Journal of Theoretical Probability, Vol. 12, No. 2, 1999

On the Distribution of the Limit of Products of I.I.D. 2×2 Random Stochastic Matrices

A. Mukherjea,[1,3] A. Nakassis,[2] and J. S. Ratti[1]

Received December 14, 1997; revised May 1, 1998

This article gives sufficient conditions for the limit distribution of products of i.i.d. 2×2 random stochastic matrices to be continuous singular, when the support of the distribution of the individual random matrices is finite.

KEY WORDS: Random matrices; limit distribution.

1. INTRODUCTION

In the 1960s and 1970s, Rosenblatt and others studied convergence of random walks taking values on compact semigroups [see Mukherjea and Tserpes[(4)]; Rosenblatt[(6)]], and properties of the limiting distributions in this abstract setting. However, very few concrete examples have been studied so far to illustrate these results. The set of 2×2 stochastic matrices, though an extremely simple (multiplicative) compact semigroup, is large enough to support some highly nontrivial cases that can shed light on a number of important questions in the above context. In this paper we study the question of continuous singularity of the limiting distribution, and then, as an application, consider the question whether the limiting distribution can arise from more than one random walk.

Let $A_1, A_2, \ldots, A_n$ be 2×2 stochastic matrices such that the first column of A_i is (x_i, y_i), where $0 \leqslant x_i \leqslant 1$, $0 \leqslant y_i \leqslant 1$. In this paper, the point (x_i, y_i) on the plane will always represent the matrix A_i. The second column of A_i is, of course, $(1 - x_i, 1 - y_i)$. Let μ be a probability measure

[1] Department of Mathematics, University of South Florida, Tampa, Florida 33620-5700.
[2] National Institute of Standards and Technology, Gaithersburg, Maryland 20899.
[3] To whom correspondence should be addressed.

0894-9840/99/0400-0571$16.00/0

such that the support of μ is given by $S(\mu)=\{A_1, A_2,..., A_n\}$ and $\mu(A_i)=p_i$, $1\leqslant i\leqslant n$. Let us assume that for each i, $1\leqslant i\leqslant n$, $0<x_i<1$ and $0<y_i<1$. Then it is well-known [see Rosenblatt[6]] that the sequence (μ^n) of convolution powers of μ (where μ^n is the distribution of $Y_1\cdot Y_2\cdots Y_n$, Y_i's being i.i.d. random matrices with distribution μ) converges weakly to a probability measure λ whose support consists of 2×2 stochastic matrices with identical rows, so that they are represented by points (x, x), $0\leqslant x\leqslant 1$. Let us define the function $G(x)$ by

$$G(x)=\lambda\{(y, y)\mid : y\leqslant x\}$$

Then G is the distribution function of λ, and since $\lambda * \mu=\lambda$, G satisfies the functional equation:

$$G(x)=\sum_{i\in\mathscr{A}} p_i G\left(\frac{x-y_i}{x_i-y_i}\right)+\sum_{i\in\mathscr{B}} p_i\left[1-G\left(\frac{x-y_i}{x_i-y_i}\right)\right]+\sum_{i\in\mathscr{C}[x]} p_i \tag{1.1}$$

where $\mathscr{A}\cup\mathscr{B}\cup\mathscr{C}=\{0, 1,..., n-1\}$, and the sets $\mathscr{A}$, $\mathscr{B}$, $\mathscr{C}$, and $\mathscr{C}[x]$ are given by

$$\mathscr{A}=\{i\mid x_i>y_i\}, \qquad \mathscr{B}=\{i\mid x_i<y_i\}, \qquad \mathscr{C}=\{i\mid x_i=y_i\}$$

and

$$\mathscr{C}[x]=\{i\mid x_i=y_i\leqslant x\}$$

For the purposes of this paper we will assume that $\mathscr{C}$ is empty. Indeed if $\mathscr{C}=\{1, 2,..., n-1\}$, then $\lambda=\mu$. If, on the other hand, $\mathscr{C}$ is a nonempty proper subset of $\{0,..., n-1\}$, the support of λ is enumerable. Indeed, if

$$\mathscr{D}_0=\{x\mid x=x_i \text{ for some } i\in\mathscr{C}\}, \qquad \text{and} \tag{1.2}$$

$$\mathscr{D}_{m+1}=\mathscr{D}_m\cup\left\{x\;\middle|\;\frac{x-y_i}{x_i-y_i}\in\mathscr{D}_m \text{ for some } i,\ i\in\mathscr{A}\cup\mathscr{B}\right\}, \qquad m=0, 1, 2,... \tag{1.3}$$

and if

$$P=\sum_{i\in\mathscr{C}} p_i \tag{1.4}$$

then repeated applications of Eq. (1.1) would show that the G-measure of $\mathscr{D}_m$ will be no less than

$$\sum_{0 \leqslant j \leqslant m} P(1-P)^j$$

It follows then that $\mathscr{D}$, the union of the sets $\mathscr{D}_m$ defined in Eq. (1.3) is an enumerable set of G-measure 1.

If $\mathscr{C}$ is empty, then $G(x)$ is continuous when all the points in $S(\mu)$ are not collinear with $(1, 0)$; and it is known that it is either continuous singular or absolutely continuous (with respect to the Lebesgue measure on $[0, 1]$). In Mukherjea and Ratti,[3] we found sufficient conditions for G to be continuous singular when the sets $\mathscr{B}$ and $\mathscr{C}$ are empty. In this article, we consider the case when $\mathscr{B}$ and $\mathscr{C}$ are not necessarily empty, and our results here are much more complete. The treatment of the case when $\mathscr{B}$ is non-empty is not at all obvious from our earlier considerations [Mukherjea and Ratti[3]], and thus, necessitates the present paper. The question when G is continuous singular was raised by Rosenblatt.[6]

Finally, we should mention that the problem was earlier considered by Nakassis[5] and Sun,[7] in the case when $S(\mu)$ has exactly two points (both below the diagonal). In Section 2, we present our main result. In Section 3, we present a nontrivial special case when the function G is like the classical Cantor function. The last section contains an application where we use results of Section 3.

2. CONTINUOUS SINGULARITY OF THE FUNCTION G

This section will establish our main result, will present extensions of the main result, and will construct examples in which the limit distribution is piece-wise polynomial.

2.1. Definitions and Terminology

As we pointed out in the introduction, 2×2 stochastic matrices are fully specified by their first column and can therefore be identified with the elements of the unit square, $[0, 1] \times [0, 1]$, and in what follows we will use the point notation to identify our stochastic matrices. Under this assumption,

$$(x', y')(x, y) = (x'(x-y) + y,\ y'(x-y) + y) \tag{2.1}$$

Clearly, whenever $x \neq y$, right multiplication by (x, y) defines an injective mapping. Moreover, if $x' = y'$, its image $(\bar{x}, \bar{y})$ will satisfy $\bar{x} = \bar{y}$. Thus if one were to identify the points in the main diagonal (i.e., $\{(a, a) \mid 0 \leqslant a \leqslant 1\}$) with the unit interval, right multiplication by (x, y) could be seen as a linear transformation of $[0, 1]$.

Following the customary notation, we define

Definition 1. For every mapping s, $s: A \to B$, every X, $X \subset A$, every Y, $Y \subset B$, and every b, $b \in B$, $Xs = \{xs \mid x \in X\}$, $Ys^{-1} = \{x \mid x \in A$ and $xs \in Y\}$, and $\{b\}\, s^{-1} = bs^{-1}$.

We note that:

- In the context of this paper A and B are sets of 2×2 stochastic matrices and s is some form of matrix multiplication, and
- the paper consistently identifies stochastic matrices with points in the unit square and the diagonal of the unit square with the unit interval because, eventually, the problems addressed involve sets of real numbers and linear mappings from $\mathscr{R}$ to $\mathscr{R}$.

Assume now that μ has finite support and that none of the points in its support lie on the main diagonal. For example,

$$S(\mu) = \{(x_i, y_i) : x_i \neq y_i,\ 0 \leqslant i \leqslant n-1\}$$

and $\mu\{(x_i, y_i)\} = p_i > 0$ with $p_0 + \cdots + p_{n-1} = 1$.

If at least one of the points in $S(\mu)$ is neither $(0, 1)$ nor $(1, 0)$, and if λ is as before, the support of λ will be some subset of the main diagonal. Thus, $\lambda = \lambda * \mu$, will be equivalent to

$$\lambda(B) = \sum_{i=0}^{n-1} p_i \lambda(B(x_i, y_i)^{-1}) \tag{2.2}$$

for every Borel set B such that $B \subset [0, 1]$.

We can now assume without loss of generality that $S(\mu)$ contains at least two points and that $(1, 0)$ is not in $S(\mu)$. For if $(1, 0)$ were in $S(\mu)$, then μ would be of the form

$$\mu = p\mu_1 + (1-p)\,\mu_2$$

where $0 < p < 1$, $S(\mu_1) = \{(1, 0)\}$, and $S(\mu_2) = S(\mu_\mu) - \{(0, 1)\}$.

Under these conditions, the iterates of μ would have the same limits as the iterates of μ_2.

Similarly, if $S(\mu)$ consisted of a single point, the support of λ would collapse to a single point.

If the support of μ does not include either $(1, 0)$ or points in the diagonal, then the mappings $(x, x) \to (x, x)(x_i, y_i)$ are linear and invertible with fixed points that lie on the intersection of the main diagonal and the line spanned by $(1, 0)$ and (x_i, y_i). Moreover, for all points other than $(0, 1)$ these mappings are contractions.

In what follows we will assume that the support of μ is finite, has at least two points, and does not contain any point on the main diagonal of the unit square or point $(1, 0)$. In addition we will use the following symbols and definitions:

Definition 2.

- T_i is the linear mapping that (x_i, y_i), $i = 0, ..., n-1$, induces over the real numbers;
- $u_i = x_i - y_i$ and $a_i = |x_i - y_i|$, $i = 0, ..., n-1$;
- f_i is the fixed point of T_i;
- $\mathfrak{F}$ is the free semigroup generated by the symbols $\{T_0, ..., T_{n-1}\}$. Note: $\mathfrak{F}$ can be homomorphically embedded in the semigroup generated by the functions $\{T_0, ..., T_{n-1}\}$; thus, when context demands it, each s in $\mathfrak{F}$ will be treated as a function.
- $\mathfrak{F}_k$ is the set of all elements in $\mathfrak{F}$ of length k;
- $[t_*, t^*]$ is the smallest closed interval whose λ-measure is one.

We note that:

- for each x and i, $xT_i = (x - f_i)\, u_i + f_i$
- $[t_*, t^*]$ depends only on the points in μ's support. It is the smallest interval that satisfies

$$[a, b]\, T_i \subset [a, b] \qquad \text{for all } i, \quad i \in \{0, ..., n-1\} \tag{2.3}$$

Equivalently, t_* and t^* are, respectively, the maximum and minimum values that simultaneously satisfy:

$$t_* \leqslant f_i \leqslant t^* \qquad \text{for all } i, \quad i = 0, ..., n-1 \tag{2.4}$$

and

$$t^* T_i^{-1} \leqslant t_* \leqslant t^* \leqslant t_* T_i^{-1} \qquad \text{for all } i, \quad i \in \mathscr{B} \tag{2.5}$$

As a rule, $[t_*, t^*]$, is a proper subset of $[0, 1]$. Nevertheless, the order-preserving linear transformation that maps $[t_*, t^*]$ onto $[0, 1]$ preserves the functional form of Eq. (2.2) as well as the u_i and a_i values, $i = 0, \ldots, n-1$. We can therefore assume that Eq. (2.2) has been normalized through the transform in question, and will rewrite it as

$$\lambda(B) = \sum_{i=1}^{n} p_i \lambda(BT_i^{-1}) \tag{2.6}$$

We shall assume therefore in what follows that unless otherwise indicated $[t_*, t^*] = [0, 1]$ and that, unless otherwise qualified, symbol I represents $[0, 1]$.

2.2. Main Result

Theorem 1. The solution of Eq. (2.6) is continuous singular (with respect to the Lebesgue measure on $[0, 1]$) if any of the following conditions holds:

(i) $a_0 + a_1 + \cdots + a_{n-1} < 1$;

(ii) $a_0 + a_1 + \cdots + a_{n-1} = 1$, and for some i, $p_i \neq a_i$;

(iii) $(a_0/p_0)^{p_0} (a_1/p_1)^{p_1} \cdots (a_{n-1}/p_{n-1})^{p_{n-1}} < 1$.

Remark 1. When either conditions (i) or (ii) earlier hold true, condition (iii) holds true as well. Indeed, if the p_i's are positive constants, and the a_i's nonnegative with constant sum T, then the maximum of

$$a_0^{p_0} a_1^{p_1} \cdots a_{n-1}^{p_{n-1}}$$

occurs at the single point where

$$\frac{a_0}{p_0} = \frac{a_1}{p_1} = \cdots = \frac{a_{n-1}}{p_{n-1}}$$

Proof. By iterating Eq. (2.6) k times we obtain:

$$\lambda(B) = \sum_{s \in \mathfrak{F}_k} \left(\prod_{i=1}^{n} p_i^{k_i(s)} \right) \lambda(Bs^{-1}) \tag{2.7}$$

where $k_i(s)$ is the number of times T_i occurs in the word s.

Therefore, if B is chosen to be Is, we obtain:

$$\lambda(Is) \geqslant \prod_{i=1}^{n} p_i^{k_i(s)} \tag{2.8}$$

while the length of Is equals

$$l(Is) = \prod_{i=1}^{n} a_i^{k_i(s)} \tag{2.9}$$

Assume now that ε and d are positive constants such that

$$\left(\frac{a_0}{p_0}\right)^{z_0} \left(\frac{a_1}{p_1}\right)^{z_1} \cdots \left(\frac{a_{n-1}}{p_{n-1}}\right)^{z_{n-1}} < d < 1$$

for all $(z_0, z_1, ..., z_{n-1})$ that satisfy

$$-\varepsilon < z_i - p_i < \varepsilon \qquad \text{for} \quad i = 0, 1, ..., n-1$$

Such constants do exist under our assumptions. Let us define the set $\mathfrak{F}_k^*$ by

$$\mathfrak{F}_k^* = \{s \in \mathfrak{F}_k : k(p_i - \varepsilon) < k_i < k(p_i + \varepsilon),\ 1 \leqslant i \leqslant n\}$$

Then we have:

$$\lambda\left(\bigcup_{s \in \mathfrak{F}_k^*} Is\right) \geqslant \sum_{s \in \mathfrak{F}_k^*} \prod_{i=0}^{n-1} p_i^{k_i(s)} = \sum \binom{k}{k_0, ..., k_{n-1}} \prod_{i=0}^{n-1} p_i^{k_i} \tag{2.10}$$

when the last summation is over all $(k_0, ..., k_{n-1})$ such that $k(p_i - \varepsilon) < k_i < k(p_i + \varepsilon)$, $1 \leqslant i \leqslant n-1$, and $k_0 + \cdots + k_{n-1} = k$.

By virtue of the central limit theorem, if ε is kept constant and $k \to +\infty$, this sum tends to 1.

Now let us compute an upper bound for $l(\bigcup_{s \in \mathfrak{F}_k^*} Is)$.

We observe that since for every s

$$\frac{l(Is)}{\lambda(Is)} = \prod_{i=0}^{i=n-1} \left(\frac{a_i}{p_i}\right)^{k_i}$$

for all s, $s \in \mathfrak{F}_k^*$, the right hand term in Eq. (2.2) is bounded by d^k.

Given that the sum in Eq. (2.10) cannot exceed 1, it ensues that

$$l\left(\bigcup_{s \in \mathfrak{F}_k^*} Is\right) < d^k \qquad \square$$

Remark 2. There are situations when we have both

$$S = a_0 + \cdots + a_{n-1} > 1 \qquad \text{and} \qquad P = \left(\frac{a_0}{p_0}\right)^{p_0} \cdots \left(\frac{a_{n-1}}{p_{n-1}}\right)^{p_{n-1}} < 1 \tag{2.11}$$

Indeed, assume that we start with a set of a_i's that are not identical to the p_i's and sum up to $S = 1$. Then, $P < 1$. Since P is continuous in its arguments, it follows that small increases in the a_i's will result in parameters that satisfy Eq. (2.11).

2.3. Extensions of the Main Theorem

As in Nakassis[5] for $n = 2$, we can show that λ is continuous singular when

$$\left(\frac{a_0}{p_0}\right)^{p_0} \left(\frac{a_1}{p_1}\right)^{p_1} \cdots \left(\frac{a_{n-1}}{p_{n-1}}\right)^{p_{n-1}} = 1$$

provided that not all ratios a_i/p_i equal 1. The purpose of this subsection is to show that this proof can be extended to cover all n.

Under the circumstances at least one of the ratios is less than 1 and at least one exceeds 1. Let L consist of all indices for which a_i/p_i is minimum and let g be maximum value of $(a_i/p_i)/(a_j/p_j)$ when i ranges over L and j over $H = \{0,..., n-1\} - L$. Clearly, $g < 1$. Assume that $t(k)$ is any unbounded increasing sequence that is $o(\sqrt{k})$ and note that the variance of the number of successes in k Bernoulli experiments is of the same order as $\sqrt{k}$.

Let us define $\mathfrak{F}_k$ as in the previous section and let

$$\mathfrak{F}_k^{\bullet} = \{s : s \in \mathfrak{F}_k \text{ and } k_i \leqslant kp_i - t(k) \text{ for all } i \text{ in } H \text{ and } k_0 + \cdots + k_{n-1} = k\}$$

For the same reasons as in the previous subsection we have:

$$\lambda\left(\bigcup_{s \in \mathfrak{F}_k^{\bullet}} Is\right) \geqslant \sum_{s \in \mathfrak{F}_k^{\bullet}} \prod_{i=0}^{n-1} p_i^{k_i(s)} = \sum \binom{k}{k_0, ..., k_{n-1}} \prod_{i=0}^{n-1} p_i^{k_i} \tag{2.12}$$

and

$$l\left(\bigcup_{s \in \mathfrak{F}_k^{\bullet}} Is\right) \leqslant \sum \binom{k}{k_0, ..., k_{n-1}} \prod_{i=0}^{n-1} a_i^{k_i} \tag{2.13}$$

where the summation is as in Eq. (2.12).

We note that if s is in $\mathfrak{F}_k^*$, then we have:

$$\frac{\prod_{i=0}^{n-1} a_i^{k_i}}{\prod_{i=1}^{n} p_i^{k_i}} = \prod_{i=0}^{n-1} \left(\frac{a_i}{p_i}\right)^{k_i} = \left(\prod_{i=0}^{n-1} \left(\frac{a_i}{p_i}\right)^{kp_i}\right)\left(\prod_{i=0}^{n-1} \left(\frac{a_i}{p_i}\right)^{z_i}\right) = \prod_{i=0}^{n-1} \left(\frac{a_i}{p_i}\right)^{z_i} \tag{2.14}$$

where by definition $z_i = k_i - kp_i$, $0 \leqslant i \leqslant n-1$ and by virtue of our assumptions $z_i \leqslant -t(k)$, $i \in H$, and $z_0 + \cdots + z_{n-1} = 0$.

It ensues that the expression in Eq. (2.14) is bounded from earlier by $g^{t(k)}$. Thus, we have a sequence of sets whose G-measure approaches a positive constant (Central Limit Theorem) while the (Lebesgue measure)/(G-measure) ratio converges to 0.

2.4. Piece-Wise Polynomial Solutions

As shown by Nakassis,[5] for $n=2$, appropriate choices of the parameters can lead to solutions G which are piece-wise polynomial. This subsection will show how these results can be extended to cover n-point support measures.

The cited reference shows that if the equation to solve is of the form

$$G(x) = (1/2)\, G(ax) + (1/2)\, G(a(x-1)+1) \tag{2.15}$$

with a such that $a^m = 2$, then, $G(x)$ is piece-wise polynomial of degree not exceeding m.

We note that:

- Equation (2.15) can be thought of as a starting point for constructing n-term equations as in Eq. (2.6) that have continuous solutions. It suffices to apply it iteratively and selectively (e.g., express $G(ax)$ as $(1/2)\, G(a^2x) + (1/2)\, G(a(ax-1)+1)$.
- The solution of Eq. (2.15) is symmetric around (1/2) and satisfies $G(x) = 1 - G(1-x)$ for all x. Thus one can use this type of substitution to obtain equations of the same form as Eq. (2.6) that result from points both above and below the main diagonal and which admit piece-wise polynomial solutions.
- The fact that all these apparently different equations admit the same solution points out some of the difficulties in establishing under what conditions the iterates of two different measures μ converge to the same measure.

The $n=2$ case results can be extended to measures of finite support as follows:

Proposition 1. If all the p_i values are equal to $(1/n)$, if all the u_i values are positive and equal to the mth root of $(1/n)$, and if the fixed i points f_i are equally spaced (i.e., $f_i = i/(n-1)$ for all i, $i = 0,..., n-1$), then $G(x)$ is piece-wise polynomial and of degree not exceeding m.

Proof. Let $c = (c_0, c_1,..., c_{m-1})$ represent a point in $\{0, 1\}^m$, let t be the positive mth root of n, and let $T = 1 + t + \cdots + t^{m-1}$. Let us define in Definition 4.

Definition 4.

- $w(c) = (-1)^S$ where S counts the nonzero terms in c;
- $N(c) = (c_0 + c_1 t + \cdots + c_{m-1} t^{m-1})/T$;
- $h(x) = \sum_c w(c)(x - N(c))^{m-2} |x - N(c)|$.

We observe that for $j = 0, 1,..., m-1$,

$$\sum_c N(c)^j = 0$$

Indeed, terms of the form $t^{i_0} t^{i_1} \cdots t^{i_{j-1}}$ occur in all c for which $c_{i_0} = c_{i_1} = \cdots = c_{i_{j-1}} = 1$. There are 2^{m-j} such c and for exactly half of them $w(c) = 1$ while for the other half $w(c) = -1$.

It ensues that if $x < 0$ or $x > 1$ and one can dispense with the absolute values, $h(x) = 0$.

We also observe that $h(x)$ satisfies

$$h(x) = \sum_{i = 0, 1,..., n-1} \frac{p_i}{u_i} h(xT_i^{-1}) \tag{2.16}$$

Indeed, fix for the moment all entries in c except for c_0 so that c can take exactly two values $c^0 = (0, c_1,..., c_{m-1})$ and $c^1 = (1, c_1,..., c_{m-1})$. Straightforward manipulations show that for all c we have:

If $X(i, c) = \dfrac{t}{n} (t(x - f_i) + f_i - N(c))^{m-2} |t(x - f_i) + f_i - N(c)|$

then $X(i, c) = (x - M(i, c))^{m-2} |x - M(i, c)|$

where $M(i, c) = \dfrac{f_i(t-1)}{n-1} + \dfrac{N(c)}{t} = \dfrac{1}{T}\left(\dfrac{c_0 + i}{t} + c_1 + c_2 t + \cdots + c_{m-1} t^{m-2}\right)$

But, $M(i, c^1) = M(i+1, c^0)$ for all i, $i = 0, 1, \ldots, n-2$ while $w(c^0) + w(c^1) = 0$. In short, when all terms $w(c)\, X(i, c)$ under consideration are added, all but two will cancel out and their sum will collapse to

$$w(c^0)\, X(0, c^0) + w(c^1)\, X(n-1, c^1)$$

But the first term corresponds to the summand of $h(x)$ that one would obtain if c were rotated one place to the left while the second similarly corresponds to the summand obtained through single rotation to the left because $[1 + (n-1)]/t = t^{m-1}$. It ensues that $h(x)$ indeed satisfies Eq. (2.16) and that its integral, $H(x)$, is bounded, is not identically 0, is constant outside $[0, 1]$, and satisfies Eq. (2.6). Under these conditions, $H(x)$ and $G(x)$ differ only by a multiplicative constant and, therefore, $G(x)$ is piece-wise polynomial and of degree at most m. □

3. GEOMETRIC CONSIDERATIONS

Conditions (i) and (ii) of the main theorem show that in some instances we can deduce the singularity of G, with little or no information other than the support of μ. This section will examine the case in which the set IT_i do not overlap (i.e., each point in $(0, 1)$ belongs to at most one IT_i). We will also assume that $(0, 1)$ is not in the support of μ so that all T_i's are centered around xs^{-1}, it ensues that if xs^{-1} were within $[0, 1]$, $(x-e, x+e)s^{-1}$ would contain $[0, 1]$ and its G-measure would be positive. Hence, the G-measure of $(x-e, x+e)$ would have been positive, a contradiction. Therefore, if $0 < x < 1$ and $G(x-e) = G(x+e)$ for some positive e, than $x \in O_k$ for some well chosen k.

The sixth is an immediate consequence of the fifth and of the observation that some points such as $\{f_i \mid i = 0, \ldots, n-1\}$ are not members of O_k for any k and, hence are not in O^*.

We note that:

- The one-step descendants of $\bar{I}s$, $\{\bar{I}T_i s \mid i = 0, \ldots, n-1\}$, can be obtained through the removal from $\bar{I}s$ of the open set Os that is in the same relative position to $\bar{I}s$ as O is to $[0, 1]$. Since the T_i's are not necessarily order preserving, it is entirely conceivable that s maps 1 onto the lower endpoints of $\bar{I}s$ and 0 onto the upper one. In this instance, when we traverse $\bar{I}s$ from right to left we see Os to be in the same relative position with respect to $\bar{I}s$ as O is with respect to $\bar{I}$.
- The support of G is C^* and is found through a Cantor type construction.

- While the O_k's and C_k's are linked to the initial measure μ, their limits O^* and C^* are fully determined from the limit of μ^n, λ.

4. AN APPLICATION

In this section, we will exploit the results in Section 3 to establish a result on the one-to-one correspondence between probability measures on 2×2 stochastic matrices with n-point support and the weak limits of their convolution powers. Briefly speaking, we consider the following problem:

Let μ and $\bar{\mu}$ be probability measures such that

$$S(\mu) = \{(x_i, y_i): 0 \leqslant i < n, x_i \neq y_i\}$$

and

$$S(\bar{\mu}) = \{(\bar{x}_i, \bar{y}_i): 0 \leqslant i < n, \bar{x}_i \neq \bar{y}_i\}$$

such that $\mu^k \to \lambda$, $\bar{\mu}^k \to \bar{\lambda}$ weakly as $k \to \infty$. Then under what conditions on μ and $\bar{\mu}$, does $\lambda = \bar{\lambda}$ imply $\mu = \bar{\mu}$? Besides being a very natural and also difficult problem in the general context, this problem has implications in the theory of attractors and iterated function systems. For details, see Dhar *et al.*[2]

Theorem 2. Let μ and $\bar{\mu}$ be two probability measures on 2×2 stochastic matrices such that

$$S(\mu) = \{A_i(x_i, y_i): 0 \leqslant i < n, x_i \neq y_i\} \qquad \text{and}$$
$$S(\bar{\mu}) = \{\bar{A}_i(\bar{x}_i, \bar{y}_i): 0 \leqslant i < n, \bar{x}_i \neq \bar{y}_i\}$$

such that:

- μ and $\bar{\mu}$ are strongly separated;
- For each i, the transforms T_i and $\bar{T}_i$ defined by A_i and $\bar{A}_i$ have the same fixed point and
- For each i, the linear transforms induced by A_i and $\bar{A}_i$, T_i and $\bar{T}_i$, are either both order preserving or both order reversing contractions (i.e., u_i and $\bar{u}_i$ are of the same sign and none equals -1 [otherwise said, $(0, 1)$ is not in the support of either measure]).

Then, the weak limits of μ^n and $\bar{\mu}^n$ are equal if and only if $\mu = \bar{\mu}$.

Proof. For the reasons pointed out by Dhar *et al.*[2] [Prop. 3.4] if μ and $\bar{\mu}$ have the same limit λ, then

$$O^* = \bar{O}^* \quad \text{and} \quad C^* = \bar{C}^*$$

Let O^M be a connected component of O^* of maximum length. Given that the components of O_{i+1} are obtained through contraction mappings of the components of O_i, O^M must belong to both O_0 and $\bar{O}_0$. Let $I(T_i) = (\alpha_i, \beta_i)$ and $I(\bar{T}_j) = (\bar{\alpha}_j, \bar{\beta}_j)$ be the elements of C_1 and $\bar{C}_1$ that about O^M. One of them must contain the fixed point (f_i or $\bar{f}_j$) of the other. Since μ and $\bar{\mu}$ are strongly separated, this implies that $i = j$. But then, the linear transforms T_i and $\bar{T}_i$ agree on two points (same fixed point and they both map either 0 or 1 onto β_i) and must, therefore, be equal. If $i > 0$, let O^L and $\bar{O}^L$ be the connected components of O and $\bar{O}$ that end at α_i. They intersect and since the components of O^* cannot intersect without being equal, in ensues that $O^L = \bar{O}^L$ and that $\beta_{i-1} = \bar{\beta}_{i-1}$. Thus we can prove that for $j = 0, \ldots, i$, $T_j = \bar{T}_j$.

This technique applies, just as well, on indices $i + 1$ through $n - 1$, so that $T_i = \bar{T}_i$ for $i = 0, \ldots, n - 1$.

Once this is established, examination of the probabilities λ assigns to the intervals IT_i shows that $p_i = \bar{p}_i$, $i = 0, \ldots, n - 1$. Thus, $\mu = \bar{\mu}$. □

REFERENCES

1. Anderson, T. W. (1984). *An Introduction to Multivariate Statistical Analysis*, Second Edition, Wiley.
2. Dhar, S., Mukherjea, A., and Ratti, J. S. (1999). A nonlinear functional equation arising from the convolution iterates of a probability measure on 2 × 2 stochastic matrices. *J. Nonlinear Analysis* **36** (2), 151–176.
3. Mukherjea, A., and Ratti, J. S. (1997). Continuous singularity of the weak limit of the convolution powers of a discrete probability measure on 2 × 2 stochastic matrices. *J. Theor. Prob.* **10** (2), 499–506.
4. Mukherjea, A., and Tserpes, N. A. (1976). *Measures on Topological Semigroups*, Springer LNM 547.
5. Nakassis, A. (1979). Limit behavior of the convolution iterates of a probability measure on a semigroup of matrices. *J. Math. Anal. Appl.* **70**, 337–347.
6. Rosenblatt, M. (1971). *Markov Processes: Structure and Asymptotic Behavior*, Springer.
7. Sun, T. C. (1975). Limits of convolutions of probability measures on the set of 2 × 2 stochastic matrices. *Bull. Inst. Math.* (Acad. Sinica) **3** (2), 235–248.