Proc. Indian Acad. Sci. (Math. Sci.) Vol. 124, No. 4, November 2014, pp. 603–612.

Limit distributions of random walks on stochastic matrices

SANTANU CHAKRABORTY[1] and ARUNAVA MUKHERJEA[2]

[1]Department of Mathematics, University of Texas – Pan American,
1201 West University Drive, Edinburg, Texas 78539, USA
[2]1216 La Brad Lane, Tampa, Florida 33613, USA
E-mail: arunava.mukherjea@gmail.com

MS received 18 May 2012; revised 18 December 2013

Abstract. Problems similar to *Ann. Prob.* **22** (1994) 424–430 and *J. Appl. Prob.* **23** (1986) 1019–1024 are considered here. The limit distribution of the sequence $X_n X_{n-1} \cdots X_1$, where $(X_n)_{n\geq 1}$ is a sequence of i.i.d. 2×2 stochastic matrices with each X_n distributed as μ, is identified here in a number of discrete situations. A general method is presented and it covers the cases when the random components C_n and D_n (not necessarily independent), $(C_n,\ D_n)$ being the first column of X_n, have the same (or different) Bernoulli distributions. Thus $(C_n,\ D_n)$ is valued in $\{0,\ r\}^2$, where r is a positive real number. If for a given positive real r, with $0 < r \leq \frac{1}{2}$, $r^{-1}C_n$ and $r^{-1}D_n$ are each Bernoulli with parameters p_1 and p_2 respectively, $0 < p_1,\ p_2 < 1$ (which means $C_n \sim p_1\delta_{\{r\}} + (1 - p_1)\delta_{\{0\}}$ and $D_n \sim p_2\delta_{\{r\}} + (1 - p_2)\delta_{\{0\}}$), then it is well known that the weak limit λ of the sequence μ^n exists whose support is contained in the set of all 2×2 rank one stochastic matrices. We show that $S(\lambda)$, the support of λ, consists of the end points of a countable number of disjoint open intervals and we have calculated the λ-measure of each such point. To the best of our knowledge, these results are new.

Keywords. Random walk; stochastic matrices; limiting measure.

2000 Mathematic Subject Classification. 60B10.

1. Introduction

In [1], it was proven that if $(X_n)_{n\geq 1}$ is a sequence of $d \times d$ i.i.d. stochastic matrices such that $P(\min_{i,j}(X_1)_{ij} = 0) < 1$, then $Y = \lim_{n\to\infty} X_n X_{n-1} \cdots X_1$ exists almost surely and $P(Y$ has rank $1) = 1$; furthermore, if for any Borel B of $d \times d$ stochastic matrices (with usual R^{d^2}-topology), we denote $\mu(B) = P(X_1 \in B)$ and $\lambda(B) = P(Y \in B)$, then λ is the unique solution of the convolution equation $\lambda \star \mu = \lambda$. Let us quickly note here that this wonderful result of Chamayou and Letac also holds under the (slightly weaker) condition that $\mu^m(\mathbb{P}) > 0$ for some positive integer m (as opposed to just 1, instead of m, considered in [1]), where μ^m is the distribution of the product $X_m \cdots X_1$ and $\mathbb{P}$ is the set of $d\times d$ strictly positive stochastic matrices. The reason is as follows: the Chamayou and Letac result shows that under the weaker condition, the subsequence $Y_{nm} = X_{nm}X_{nm-1} \cdots X_1$ converges almost surely to some $d \times d$ rank one stochastic matrix, Y_0, and consequently, any subsequence $X_{n_k} X_{n_k-1} \cdots X_1$ with $n_k > s_k m$ (for some s_k), will also converge almost surely to a $d \times d$ stochastic matrix $VY_0(=Y_0$, as Y_0 has rank one), where V is a limit point of the product subsequence $X_{n_k} X_{n_k-1} \cdots X_{s_k m+1}$. This establishes our observation.

In the same paper, Chamayou and Letac (see also [4]) tried to identify λ in the case when the rows of X_1 above are independent, and for $1 \le i \le d$, the i-th row of X_1 has Dirichlet distribution with positive parameters $\alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{id}$, and they were successful in the case when $\sum_{j=1}^{d} \alpha_{ij} = \sum_{j=1}^{d} \alpha_{ji}$, $1 \le i \le d$. Indeed, there are only very few (other than those given in [1, 2, 4]) examples in the literature even for 2×2 stochastic matrices when the limit distribution λ has been identified completely in the above context.

In this paper, we consider 2×2 i.i.d. stochastic matrices $(X_n)_{n \ge 1}$ with $X_n = \begin{pmatrix} C_n & 1 - C_n \\ D_n & 1 - D_n \end{pmatrix}$, such that each X_n is distributed as μ and we subsequently identify the distribution λ, the distribution of $\lim_{n\to\infty} X_n X_{n-1} \cdots X_1$ in the case when $r^{-1}C_n$ and $r^{-1}D_n$, r being a positive real number satisfying $0 < r \le \frac{1}{2}$, are each Bernoulli (but with possibly different parameters p_1 and p_2, $0 < p_1, p_2 < 1$). Here C_n and D_n are not necessarily independent. As far as we know, our results and methods are all new.

What we already know is the following: When λ is the weak limit of $(\mu^n)_{n \ge 1}$ and S contains a rank one matrix, then the support of λ, $S(\lambda)$ consists of all rank one stochastic matrices in $S = \overline{\cup_{n=1}^{\infty} S(\mu^n)}$, where

$$S(\mu^n) = \overline{\{A_1 A_2 \cdots A_n \mid \text{for each } i,\ A_i \in S(\mu),\ 1 \le i \le n\}}$$

and n is a positive integer. This is an algebraic fact for the support of an idempotent probability measure (note that $\lambda = \lambda \star \lambda$, see [3]). The results we present here are complete though, as will be seen, computationally somewhat complicated. Our methods are necessarily different from those used by Chamayou and Letac and also by Van Assche. Using our methods, one can solve this problem completely for other discrete distributions on $d \times d$ stochastic matrices X with $d > 2$. However, computations are expected to be challenging. We believe that this problem of identifying the limit distribution is important because products of i.i.d. random matrices have been studied in numerous different contexts. In what follows, a strictly positive matrix will mean a matrix with each entry positive.

Let $(X_i)_{i \ge 1}$, as before, be i.i.d. $d \times d$ stochastic matrices such that for some positive integer $m \ge 1$,

$$\mu^m(\mathbb{P}) > 0 \tag{1}$$

(recall that $\mathbb{P}$ is the set of $d \times d$ strictly positive stochastic matrices in S). Then the sequence $(\mu^n)_{n \ge 1}$, where $\mu(B) = P(X_1 \in B)$ for Borel sets B of $d \times d$ stochastic matrices, converges weakly to a probability measure λ and $S(\lambda)$ consists of all rank one stochastic matrices in $S = \overline{\cup_{n=1}^{\infty} S(\mu^n)}$ such that $\lambda(\mathbb{P}) > 0$.

Let us prove this result.

By Theorem 2.7(i), page 87 in [3], it follows that $\frac{1}{n} \sum_{i=1}^{n} \mu^i$ converges weakly to a probability measure λ such that $\lambda = \lambda \star \lambda = \lambda \star \mu = \mu \star \lambda$. Then, by our assumption, $\mu^m(\mathbb{P}) > 0$ for some positive integer m. Since $\mathbb{P} S \mathbb{P} \subset \mathbb{P}$ and $\mu^m \star \lambda \star \mu^m = \lambda$, it follows that $\lambda(\mathbb{P}) \ge \mu^m(\mathbb{P}) \lambda(S) \mu^m(\mathbb{P}) > 0$. Since $\mathbb{P}$ is an open subset of S, $S(\lambda) \cap \mathbb{P} \ne \emptyset$. Let $x \in S(\lambda) \cap \mathbb{P}$. Then, $x S(\lambda) x \subset \mathbb{P}$. By Theorem 2.2, page 74 of [3], $x S(\lambda) x$ is a compact group of strictly positive matrices and consequently, by Corollary 1.8 in [3], $x S(\lambda) x$ is a single element e $(= e^2)$ in $\mathbb{P}$. Since e is idempotent and strictly positive, e must have rank one. Then it follows from Theorem 2.2 and Proposition 1.9 in [3] that $S(\lambda)$ consists of all rank one matrices in S. Since $e S(\lambda) e = x S(\lambda) x$ is a single idempotent element, it follows by Theorem 2.7(iii) in [1] that μ^n converges weakly to λ. The proof is complete.

Let us also mention that under condition (1), if S_μ (and consequently, S itself) consists of only $d \times d$ bistochastic matrices, then S_λ must be a singleton since there is only one rank one $d \times d$ bistochastic matrix.

Thus, for $d = 2$, λ is the unique mass at $\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$.

From now on, we will often denote the matrix $\begin{pmatrix} x & 1-x \\ x & 1-x \end{pmatrix}$ by simply x, when there is no fear of confusion. Thus, for the limiting measure λ, $\lambda(x)$ will mean $\lambda \begin{pmatrix} x & 1-x \\ x & 1-x \end{pmatrix}$ and if we write that the support of λ, $S(\lambda)$ is contained in [0, 1], then this means the following:

$$S(\lambda) \subset \left\{ \begin{pmatrix} x & 1-x \\ x & 1-x \end{pmatrix} : 0 \le x \le 1 \right\}.$$

In §2, we state and prove our main results omitting the details which can be easily worked out by the reader. And in §3, we present alternative proofs of these results. We would like to point out here that we covered the cases $0 < r \le \frac{1}{2}$ and $r = 1$. Note that the case $\frac{1}{2} < r < 1$ is not considered here and left out for future consideration.

2. The main results

Consider 2×2 i.i.d. stochastic matrices $(X_n)_{n \ge 1}$ with $X_n = \begin{pmatrix} C_n & 1-C_n \\ D_n & 1-D_n \end{pmatrix}$, such that each X_n is distributed as μ. Also, assume that for a given r with $r = 1$ or $0 < r \le \frac{1}{2}$, both $r^{-1}C_n$ and $r^{-1}D_n$ are Bernoulli with parameters p_1 and p_2 respectively. Then, it is clear that the support of μ, $S(\mu)$ is given by

$$S(\mu) = \left\{ \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ r & 1-r \end{pmatrix}, \begin{pmatrix} r & 1-r \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} r & 1-r \\ r & 1-r \end{pmatrix} \right\}.$$

Let the μ-masses at these points be denoted by $p_{00},\ p_{01},\ p_{10},\ p_{11}$ respectively so that $p_{00} + p_{01} = q_1$, $p_{00} + p_{10} = q_2$, $p_{10} + p_{11} = p_1$ and $p_{01} + p_{11} = p_2$, where $q_i = 1 - p_i$ for $i = 1, 2$.

Let λ be the distribution of $\lim_{n \to \infty} X_n X_{n-1} \cdots X_1$. For $r = 1$, one can easily observe that λ follows a Bernoulli distribution with parameter entirely dependent on the probability mass function of μ, namely,

$$\lambda(0) = \frac{p_{00}(1 - p_{10}) + p_{11} p_{01}}{(1 - p_{10})^2 - p_{01}^2}.$$

For $0 < r \le \frac{1}{2}$, the support of μ^n, $S(\mu^n)$ and consequently S is contained in the set

$$\left\{ \begin{pmatrix} x & 1-x \\ y & 1-y \end{pmatrix} : 0 \le x \le r,\ 0 \le y \le r \right\}.$$

Also, it is known that the relation $\lambda \star \mu = \lambda$ holds and the support of λ, namely, $S(\lambda)$ consists of all rank one matrices in S. As a result,

$$S(\lambda) \subset \{x : 0 \leq x \leq r\}, \quad \text{recall that } x \text{ stands for} \begin{pmatrix} x & 1-x \\ x & 1-x \end{pmatrix}.$$

Moreover, exploiting the identity $\lambda \star \mu = \lambda$, we have

$$\lambda(0) = \frac{p_{00}}{1-p_{10}}, \quad \lambda(r) = p_{11} + \lambda(0) p_{01} = \frac{p_{11}(1-p_{10}) + p_{00}p_{01}}{1-p_{10}}$$

and for other points x with $0 < x < r$ with positive λ-masses, we have

$$\lambda(x) = \lambda(r^{-1}x)p_{10} + \lambda(1-r^{-1}x)p_{01}. \tag{2}$$

For further details on the nature of λ, we need two propositions (Propositions 2.1 and 2.2) for taking care of the cases $0 < r < \frac{1}{2}$ and $r = \frac{1}{2}$.

PROPOSITION 2.1

For $0 < r < \frac{1}{2}$, *we have the following*:

(i) *For every positive integer* i, *there are exactly* 2^{i-1} *points that have positive* λ*-masses which are polynomials in* r. *These polynomials are of the form* $\sum_{j=1}^{k}(-1)^{j-1}r^{i_j}$ *for* $1 \leq i_1 < i_2 < i_3 < \cdots < i_k = i$ *for some* $k \leq i$.

(ii) *Each such point has* λ*-measure equal to* $\lambda(r)\, p_{10}^{i-1-k} p_{01}^{k}$. *For every* $i > 1$, *the sum of the* λ*-masses of all* 2^{i-1} *points that have positive* λ*-masses equals* $\lambda(r)\, [p_{10} + p_{01}]^{i-1}$.

(iii) *The sum of the* λ*-masses of all the polynomials in* r *of all finite degrees in* $[0, r]$ *with positive* λ*-masses along with the* λ*-mass at zero equals* 1.

Proof.

Part (*i*): We start with the interval $[0, r]$ where the two points 0 and r have positive λ-masses. So, r is a polynomial in r of degree 1 with positive λ-mass. Starting with degree 1, we generate polynomials in r of higher degree with positive λ-masses by making use of the identity (2). It is observed that one can obtain two points with positive λ-masses which are polynomials in r of degree 2, namely, r^2 and $r - r^2$. Similarly, there are four points with positive λ-masses which are polynomials in r of degree 3, namely, $r^3, r^2 - r^3, r - r^2 + r^3$ and $r - r^3$. Continuing like this, for $i > 1$, there are 2^{i-1} polynomials in r of degree i with positive λ-masses in $[0, r]$.

The 2^{i-1} polynomials of degree i with positive λ-masses along with the previous $\sum_{l=1}^{i-2} 2^l$ polynomials of less degrees partition the whole interval $[0, r]$ into $2^i - 1$ intervals of the form $[a_1^{(i)}, b_1^{(i)}], (b_1^{(i)}, a_2^{(i)}), [a_2^{(i)}, b_2^{(i)}], (b_2^{(i)}, a_3^{(i)}), [a_3^{(i)}, b_3^{(i)}], \ldots, [a_{2^{i-1}-1}^{(i)}, b_{2^{i-1}-1}^{(i)}], (b_{2^{i-1}-1}^{(i)}, a_{2^i}^{(i)}), [a_{2^{i-1}}^{(i)}, b_{2^{i-1}}^{(i)}]$ where the 2^{i-1} polynomials of degree i with positive λ-masses in $[0, r]$ are $b_1^{(i)}, a_2^{(i)}, b_3^{(i)}, a_4^{(i)}, \ldots, b_{2^{i-1}-1}^{(i)}, a_{2^{i-1}}^{(i)}$. Other $a_j^{(i)}$ s and $b_j^{(i)}$ s are polynomials of degree less than i. Using (2), it follows that for $i > 1$, each $(b_j^{(i)}, a_{j+1}^{(i)})$ has λ-measure zero for $j = 1, 2, \ldots, 2^{i-1} - 1$ and the other 2^{i-1} intervals at stage i have nonzero probability.

We also observe that the 2^{i-1} polynomials of degree i with positive λ-masses are related to the 2^{i-2} intervals of nonzero probability at the previous stage by adding r^i to the left end points of each of these 2^{i-2} intervals and subtracting r^i from the right end points of each of these 2^{i-2} intervals. Thus, it follows that, for each such point at the i-th stage, $b_{2l-1}^{(i)} = a_{2l-1}^{(i)} + r^i$ and $a_{2l}^{(i)} = b_{2l}^{(i)} - r^i$ for $l = 1, 2, \ldots, 2^{i-2}$ for $i > 1$. Also, as expected, for the older points (polynomials of degree $i-1$), $a_{2l-1}^{(i)} = a_l^{(i-1)}$ and $b_{2l}^{(i)} = b_l^{(i-1)}$ for $l = 1, 2, \ldots, 2^{i-2}$ for $i > 1$.

Thus, for $i = 1$, there is only one interval, namely, $[a_1^{(1)}, b_1^{(1)}]$ with $a_1^{(1)} = 0$ and $b_1^{(1)} = r$ having positive λ-masses. For $i = 2$, there are three intervals, namely, $[a_1^{(2)}, b_1^{(2)}]$, $(b_1^{(2)}, a_2^{(2)})$, $[a_2^{(2)}, b_2^{(2)}]$ with $a_1^{(2)} = 0$, $b_1^{(2)} = r^2$, $a_1^{(2)} = r - r^2$, $b_2^{(2)} = r$ and $b_1^{(2)}$ and $a_2^{(2)}$ are the only two new points (of degree two) with positive λ-masses. Similarly, for $i = 3$, there are seven intervals and four new points with positive λ-masses, namely, $b_1^{(3)} = r^3$, $a_2^{(3)} = r^2 - r^3$, $b_3^{(3)} = r - r^2 + r^3$ and $a_4^{(3)} = r - r^3$. In all these cases, the new points of positive λ-masses are generated according to the observation in the previous paragraph.

Continuing this way, for every i, the 2^{i-1} polynomials in r of degree i with positive λ-masses in $[0, r]$ can be generated and each such polynomial of degree i is of the form $\sum_{j=1}^{k} (-1)^{j-1} r^{i_j}$ for $1 \le i_1 < i_2 < i_3 < \cdots < i_k = i$ for some $k \le i$. For the general case, induction on i may be used.

Part (*ii*): Using (2), the two polynomials of degree 2 with positive λ measures have masses equal to $\lambda(r)\, p_{10}$ and $\lambda(r)\, p_{01}$ respectively and the four polynomials of degree 3 with positive λ-measures have masses equal to $\lambda(r)\, p_{10}^2$, $\lambda(r)\, p_{10} p_{01}$, $\lambda(r)\, p_{01}^2$ and $\lambda(r)\, p_{10} p_{01}$ respectively. In general, consider a typical point in the support of λ. As we have seen in part (i) above, it is a polynomial of some degree $i > 1$ having the form $\sum_{j=1}^{k} (-1)^{j-1} r^{i_j}$ for some $1 \le i_1 < i_2 < i_3 < \cdots < i_k = i$ and for some $k \le i$. By mathematical induction, it follows that such a point has λ-measure equal to $\lambda(r)\, p_{10}^{i-1-k} p_{01}^k$.

One also observes that, for every $i > 1$, $b_{2^{i-2}} = r^2$ and $a_{2^{i-2}+1} = r - r^2$ so that for every i, there are two kinds of polynomials in r of degree i in $[0, r]$ with positive λ-masses: the first kind are numerically less than or equal to r^2 having positive λ-masses and the second kind are numerically greater than or equal to $r - r^2$. In each kind, there are exactly 2^{i-2} polynomials. By (2), the sum of the λ-masses for polynomials of the first kind is $\lambda(r)[p_{10} + p_{01}]^{i-2} p_{10}$ and that for the second kind is equal to $\lambda(r)[p_{10} + p_{01}]^{i-2} p_{01}$. As a result, the sum of the λ-masses for all the polynoimals in r of degree i in $[0, r]$ with positive λ-masses equals $\lambda(r)[p_{10} + p_{01}]^{i-1}$. This can also be proved without much difficulty by using induction on i.

Part (*iii*): The sum of the λ-masses of each of these polynomials in r of all finite degrees in $[0, r]$ with positive λ-masses along with the λ-mass at zero equals 1. Thus, if A_i denotes the collection of all 2^{i-1} points with positive λ-masses for $i = 1, 2, 3, \ldots$ and if we write $A = \bigcup_{i=1}^{\infty} A_i$, then it is clear that the support of λ, namely, $S(\lambda)$ equals $\{0\} \cup A$ and satisfies

$$\lambda(0) + \lambda(A) = \lambda(0) + \sum_{i=1}^{\infty} \lambda(A_i) = \lambda(0) + \lambda(r) \left[\sum_{i=1}^{\infty} (p_{10} + p_{01})^{i-1} \right] = 1.$$

From the above description of $S(\lambda)$ for $r < \frac{1}{2}$, we observe that $S(\lambda)$ turns out to be a countable closed set contained in $[0, r]$ consisting of 0 and the set A of polynomials in r:

$$A = \{r;\ r^2,\ r - r^2;\ r^3,\ r^2 - r^3,\ (r - r^2) + r^3,\ r - r^3;\ \ldots\}.$$

So, it is clear that the construction of the set A follows a Cantor-set type construction for $r < \frac{1}{2}$ and the proof is complete. The interested reader may contact anyone of the authors for the details of the proofs for parts (i) and (ii). □

PROPOSITION 2.2

For $r = \frac{1}{2}$, *we have the following*:

(i) *The only points that have positive λ-masses are the dyadic rationals in* $[0, \frac{1}{2}]$. *Thus, for every* i, *there are exactly* 2^{i-2} *dyadic rationals of the form* $\frac{k}{2^i}$ *with* $k \le 2^{i-1}$ *and* k *odd with positive λ-mass.*

(ii) *A typical point has λ-measure equal to* $\lambda\left(\frac{1}{2}\right)(p_{10} + p_{01})\, p_{10}^{i-1-k}\, p_{01}^{k-1}$ *for some positive integer* k. *For every* $i > 1$, *the sum of the λ-masses of all* 2^{i-2} *points that have positive λ-masses equals* $\lambda\left(\frac{1}{2}\right)[p_{10} + p_{01}]^{i-1}$.

(iii) *The sum of the λ-masses of all dyadic rationals in* $[0, \frac{1}{2}]$ *along with the λ-mass at zero equals* 1. *Thus, for* $r < \frac{1}{2}$, *for any positive integer* i, *the number of polynomials in* r *of degree* i *is exactly twice as many as the number of dyadic rationals of the form* $\frac{k}{2^i}$ *with* $k \le 2^{i-1}$ *and* k *odd.*

Proof. Substituting $r = \frac{1}{2}$ in the above steps (of the case $r < \frac{1}{2}$), we have $r^2 = r - r^2$ at stage 2, $r^3 = r^2 - r^3$ and $r - r^2 + r^3 = r - r^3$ at stage 3 and so on. In general, at any stage $i > 1$, $b_j^{(i)} = a_{j+1}^{(i)}$ for $j = 1, 2, \ldots, 2^{i-1} - 1$ so that for $r < \frac{1}{2}$, for any positive integer i, the number of polynomials in r of degree i is exactly twice as many as the number of polynomials in $\frac{1}{2}$ of degree i. But these polynomials in $\frac{1}{2}$ of degree i are exactly same as the dyadic rationals of the form $\frac{k}{2^i}$ with $k \le 2^{i-1}$ and k odd and thus (i) follows.

Now, note that the open intervals with λ-measure zero mentioned in part (i) of Proposition 2.1 are empty sets when $r = \frac{1}{2}$. Now, for any pair of $b_j^{(i)}$ and $a_{j+1}^{(i)}$, note that if the interval $(b_j^{(i)}, a_{j+1}^{(i)})$ appears for the first time at the i-th stage, then, j is odd, say, $j = 2l - 1$. Now considering l odd and even, we have two cases. If l is odd, then, let $l = 2m - 1$ so that $b_{2l-1}^{(i)} = a_m^{(i-2)} + r^i$ and $a_{j+1}^{(i)} = a_{2l}^{(i)} = a_m^{(i-2)} + r^{i-1} - r^i$. Now, if $a_m^{(i-2)} = \sum_{t=1}^{k}(-1)^t \left(\frac{1}{2}\right)^{i_t}$ then using the arguments in the case $r < \frac{1}{2}$, $\lambda(a_m^{(i-2)}) = \lambda\left(\frac{1}{2}\right) p_{10}^{i-3-k} p_{01}^{k}$ so that $\lambda(b_{2l-1}^{(i)}) + \lambda(a_{2l}^{(i)}) = \lambda\left(\frac{1}{2}\right) p_{10}^{i-3-k} p_{01}^{k+1}(p_{10} + p_{01})$. The same argument can be repeated when l is even and one gets the same λ-measure for $\lambda(b_{2l-1}^{(i)}) + \lambda(a_{2l}^{(i)})$ for the case l even also. But since, as argued before, for every j, we have $b_j^{(i)} = a_{j+1}^{(i)}$, so the above measure is the λ-measure of a single typical point in the case $r = \frac{1}{2}$.

Now for the remaining part of the proof, we follow the argument described for the case $r < \frac{1}{2}$ closely and show that the sum of the λ-masses for all dyadic rationals at stage $i > 1$ equals $\lambda\left(\frac{1}{2}\right)[p_{10} + p_{01}]^{i-1}$ and (ii) follows.

Finally, one observes that, for the case $r = \frac{1}{2}$ also, the sum of the λ-masses at all the dyadic rationals in $[0, \frac{1}{2}]$ equals 1 and (iii) follows. Thus, the proof of the proposition is complete. □

From the proof, it is clear that for $r < \frac{1}{2}$, for any positive integer i, the number of polynomials in r of degree i is exactly twice as many as the number of dyadic rationals of the form $\frac{k}{2^i}$ with $k \leq 2^{i-1}$ and k odd. The reader may be interested to check the special case scenario considering C_n and D_n to be independent and/or identically distributed.

3. Alternative proofs of the results in Section 2

In this section, we present alternative proofs of our results due to the referee.

Let the 2×2 i.i.d. stochastic matrices $(X_n)_{n\geq 1}$, C_n, D_n, the probability distribution μ of X_n, the support $S(\mu)$ of μ be as in §2. Also, we continue to assume that for a given r with $r = 1$ or $0 < r \leq \frac{1}{2}$, both $r^{-1}C_n$ and $r^{-1}D_n$ are Bernoulli with parameters p_1 and p_2 respectively as in §2.

Let λ be the distribution of $\lim_{n\to\infty} X_n X_{n-1} \cdots X_1$. Then λ is the distribution of

$$\sum_{k=1}^{\infty} D_k \prod_{i=1}^{k-1} (C_i - D_i). \tag{3}$$

From this, the fact that $D_k \prod_{i=1}^{k-1}(C_i - D_i)$ can take only the values 0 and $\pm r^k$ as well as the fact that with probability 1 it is 0 when k is large enough, we infer that λ is concentrated on a set of numbers of the form $\sum_{k=1}^{\infty} c_k r^k$, where $\mathbf{c} = (c_k)_{k\leq 1}^{\infty}$ is a sequence of $-1,\ 0,\ 1$. We will describe it more properly in the following paragraphs.

First, we consider the case $0 < r < \frac{1}{2}$. We show below that the support of λ, $S(\lambda)$, which is the set of all finite degree polynomials in r (having positive λ-masses) of the form $\sum_{j=1}^{k}(-1)^{j-1}r^{i_j}$ for some $1 \leq i_1 < i_2 < i_3 < \cdots < i_k = i$ for some $k \leq i$ and $i \geq 1$, is bijective to the following set:

$$\mathscr{B} = \{\mathbf{b} : \ \mathbf{b} \text{ is an infinite sequence of 0's and 1's with finite number of 1's}\}$$

so that if $\mathbf{b} \in \mathscr{B}$ is a sequence $b_1 b_2 \cdots$, then the bijection from $\mathscr{B}$ to $S(\lambda)$ is obtained in two stages. At the first stage, $\mathscr{B}$ is mapped to another set $\mathscr{C}$ as follows:

$$\phi(b_1 b_2 \cdots) = (\phi_1(b_1 b_2 \cdots),\ \phi_2(b_1 b_2 \cdots),\ \cdots),$$

where ϕ_i's take values -1, 0 or 1. Let $i_1, i_2, \ldots, i_k$ be the indices such that $b_j = 1$ if $j = i_1, i_2, \ldots, i_k$ and $b_j = 0$ otherwise. Then, $\phi_{i_l}(b_1 b_2 \cdots) = (-1)^{l-1}$ for $l = 1, \ldots, k$ and $\phi_j(b_1 b_2 \cdots) = 0$ otherwise. Thus, $\mathscr{C}$ is the following set:

$$\begin{aligned}\mathscr{C} \ = \ \{\mathbf{c} : \ & \mathbf{c} \text{ is an infinite sequence of 0's, 1's and} \\ & -1\text{'s with finite numbers of 1's and } -1\text{'s}\}\end{aligned}$$

so that if $\mathbf{c} = c_1c_2\cdots$ is a typical point in $\mathscr{C}$, then $c_j = 0$ if $b_j = 0$ and $c_j = 1$ or -1 if $b_j = 1$. As per the above description, $c_{i_l} = \phi_{i_l}(b_1b_2\cdots) = (-1)^{l-1}$. Then, finally, $\mathscr{C}$ is mapped to $S(\lambda)$ using the following map:

$$\psi(c_1c_2\cdots) = \sum_{i=1}^{\infty} c_j r^j.$$

It is clear that the above definition clearly covers all polynomials in r belonging to $S(\lambda)$. Thus, the bijection f between $\mathscr{B}$ and $S(\lambda)$ is as follows:

$$f(b_1b_2\cdots) = \psi\circ\phi(b_1b_2\cdots) = \sum_{j=1}^{\infty}\phi_j(b_1b_2\cdots)\,r^j.$$

To proceed with the proofs, denote the infinite sequence $000\cdots$ by $\mathbf{0}$ and the infinite sequence $100\cdots$ by $\mathbf{1}$. One defines a Markov transition kernel on $\mathscr{B}$ as follows:

(1) $p(\mathbf{0},\ \mathbf{0}) = p_{00} + p_{10} = q_2,\ p(\mathbf{0}, \mathbf{1}) = p_{01} + p_{11} = p_2.$
(2) For $\mathbf{b} \neq \mathbf{0}$, we have, $p(\mathbf{b},\ \mathbf{0}) = p_{00},\ p(\mathbf{b},\ 0\mathbf{b}) = p_{10},\ p(\mathbf{b},\ 1\mathbf{b}) = p_{01},\ p(\mathbf{b},\ \mathbf{1}) = p_{11}.$

Now let us make the following observations.

PROPOSITION 3.1

The above Markov chain is irreducible, aperiodic and positive recurrent. Its stationary distribution π *is given by*

$$\pi(\mathbf{0}) = \frac{p_{00}}{1-p_{10}}, \quad \pi(\mathbf{1}) = p_{11} + \frac{p_{00}p_{01}}{1-p_{10}} \tag{4}$$

and

$$\pi(\mathbf{b}) = p_{10}^{j}p_{01}^{k-1-j}\pi(\mathbf{1}),$$

when $\mathbf{b} = b_1\cdots b_k000\cdots$ *with* $b_k = 1$ *and* j *is the size of* $\{i < k : b_i = 0\}$.

Proof. Irreducibility and aperiodicity follow easily. Using the equation $\boldsymbol{\pi}(\mathbf{b}') = \sum_{\mathbf{b}\in\mathscr{B}}\boldsymbol{\pi}(\mathbf{b})\,p(\mathbf{b},\ \mathbf{b}')$ shows (4). If $\mathbf{b}' \neq \mathbf{0}$ or $\mathbf{1}$, there exists $\mathbf{b}'' \neq \mathbf{0}$ such that $\mathbf{b}' = \mathbf{b}\mathbf{b}''$ with $b = 0$ or 1. Therefore, $p(\mathbf{b},\ \mathbf{b}') = p(\mathbf{b},\ \mathbf{b}\mathbf{b}'') \neq 0$ if and only if $\mathbf{b} = \mathbf{b}''$. As a result, $p(\mathbf{b}'',\ b\mathbf{b}'') = p_{01}$ if $b = 1$ and p_{10} if $b = 0$. Also, $\boldsymbol{\pi}(\mathbf{b}\mathbf{b}'') = p_{01}\boldsymbol{\pi}(\mathbf{b}'')$ or $p_{10}\boldsymbol{\pi}(\mathbf{b}'')$ accordingly. If $\mathbf{b} = b_1\cdots b_k000\cdots$ with $b_k = 1$, then it follows by iteration that if j is the number of $i < k$ such that $b_i = 0$, we have the following:

$$\boldsymbol{\pi}(\mathbf{b}) = \gamma_{b_1}\cdots\gamma_{b_k}\boldsymbol{\pi}(\mathbf{1}) = p_{10}^{j}p_{01}^{k-1-j}\boldsymbol{\pi}(\mathbf{1}),$$

where $\gamma_b = p_{10}$ for $b = 0$ and $\gamma_b = p_{01}$ for $b = 1$. □

Remark 3.1. One easily checks that

$$\begin{aligned}\sum_{\mathbf{b}\in\mathscr{B}}\boldsymbol{\pi}(\mathbf{b}) &= \boldsymbol{\pi}(\mathbf{0}) + \boldsymbol{\pi}(\mathbf{1})\sum_{k=1}^{\infty}\sum_{j=0}^{k-1}\binom{k-1}{j}p_{10}^{j}p_{01}^{k-1-j}\\ &= \boldsymbol{\pi}(\mathbf{0}) + \boldsymbol{\pi}(\mathbf{1})\frac{1}{p_{00}+p_{11}} = 1.\end{aligned}$$

PROPOSITION 3.2

Let f be as introduced before Proposition 3.1. *The probability distribution λ is the image of π by the map $\mathbf{b} \mapsto f(\mathbf{b})$. In other words,*

$$\lambda = \sum_{\mathbf{b}\in\mathscr{B}} \boldsymbol{\pi}(\mathbf{b})\delta_{f(\mathbf{b})}$$

Proof. Consider the Markov chain (Y_n) on $\mathscr{B}$ of our Proposition 3.1. Then, under the set up of §2, we have

$$f(Y_{n+1}) = (C_n - D_n)f(Y_n) + D_n = C_n f(Y_n) + D_n (1 - f(Y_n)).$$

Therefore, if the distribution of the random sequence Y of $\mathscr{B}$ is the stationary distribution $\boldsymbol{\pi}$ of the Markov chain (Y_n), then the distribution of $f(Y_n)$ is λ.

Thus, the case $r < \frac{1}{2}$ follows from Propositions 1 and 2.

For the case $r = \frac{1}{2}$, we have the following remark. □

Remark 3.2. For the case $r = \frac{1}{2}$, the map f introduced above is not one-to-one as $\left(\frac{1}{2}\right)^2 = f(01\cdots) = f(11\cdots) = \frac{1}{2} - \left(\frac{1}{2}\right)^2$. Thus, the proof in this case needs to be worked out separately.

PROPOSITION 3.3

If $r = \frac{1}{2}$ and $\mathbf{b} \neq \mathbf{0}, \mathbf{1}$, then $\lambda(f(\mathbf{b})) = p_{10}^{j} p_{01}^{m-j}(p_{10} + p_{01})\boldsymbol{\pi}(\mathbf{1})$.

Proof. We only give the main ideas of the proof. The details are left to the reader.

If $\mathbf{b} \neq \mathbf{0}, \mathbf{1}$, then there exists a finite sequence $\mathbf{b}^*$ of 0's and 1's such that $\mathbf{b} = \mathbf{b}^*01000\cdots$ or $\mathbf{b} = \mathbf{b}^*11000\cdots$ and we can call $\mathbf{b}^*$ the type of $\mathbf{b}$. Then $f(\mathbf{b}) = f(\mathbf{b}')$ if and only if $\mathbf{b}$ and $\mathbf{b}'$ have the same type. If $\mathbf{b}^*$ has j zeros and length m, then $\lambda(f(\mathbf{b})) = \boldsymbol{\pi}(\mathbf{b}^*01000\cdots) + \boldsymbol{\pi}(\mathbf{b}^*11000\cdots) = p_{10}^{j} p_{01}^{m-j}(p_{10} + p_{01})\boldsymbol{\pi}(\mathbf{1})$ etc. The part '$\Rightarrow$' is less obvious and involves a little discussion about the two ways of representing an odd integer as a sum $\sum_{j=1}^{k}(-1)^j 2^{p_j}$ when $p_1 > p_2 > \cdots > p_k = 1$.

Acknowledgement

Our original proof assumed that C_n and D_n were independent. However, the referee helped us observe that the proof still worked even when C_n and D_n were not independent. We also thank the referee for his shorter alternative proofs.

References

[1] Chamayou J F and Letac G, A transient random walk on stochastic matrices with Dirichlet distributions, *Ann. Prob.* **22** (1994) 424–430

[2] Chassaing P, Letac G and Mora M, Brocot sequences and random walks in SL(2, *R*), *Springer LNM* **1034** (1985) 37–50

[3] Hognas G and Mukherjea A, Probability measures on semigroups (2nd edition) (2011) (Springer, New York)

[4] Van Assche W, Products of 2×2 stochastic matrices with random entries, *J. Appl. Prob.* **23** (1986) 1019–1024