is defined by a contraction mapping on S. In this section we study a class of random dynamical systems with this property. For obvious reasons, we term them random contractions.

The results in this section are due, in the generality we present them, to Norman [25]. The reader should consult his book as well as Iosifescu [15] for interesting applications to mathematical models of learning.

DEFINITION 6.2: A *random contraction* is a random dynamical system with the following additional properties. (a) There is a constant M such that for all $A \in \mathbf{E}$

$$ |Q(s,A) - Q(t,A)| \leqq Md(s,t). $$

(b) Define

$$ r \equiv \sup_{s \neq t} \int Q(s,de) \frac{d(\theta(s,e),\theta(t,e))}{d(s,t)} $$

Then $r < 1$.

Condition (b) is just the requirement that θ be, on "average," a contraction mapping. For a somewhat more general version of (b), the reader should consult Norman [25, p. 31].

It follows from condition (a) and Proposition 5.6 that the m-operator defined by a random contraction is stable. But, as we shall soon see, much more is true of such operators.

Let $L(S)$ denote the set of all real valued Lipschitz functions on S. Thus a function f is in $L(S)$ if and only if

$$ \sup_{s \neq t} \frac{|f(s) - f(t)|}{d(s,t)} \equiv m(f) < \infty. $$

If $|f|_\infty$ denotes the sup norm of f, then define the Lipschitz norm of f, $|f|_L$ by the equation

$$ |f|_L \equiv |f|_\infty + m(f). $$

It is easy to see that $|\ |_L$ makes $L(S)$ into a Banach space. The following lemma shows that T is in fact a continuous linear operator on $L(S)$.

LEMMA 6.3 (Norman [25, p. 35]): *If T is an m-operator defined by a random contraction, then there is a positive constant R such that*

$$ |Tf|_L \leqq r|f|_L + R|f|_\infty. $$

WARNING 6.4: The m-operator T defined by a random contraction need not have norm equal to one when regarded as an operator on $L(S)$.

Here is the first principal result about random contractions.

THEOREM 6.5 (Norman [**25**, p. 50]). *Suppose the state space S is compact and that T is defined by a random contraction. Then T is a quasi-compact operator on L(S).*

Notice that Theorem 6.5 requires the state space to be compact. This assumption will be relaxed for a wide class of interesting cases in Theorem 6.11.

LEMMA 6.6: *Let T be an operator on C(S) that is quasi-compact when regarded as an operator on L(S). Then if $f \in C(S)$ is uniformly continuous, $\{A_n(T)f\}$ is uniformly convergent.*

PROOF: It is well known (see, for example, Dudley [**8**, Lemma 8]) that the closure of $L(S)$ in $C(S)$ with respect to the sup norm is precisely the Banach subspace of all uniformly continuous functions on S.

Let $f \in C(S)$ be uniformly continuous. We shall show that the sequence $\{A_n(T)f\}$ is Cauchy with respect to $|\ |_\infty$. Fix $\epsilon > 0$. Choose an element $g \in L(S)$ such that $|g - f|_\infty < \epsilon/4$. Abbreviate $A_n(T)$ by A_n and note that

$$|A_n f - A_m f|_\infty = |A_n f - A_n g + A_n g - A_m g + A_m g - A_m f|_\infty$$

$$\leqq |A_n|_\infty |f - g|_\infty + |A_m|_\infty |f - g|_\infty + |A_n g - A_m g|_\infty$$

$$\leqq 2|f - g|_\infty + |A_n g - A_m g|_\infty.$$

Since T is quasi-compact on $L(S)$, the sequence $\{A_n g\}$ is Cauchy with respect to $|\ |_L$ and therefore with respect to $|\ |_\infty$. Hence there is an integer n_0 such that when $n, m \geqq n_0$, $|A_n g - A_m g| < \epsilon/2$. Then $n, m \geqq n_0$ implies

$$|A_n f - A_m f|_\infty < 2\frac{\epsilon}{4} + \epsilon/2 = \epsilon. \qquad\qquad Q.E.D.$$

COROLLARY 6.7: *Let S be compact and T be defined by a random contraction. Then T is equicontinuous.*

PROOF: Since S is compact, T is tight. Furthermore, when S is compact, every element of $C(S)$ is uniformly continuous. Hence T is uniformly mean stable by Lemma 6.6. $\qquad\qquad Q.E.D.$

In view of Theorem 6.5, it is natural to ask what convergence results for probabilities are implied by the quasi-compactness of an operator T on $L(S)$. To clarify the situation, we must first describe the relationships between the dual Banach space $L(S)^*$ and the set of probabilities on S.

First of all, it is clear that any probability μ is a continuous, linear functional on $L(S)$, and hence an element of $L(S)^*$. Recall that the norm on $L(S)^*$ is

defined as follows. If $x^* \in L(S)^*$, then

$$|x^*| \equiv \sup_{\substack{f \in L(S) \\ |f|_L \le 1}} |x^*(f)|.$$

Of course, the norm on $L(S)^*$ induces a metric ρ on $L(S)^*$ via the equation

$$\rho(x^*, y^*) \equiv |x^* - y^*|.$$

If μ, λ are probabilities, then

$$\rho(\mu, \lambda) = \sup_{\substack{f \in L(S) \\ |f|_L \le 1}} \left| \int f \, d\mu - \int f \, d\lambda \right|.$$

THEOREM 6.8 (Dudley [8, Theorem 12]): *Suppose that S is a separable metric space. Then the topology on the set of probabilities on S induced by the metric ρ is the weak topology (see Definition 2.5).*

COROLLARY 6.9: *Suppose that S is compact and that T is defined by a random contraction. Then for any probability μ, the sequence $\{A_n(T^*)\mu\}$ converges at an arithmetic rate in a metric defining the weak topology to an invariant probability.*

PROOF: Since T is quasi-compact on $L(S)$, T^* is quasi compact on $L(S)^*$. Hence Theorem 3.4 applies and we conclude that the sequence of operators $\{A_n(T^*)\}$ on $L(S)^*$ is convergent at an arithmetic rate to a projection. In particular, for any probability μ, the sequence $\{A_n(T^*)\mu\}$ converges at an arithmetic rate in the metric ρ to some element of $L(S)^*$. But since S is compact, the set of probabilities on S is weakly compact. Thus the limit must itself be a probability. $\qquad Q.E.D.$

Since Corollary 6.7 asserts that, for compact state spaces, the m-operators induced by random contractions are equicontinuous, it follows that the validity of the uniqueness criterion 2.11 suffices to imply the uniqueness of the invariant probability. But since, in these circumstances, T is actually quasi-compact on $L(S)$ we can show that the validity of the generalized uniqueness criterion 3.5 suffices to prove the convergence of the sequence $\{T^n\}$ of operators on $L(S)$.

THEOREM 6.10: *Suppose S is compact and T is defined by a random contraction. If the transition probability P defining T satisfies 3.5, then 1 is the only proper value of T of modulus one.*

PROOF: First of all, we will say that a complex number α is a proper value of T if and only if it is a proper value of $\hat{T}$ operating on the Banach space $\hat{L}(S)$ of all complex valued Lipschitz functions on S; the operator $\hat{T}$ is defined by the equation $\hat{T}g = Tf_1 + iTf_2$ where $g = f_1 + if_2, f_1, f_2 \in L(S)$.

Now if α is a proper value of T of modulus one, [25, Theorem 5.1, p. 57] shows that there is an integer k such that $\alpha^k = 1$. Thus the argument in Theorem 3.9 shows that if, for each n, the only fixed points of T on $L(S)$ are the constant functions, then $\alpha = 1$.

Since T is quasi-compact on $L(S)$, so is T^k for every k. Hence Lemma 6.6 implies that T^k is equicontinuous for each k. The argument in Theorem 3.10 then shows that the only fixed points of T in $L(S)$ are the constant functions.

Therefore, the only proper value of T of modulus one is 1.   Q.E.D.

We close this section by extending some of these results to the case of noncompact S.

Suppose we are given a random dynamical system. Consider the following three conditions: (a) Let d be the metric on S. Then

$$\sup_{s,t} d(s,t) < \infty.$$

(b) There is a probability λ on $(E, \mathbf{E})$ such that $Q(s, A) \geqq \alpha\lambda(A)$ for some positive constant $\alpha > 0$. (c) Define, assuming that (b) holds,

$$\bar{r} \equiv \sup_{s \neq t} \int \frac{d(\theta(s,e), \theta(t,e))}{d(s,t)} \lambda(de).$$

Then $\bar{r} < 1$.

THEOREM 6.11 (Norman [25, p. 67, Theorem 1.1]): *Suppose T is defined by a random dynamical system satisfying (a), (b), and (c). Then the sequence of operators $\{T^n\}$ on $L(S)$ converges at a geometric rate to a projection whose range consists of only the constant functions.*

PROOF: Norman proves all assertions except convergence at a geometric rate. But since T is quasi-compact and the norms of T^k are uniformly bounded, the Corollary on p. 205 of [34] shows that the convergence occurs at a geometric rate.
                                                                                 Q.E.D.

COROLLARY 6.12: *Suppose in addition to the hypothesis of Theorem 6.11, that the operator T is tight. Then, for any probability μ, the sequence of probabilities $\{T^{*n}\mu\}$ converges at a geometric rate in the metric ρ to the unique invariant probability for T.*

PROOF: Theorem 6.11 asserts that $\{T^{*n}\mu\}$ converges at a geometric rate to some element of $L(S)^*$. Since T is tight, Proposition 2.8, together with the fact that ρ induces the weak topology on probabilities, implies that this limit is itself a probability.

We now show that the limiting probability (which is obviously invariant) is unique. Suppose not. Then there are at least two distinct invariant probabilities

λ_1, λ_2. Theorem 6.1, p. 40 of [27] allows us to find a uniformly continuous function g such that $\int g \, d\lambda_1 \neq \int g \, d\lambda_2$. Since $L(S)$ is sup norm dense in the Banach space of all uniformly continuous functions (Dudley [8, Lemma 8]), we may approximate g by a Lipschitz function f so that $\int f \, d\lambda_1 \neq \int f \, d\lambda_2$. But then

$$(f, \lambda_1) = (f, A_n(T^*)\lambda_1)$$

$$\neq (f, A_n(T^*)\lambda_2) = (f, \lambda_2).$$

Since $(f, A_n(T^*)\lambda_i) = (A_n(T)f, \lambda_i)$, and since $h \equiv \lim A_n(T)f$ exists in $L(S)$, we conclude $(h, \lambda_1) \neq (h, \lambda_2)$. But then h cannot be constant. Since h is an invariant function for T, this contradicts the conclusion of Theorem 6.11. Thus $\lambda_1 = \lambda_2$, and the invariant probability for T is unique.   Q.E.D.

Bell Laboratories, Murray Hill, New Jersey

APPENDIX

This Appendix contains references for the undefined mathematical terms used in this paper. For definitions of any terms omitted from this Appendix and for general mathematical background the reader should consult [1, 9, 14, 19, and 30].

Term	Reference
adjoint operator	[9, p. 478]
Banach space	[9, p. 59]
Banach space dual	[9, p. 61]
Borel σ-field (or algebra)	[14, p. 219]
continuous linear operator	[9, p. 60]
linear functional	[9, p. 38]
measure	[14, p. 31]
measurable space	[14, p. 73]
measurable set	[14, p. 73]
norm	[9, p. 59]
regular measure	[9, p. 137]
σ-field (or algebra)	[14, p. 28]
set function	[14, p. 30]

REFERENCES

[1] BARTLE, R.: *The Elements of Integration.* New York: Wiley, 1966.
[2] BHATTACHARYA, R. N., AND M. K. MAJUMDAR: "On Global Stability of Some Stochastic Economic Processes: A Synthesis," in *Quantitative Economics and Development,* ed. by L. Klein, M. Nerlove, and S. C. Tsiang. New York: Academic Press, 1980.
[3] BREIMAN, L.: *Probability.* Reading, Mass.: Addison-Wesley, 1968.
[4] BROCK, W., AND M. MAJUMDAR: "Global Asymptotic Stability Results for Multisector Models of Optimal Growth Under Uncertainty When Future Utilities are Discounted," *Journal of Economic Theory,* 18(1978), 225–243.
[5] BROCK, W., AND L. MIRMAN: "Optimal Economic Growth and Uncertainty: The Discounted Case," *Journal of Economic Theory,* 4(1972), 479–513.
[6] CROSS, J.: "A Stochastic Learning Model of Economic Behavior," *Quarterly Journal of Economics,* 87(1973), 239–266.

[7] DALEY, D. J.: "Stochastically Monotone Markov Chains," *Z. Wahrscheinlichkeitstheorie verw. Geb.*, 10(1968), 305–317.

[8] DUDLEY, R.: "Convergence of Baire Measures," *Studia Mathematica*, 27(1966), 251–268.

[9] DUNFORD, N., AND J. SCHWARTZ: *Linear Operators, Part I.* New York: Interscience, 1958.

[10] FOLEY, D., AND M. HELLWIG: "Asset Management With Trading Uncertainty," *Review of Economic Studies*, 42(1975), 327–346.

[11] FUTIA, C.: "Schumpeterian Competition," *Quarterly Journal of Economics*, June 1980, 675–695.

[12] GRANDMONT, J. M., AND W. HILDENBRAND: "Stochastic Processes of Temporary Equilibria," *Journal of Mathematical Economics*, 1(1974), 247–277.

[13] GREEN, J., AND M. MAJUMDAR: "The Nature of Stochastic Equilibria," *Econometrica*, 43(1975), 647–660.

[14] HALMOS, P. R.: *Measure Theory*. Princeton: Van Nostrand, 1950.

[15] IOSIFESCU, M., AND R. THEODORESCU: *Random Processes and Learning*. New York: Springer-Verlag, 1969.

[16] INABA, F. S.: "Consumption Loans Under Uncertainty," *Journal of Economic Theory*, 15(1977), 26–53.

[17] JAMISON, B.: "Asymptotic Behavior of Successive Interates of Continuous Functions Under a Markov Operator," *Journal of Mathematical Analysis and Applications*, 9(1964), 203–214.

[18] ———: "Ergodic Decomposition Induced by Certain Markov Operators," *Transactions of the American Mathematical Society*, 117(1965), 451–468.

[19] KELLEY, J.: *General Topology*. New York: Van Nostrand, 1955.

[20] LUCAS, R. E.: "Expectations and the Neutrality of Money," *Journal of Economic Theory*, 4(1972), 103–124.

[21] ———: "Asset Prices in an Exchange Economy," *Econometrica*, 46(1978), 1429–1445.

[22] MIRMAN, L.: "The Steady State Behavior of a Class of One Sector Growth Models With Uncertain Technology," *Journal of Economic Theory*, 6(1973), 219–242.

[23] NELSON, R. R., AND S. G. WINTER: "Factor Price Changes and Factor Substitution in an Evolutionary Model," *The Bell Journal of Economics*, 6(1975), 466–487.

[24] NEVEU, J.: *Mathematical Foundations of the Calculus of Probability*. San Francisco: Holden-Day, 1965.

[25] NORMAN, M. F.: *Markov Processes and Learning Models*. New York: Academic Press, 1972.

[26] OREY, S.: *Limit Theorems for Markov Chain Transition Probabilities*. New York: Van Nostrand, 1971.

[27] PARTHASARATHY, K.: *Probability Measures on Metric Spaces*. New York: Academic Press, 1967.

[28] REVUZ, D.: *Markov Chains*. Amsterdam: North Holland, 1975.

[29] ROBERTSON, A. P., AND W. J. ROBERTSON: *Topological Vector Spaces*. Cambridge: Cambridge University Press, 1973.

[30] ROYDEN, J.: *Real Analysis*. London: The Macmillan Co., 1968.

[31] SANGHVI, A. P., AND M. J. SOBEL: "Bayesian Games as Stochastic Processes," *International Journal of Game Theory*, 5(1976), 1–22.

[32] SCHMALENSEE, R.: "Alternative Models of Bandit Selection," *Journal of Economic Theory*, 10(1975), 333–342.

[33] SINE, R.: "Geometric Theory of a Single Markov Operator," *Pacific Journal of Mathematics*, 27(1968), 155–166.

[34] YOSIDA, K., AND S. KAKUTANI: "Operator-Theoretical Treatment of Markoff's Process and Mean Ergodic Theorem," *Annals of Mathematics*, 42(1941), 188–228.