2006

Some problems on products of random matrices

Edgardo S. Cureg
University of South Florida

Some Problems on Products of Random Matrices

by

Edgardo S. Cureg

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctorate of Philosopy
Department of Mathematics
College of Arts and Sciences
University of South Florida

Major Professor: Arunava Mukherjea, Ph.D.
Athanassios Kartsatos, Ph.D.
Jogindar Ratti, Ph.D.
Stephen Suen, Ph.D

Date of Approval:
November 14, 2006

Keywords: weak convergence, stochastic matrices, circulant matrices, Lyapunov exponents, random Fibonacci sequences

Dedication

To the memory of my late father, Bernardo C. Cureg, Sr. (1931-1991), who taught me the value of education.

Acknowledgments

This work owes much of its development to the helpful support, guidance, suggestions, ideas, criticism, and patience of the author's major professor, Dr. Arunava Mukherjea, whom he sincerely thanks for everything concerned with the inception, production, and completion of this dissertation. It has been more than five years after Dr. Mukherjea gave a course in random matrices at USF in the summer of 2001, thus introducing the author to this field, and since then it has been an incredible learning experience – no, journey – for him. Thank you so much, Dr. Mukherjea!

The author would also like to thank his committee members, Drs. Dennis Killinger, Athanassios Kartsatos, Jogindar Ratti, and Stephen Suen, for reviewing and evaluating this dissertation. He also extends his sincerest gratitude to Ayako Ossowski, Nancy Morris, and Sarina Maldonado, for all the help with the paperwork and documentation involved with this dissertation; and to Dr. David Kephart, one of his most trusted friends and whose mathematical acuity he so admires and envies, not only for providing the LaTeX template for this dissertation, but for his faithful friendship since the very beginning.

The author wishes to acknowledge the assistance provided by the National Science Foundation through their Graduate Research Fellowship Program. NSF has provided the author not only with funds for tuition and stipend from Summer 2001 to Summer 2004, but, more importantly, the opportunity to conduct collaborative work with international researchers in Taiwan and Finland. He extends his deepest appreciation to Prof. Göran Högnäs of Åbo Akademi University in Turku, Finland, for hosting him as an NSF international research travel grantee during the summer of 2004, and for inviting him to the 27th Finnish Summer School on Probability. He also thanks Prof. Yeong-Nan Yeh of the Institute of Mathematics, Academia Sinica, Taipei, Taiwan, for hosting him as a participant to the 2002 NSF Summer Programs in Japan, Korea, and Taiwan (now called NSF East Asia and Pacific Summer Institutes for U.S. Graduate Students (EAPSI)).

Thanks are also due Prof. Nataša Jonoska of USF's Department of Mathematics, as well as Profs. Arlene Pascasio, Blessilda Raposa, and Leonor Ruivivar of the Mathematics Department, De La Salle University, Manila, Philippines, for their support of the author's application for the NSF Graduate Research Fellowship.

The author would also like to thank Japan's Ministry of Education, Culture, Sports, Science and Technology (formerly known as *Monbusho*) for the scholarships it provided from 1988 to 1992 at Kumamoto National College of Technology and from 1992 to 1994 at Toyohashi University of Technology, and the Tokyo-based IBM Asia for a 1994-1996 fellowship at TUT. He offers his deepest thanks to his Japanese friend, Takiko "Tina" Nagaya, for her love and generosity through all these years, and to his sister Leila Taghap, whom he loves dearly and to whom he is forever indebted, for her financial support of his early education in the Philippines.

Finally, the author would like to thank his best friend, Dr. Harry Brown, for the moral support during times it was most needed. You know the number, Harry!

Table of Contents

List of Figures ii

Abstract iii

Chapter 1 Introduction 1

Chapter 2 Weak Convergence in Stochastic Matrices 8
- 2.1 Introduction 8
- 2.2 Preliminaries 10
- 2.3 4×4 Stochastic Matrices 21
- 2.4 The $d \times d$ case 36

Chapter 3 Weak Convergence in Circulant Matrices 49
- 3.1 $d \times d$ Circulant Matrices 49
- 3.2 3×3 Circulant Matrices 56
- 3.3 4×4 Circulant Matrices 66
- 3.4 $d \times d$ Toeplitz Matrices 76

Chapter 4 Numerical Calculation of Lyapunov Exponents for Some Random Fibonacci Recurrences 79
- 4.1 Random Matrix Products and Random Fibonacci Sequences 79
- 4.2 Numerical Calculation of the Lyapunov Exponents 85

References 104

About the Author . . . . End Page

List of Figures

1 $\nu_1(p)$ for different values of p 88
2 $\nu_2(p)$ for different values of p 89
3 $\nu_\beta(p)$, where $\beta = 1/2$, for different values of p 90
4 $\nu_\beta(p)$, where $\beta = 3/4$, for different values of p 91
5 $\nu_\beta(p)$, where $\beta = 2$, for different values of p 92
6 $\nu_\beta(p)$, where $\beta = 8$, for different values of p 93
7 $\nu_\beta(p)$ for fixed $p = 0.2$ and different values of $\beta < 1$ 94
8 $\nu_\beta(p)$ for fixed $p = 0.4$ and different values of $\beta < 1$ 95
9 $\nu_\beta(p)$ for fixed $p = 0.6$ and different values of $\beta < 1$ 96
10 $\nu_\beta(p)$ for fixed $p = 0.8$ and different values of $\beta < 1$ 97
11 $\nu_\beta(p)$ for fixed $p = 0.2$ and different values of $\beta > 1$ 98
12 $\nu_\beta(p)$ for fixed $p = 0.4$ and different values of $\beta > 1$ 99
13 $\nu_\beta(p)$ for fixed $p = 0.6$ and different values of $\beta > 1$ 100
14 $\nu_\beta(p)$ for fixed $p = 0.8$ and different values of $\beta > 1$ 101
15 γ_1 vs. p 102
16 γ_2 vs. p 102
17 γ_β vs. p for $\beta = 0.5, 0.75, 2$, and 8 103

Some Problems on Products of Random Matrices

Edgardo S. Cureg

ABSTRACT

We consider three problems in this dissertation, all under the unifying theme of random matrix products. The first and second problems are concerned with weak convergence in stochastic matrices and circulant matrices, respectively, and the third is concerned with the numerical calculation of the Lyapunov exponent associated with some random Fibonacci sequences.

Stochastic matrices are nonnegative matrices whose row sums are all equal to 1. They are most commonly encountered as transition matrices of Markov chains. Circulant matrices, on the other hand, are matrices where each row after the first is just the previous row cyclically shifted to the right by one position. Like stochastic matrices, circulant matrices are ubiquitous in the literature.

In the first problem, we study the weak convergence of the convolution sequence μ^n, where μ is a probability measure with support S_μ inside the space S of $d \times d$ stochastic matrices, $d \geq 3$. Note that μ^n is precisely the distribution of the product $X_1 X_2 \cdots X_n$ of the μ-distributed independent random variables $X_1, X_2, \ldots, X_n$ taking values in S. In [CR] Santanu Chakraborty and B.V. Rao introduced a cyclicity condition on S_μ and showed that this condition is necessary and sufficient for μ^n to <u>not</u> converge weakly when $d = 3$ and the minimal rank r of the matrices in the closed semigroup S generated by S_μ is 2. Here, we extend this result to any $d > 3$. Moreover, we show that when the minimal rank r is not 2, this result does not always hold.

The second problem is an investigation of weak convergence in another direction, namely the case when the probability measure μ's support S_μ consists of $d \times d$ circulant matrices, $d \geq 3$, which are not necessarily nonnegative. The resulting semigroup S generated by S_μ now lacking the nice property of compactness in the case of stochastic matrices, we assume tightness of the sequence μ^n to analyze the problem. Our approach is based on the work of Mukherjea and his collaborators, who in [LM] and [DM] presented a method based on a bookkeeping of the possible structure of the compact kernel K of S.

The third problem considered in this dissertation is the numerical determination of Lyapunov exponents of some random Fibonacci sequences, which are stochastic versions of the classical Fibonacci sequence

$$f_{n+1} = f_n + f_{n-1},\ n \geq 1, \text{ and } f_0 = f_1 = 1,$$

obtained by randomizing one or both signs on the right side of the defining equation and/or adding a "growth parameter." These sequences may be viewed as coming from a sequence of products of i.i.d. random matrices and their rate of growth measured by the associated Lyapunov exponent. Following techniques presented by Embree and Trefethen in their numerical paper [ET], we study the behavior of the Lyapunov exponents as a function of the probability p of choosing $+$ in the sign randomization.

Chapter 1
Introduction

To set the framework for the problems we consider in this dissertation, we recall the definition of weak convergence in the context of measures on topological semigroups [HMu]. Let S be a locally compact Hausdorff second-countable topological semigroup. Let $\mathcal{B}$ be the class of Borel subsets of S. Let $P(S)$ be the set of all regular probability measures defined on $\mathcal{B}$, i.e. measures μ that satisfy the condition that for every $\epsilon > 0$, there exists a compact set $K_\epsilon \in \mathcal{B}$ for which $\mu(S \setminus K_\epsilon) < \epsilon$. The support S_μ of $\mu \in P(S)$ is given by

$$S_\mu = \{x \in S : \mu(V) > 0 \text{ for any open set } V \text{ containing } x\}.$$

Note that S_μ is closed and $\mu(S_\mu) = 1$.

A sequence $(\mu_n)_{n\geq 1}$ in $P(S)$ is then said to be weakly convergent to $\mu \in P(S)$ if

$$\lim_{n\to\infty} \int f\, d\mu_n = \int f\, d\mu$$

for every bounded (real) continuous function f on S.

If $\mu \in P(S)$ and $X_1, X_2, \ldots, X_n$ are μ-distributed independent random variables taking values in S, then the product $X_1 X_2 \cdots X_n$ has distribution

$$\mu * \mu * \cdots * \mu \ (n \text{ factors}) \equiv \mu^n.$$

Here, the convolution product $\mu * \nu$ of $\mu, \nu \in P(S)$ is the unique regular probability measure on $\mathcal{B}$ guaranteed by the Riesz representation theorem to exist and satisfy the equation

$$\int f\, d(\mu * \nu) = \int \int f(xy)\, \mu(dx)\nu(dy)$$

for every bounded countinuous function f on S with compact support. Or, more conveniently,

$$\mu * \nu(B) = \int \mu(Bx^{-1})\, \nu(dx) = \int \nu(x^{-1}B)\, \mu(dx) \qquad (1.\ 1)$$

for $B \in \mathcal{B}$. The sets Bx^{-1} and $x^{-1}B$ in (1. 1) are defined by

$$Bx^{-1} \equiv \{y \in S \,:\, yx \in B\}$$

and

$$x^{-1}B \equiv \{y \in S \,:\, xy \in B\},$$

respectively.

The first two problems considered here may then be simply described as the determination of necessary and sufficient conditions for weak convergence of the convolution sequence $(\mu^n)_{n\geq 1}$, when μ has support S_μ consisting of either $d \times d$ stochastic matrices or $d \times d$ circulant matrices, and

$$S = \overline{\bigcup_{n\geq 1} S_\mu^n}$$

is the closed (multiplicative) semigroup generated by S_μ.

Stochastic matrices are nonnegative matrices whose row sums are all equal to one. Ubiquitous in the literature, they are most commonly encountered as transition matrices of Markov chains.

The motivation for studying weak convergence in stochastic matrices comes from the work of Mukherjea and his students. In [LM] and [DM], Lo and Mukherjea and Dhar and Mukherjea present solutions to the problem of weak convergence of convolution powers μ^n of a probability measure μ with support S_μ in $d \times d$ nonnegative and stochastic matrices, respectively, in terms of easily verifiable conditions on S_μ. As Dhar and Mukherjea note in their paper, the problem of convergence in distribution of products of $d \times d$ i.i.d. stochastic matrices is an old one, reaching as far back as Rosenblatt's 1965 work [MR2]. A method that extends to $d = 3$ Mukherjea's simple and complete result for $d = 2$, that weak convergence occurs if and only if the support of μ is not equal to

$$\left\{ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\}$$

[M1], is presented in their paper. This method is based on a general result giving a necessary and sufficient condition for weak convergence of a <u>tight</u> sequence (μ^n) of convolution powers of a probability measure μ (that is, given $\epsilon > 0$, there exists a compact subset K_ϵ of S such that $\mu^n(K_\epsilon) > 1 - \epsilon$ for all $n \geq 1$) and looks at the possible structure of the compact kernel K (which, for semigroups of stochastic matrices, is well-known to consist of matrices with the minimal rank [HMu]) of the closed semigroup S generated by S_μ. The method "works for $d > 3$ even though calculations are more involved for higher values of d."

In another paper dealing with the same subject for $d = 3$, Chakraborty and Rao [CR] present a different solution, this time based on a division of S into certain subsets according to the number of recurrent and transient classes. Their method is "too cumbersome to be carried over to higher dimensions," but it succeeded in expressing the same result obtained by Dhar and Mukherjea in more succint terms using a cyclicity property of S_μ. According to their definition, S_μ is <u>cyclic</u> if there are pairwise disjoint subsets $A_1, A_2, \ldots, A_k$ of $\{1, 2, \ldots, d\}$ such that for any s, $1 \leq s \leq k$, and for any $i \in A_s$,

$$\sum_{j \in A_{s+1}} x_{ij} = 1$$

for any element $x = (x_{ij}) \in S_\mu$. Note that $A_{k+1} = A_1$ in the sum. Chakraborty and Rao then proved in [CR] that when $d = 3$, as long as S_μ is not contained in a group of permutation matrices, μ^n does not converge weakly if and only if S_μ is cyclic.

The expansion of the methods presented in [DM] and [CR] to 4×4 stochastic matrices, as well as the generalization of Chakraborty and Rao's result in the case when the rank of the matrices in the kernel K of the semigroup S generated by S_μ, where S_μ is contained in a set of $d \times d$ stochastic matrices, $d \geq 4$, is 2, is essentially the subject of our first problem. Part of the findings in this investigation is that in the general $d \times d$ situation, this characterization of non-weak convergence of μ^n in terms of cyclicity of S_μ is no longer valid when the rank of the matrices in the kernel K of S is bigger than 2.

In obtaining our main results, we followed Dhar and Mukherjea's approach and used the general theorem mentioned above. From the proofs presented here, however, it is evident that the connection between cyclicity and non-weak convergence does not follow easily from this general result.

Leaving the domain of stochastic matrices and their nonnegativity property, we next investigate the problem of weak convergence of μ^n in the context of circulant matrices, which are not necessarily nonnegative. Circulant matrices are matrices where each row after the first is just the previous row cyclically shifted to the right by one position. Familiar examples are the identity matrix, the zero matrix, and the all 1s matrix. A 3×3 circulant matrix has the form

$$\begin{pmatrix} a & b & c \\ c & a & b \\ b & c & a \end{pmatrix}.$$

These matrices appear in many mathematical problems. A detailed account of this (with many examples) can be found in the beautiful work of Diaconis[Di]. Take, for example, Diaconis' Example 2.1 (p. 40, [Di]). He considers "a particle constrained to hop about on n points arranged in a circle. At each time the particle hops left or right with probability $\frac{1}{2}$. This is the cyclic version of the classical drunkard's walk. Index the points as $0, 1, 2, \ldots, n-1$. The chance of moving from i to j is thus

$$M_{ij} = \begin{cases} \frac{1}{2}, & \text{if } |i-j| = 1; \\ 0, & \text{otherwise.} \end{cases}$$

The matrix M is a circulant matrix with first row $(0, \frac{1}{2}, 0, \ldots, 0, \frac{1}{2})$." Diaconis goes on in his paper discussing problems involving covariance matrices, cyclic codes, etc., where one encounters circulant matrices.

Circulant matrices have a nice structure, and as far as we know, the problem of convergence in distribution of products of $d \times d$ i.i.d. random circulant matrices, or, equivalently, the problem of weak convergence of convolution powers μ^n of a probability measure μ supported on such matrices, has not been studied so far. Hence this is the second problem we consider. We clarify that in the circulant case, S, the closed semigroup generated by the support S_μ of μ, usually is not compact (in the usual topology), so we assume tightness of the sequence $(\mu^n)_{n\geq 1}$. Our main result points to the importance of the special orthogonal group $SO(d)$ in the characterization of weak convergence of (μ^n).

The third problem considered in this dissertation is the numerical determination of Lyapunov exponents of some random Fibonacci sequences. A random Fibonacci sequence is a stochastic version of the classical Fibonacci sequence

$$f_{n+1} = f_n + f_{n-1},\ n \geq 1,\ \text{and } f_0 = f_1 = 1$$

obtained by randomizing one or both of the signs on the right side of the defining equation.

For example, one version of a random Fibonacci sequence is the one originally considered by Viswanath in [Vi], given by

$$x_{n+1} = \pm x_n \pm x_{n-1} \qquad (n \geq 1) \tag{1. 2}$$

with $x_0 = x_1 = 1$, and where the signs are chosen independently and with equal probabilities. Viswanath determined the rate of growth of this random sequence. Recall that the rate of growth of a random sequence coming from a sequence of i.i.d. random matrices is the exponential of its associated Lyapunov exponent, which, by a result of Furstenberg and Kesten [FK], is equal to the almost sure limit

$$\lim_{n\to\infty} \frac{\log |x_n|}{n}.$$

In [Vi], Viswanath found the exact value of the rate of growth γ_f of the random Fibonacci recurrence (1. 2) to be

$$\lim_{n\to\infty} |x_n|^{1/n} = e^{\gamma_f} = 1.13198824\ldots \tag{1. 3}$$

with probability 1. This result was obtained using “the theory of random matrix products, Stern-Brocot division of the real line, a fractal measure, and a rounding error analysis to validate the computer calculation.” Observe that the rate of growth of the classical Fibonacci sequence is given by the golden ratio, $\frac{1+\sqrt{5}}{2} \approx 1.618$.

Viswanath actually used the random recurrence

$$x_{n+1} = \pm x_n + x_{n-1} \tag{1. 4}$$

in his calculation, since under the same initial conditions and independence and equal probability of choosing the signs, (1. 2) and (1. 4) give rise to the same quantity given in (1. 3). In fact, the recurrence

$$x_{n+1} = x_n \pm x_{n-1} \tag{1. 5}$$

could have been used as well.

In his concluding remarks on the subject, Viswanath posed a generalization of the problem in which each $\pm$ sign is still chosen independently in (1. 2) but $+$ and $-$ occur with probabilities p and $q := 1 - p$, respectively, where $0 < p < 1$. Noting that the techniques he used to calculate the Lyapunov exponent $\gamma_f(p)$ for $p = 1/2$ "do not seem to generalize easily" to arbitrary values of p, Viswanath instead calculated $\gamma_f(p)$ numerically for different values of p using Ulam's method [HMi]. The resulting graph of $\gamma_f(p)$ vs. p shows a smooth dependence of $\gamma_f(p)$ on p, a result consistent with Peres' theorem [Pe].

We emphasize that Viswanath's numerical calculation of $\gamma_f(p)$ was done for the random recurrence (1. 2). Viswanath did not consider the numerical approximation of Lyapunov exponents for the corresponding generalization to the random recurrences (1. 4) and (1. 5). Thus, here we investigate this problem. Our numerical results in this case suggest that the Lyapunov exponent for (1. 5) exhibits symmetry with respect to $p = 1/2$, whereas for (1. 4) the Lyapunov exponent monotonically increases with p, but not in the same manner as the Lyapunov exponent for (1. 2) reported by Viswanath.

In a related article [ET], Embree and Trefethen gave a numerical description of what they called the "Lyapunov constant" $\sigma(\beta) = \lim_{n\to\infty} |x_n|^{1/n}$ (with probability 1) for the random Fibonacci recurrence

$$x_{n+1} = x_n \pm \beta x_{n-1}, \tag{1. 6}$$

where $\beta > 0$, the signs are chosen independently and with equal probabilities, and $x_0 = x_1 = 1$. They found that for a certain range of values of the parameter β, the Lyapunov constant is less than 1 (resulting in the exponential decay of the solutions to the random recurrence), and for values of β outside this range, the Lyapunov constant is greater than 1 (hence the solutions grow exponentially). They further observed that σ depends on β in a non-smooth, fractal way.

In the section on "Discussions and Generalizations" of the same paper, Embree and Trefethen posed as one of the modifications of the random Fibonacci recurrence $(1.\ 6)$ the following generalization: "The coin might be weighted, so that $+$ is chosen with probability p and $-$ with probability $1 - p$." The other problem we consider here is exactly this generalization. Our results suggest that for values of $\beta < 1$, there appears to exist some $p^* = p^*(\beta)$ for which the Lyapunov exponent for $(1.\ 6)$ is 0, meaning the random Fibonacci sequence in this case neither grows nor decays.

Chapter 2
Weak Convergence in Stochastic Matrices

2.1 Introduction

Stochastic matrices are nonnegative matrices whose row sums are all equal to 1. Ubiquitous in the literature, these matrices are most commonly encountered as transition matrices of Markov chains.

We now describe briefly the problem of weak convergence in stochastic matrices. Let μ be a probability measure on $\mathcal{B}$, the Borel sets of $d \times d$ stochastic matrices. Here, d is a positive integer greater than 1. Let S_μ be the support of μ and let

$$S = \overline{\bigcup_{n\geq 1} S_\mu^n}$$

be the closed multiplicative semigroup generated by S_μ. Notice that S is then a compact Hausdorff topological semigroup (with respect to usual matrix topology and matrix multiplication). We define the convolution iterates μ^n in the usual manner. In other words, for any $B \in \mathcal{B}$,

$$\mu^{n+1}(B) = \int \mu^n\{y : yx \in B\}\, \mu(dx),$$

for all $n \geq 1$. We then say that the sequence $(\mu^n)_{n\geq 1}$ weakly converges to a probability measure ν if and only if

$$\lim_{n\to\infty} \int f\, d\mu^n = \int f\, d\nu$$

for every bounded (real) continuous function f on S.

In [DM], Dhar and Mukherjea present a solution to the problem of weak convergence of μ^n when the support S_μ of μ is contained in a set of 3×3 stochastic matrices. Their solution is expressed in terms of easily verifiable conditions on S_μ, similar in spirit to Mukherjea's simple and beautiful result that when S_μ is contained in 2×2 stochastic matrices, weak convergence occurs if and only if

$$S_\mu \neq \left\{ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\}$$

[M1]. Their method is based on a bookkeeping of the possible structure of the compact kernel K of S which, for semigroups of stochastic matrices, is well-known to consist of all matrices in S with the minimal rank [HMu].

In another paper dealing with the same problem, Chakraborty and Rao [CR] introduce a "cyclicity" property for the support S_μ, calling S_μ *cyclic* if there are pairwise disjoint subsets $A_1, A_2, \ldots, A_k$ of $\{1, 2, \ldots, d\}$ such that for any s, $1 \leq s \leq k$, and for any $i \in A_s$,

$$\sum_{j \in A_{s+1}} P_{ij} = 1$$

for any element $P = (P_{ij}) \in S_\mu$. Note that $A_{k+1} = A_1$ in the sum. For example, the set

$$\left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 1-a-b & a & b \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \right\}$$

where $a, b \geq 0$ and $0 < a + b \leq 1$, is cyclic with $A_1 = \{2\}$ and $A_2 = \{3\}$. Chakraborty and Rao's result, that in all cases except when S_μ is contained in a group of permutation matrices, the cyclicity property of the elements in S_μ is necessary and sufficient for the sequence μ^n not to converge weakly, is the same result obtained by Dhar and Mukherjea expressed in more succint terms.

One motivation for our results in this chapter is to investigate if this connection between cyclicity and weak convergence continues to hold even for $d \times d$ stochastic matrices, where $d > 3$. We will show here that, as long as the common rank of the matrices in the kernel K is 2, the equivalence of cyclicity of the support S_μ and non-weak convergence of μ^n still holds when $d > 3$, but the cyclicity property is not necessary for non-weak convergence of μ^n if the rank of the matrices in K is bigger than 2.

2.2 Preliminaries

We begin with some notations and standard definitions. From now on, unless otherwise stated, all matrices have real entries and are $d \times d$, with $d \geq 2$.

A matrix A is stochastic if its entries are all nonnegative and the sum of the entries in each row is 1.

A set S of matrices forms a semigroup if it is closed under matrix multiplication. A semigroup S is said to be left-zero (resp. right-zero) if $AB = A$ (resp. $AB = B$) for all $A, B \in S$.

A subset X of a semigroup S is called a right ideal if $XS \subseteq X$, where (as usual) $XS = \{AB : A \in X, B \in S\}$. A left ideal is defined similarly. X is a two-sided ideal, or simply an ideal, if it is simultaneously a left- and right ideal of S. The smallest (relative to set inclusion) ideal of S is called its kernel. S is called simple if it has no proper ideals.

A matrix $A \in S$ is idempotent if $A^2 = A$. If A is idempotent, and, in addition, there is no other idempotent $B \in S$ satisfying $AB = BA = B$, then A is called primitive. S is called completely simple if it is simple and it contains a primitive idempotent.

Idempotent stochastic matrices will play a major role in the analysis needed for our problem, so we next present some results that will be used in the sequel. The first such result is the following well-known structure theorem for idempotent stochastic matrices (see, for example, [M2]).

THEOREM 2.1 *Let A be a $d \times d$ idempotent stochastic matrix. Let p be the rank of A. Then there is a partition $\{T, C_1, C_2, \ldots, C_p\}$ of $\{1, 2, \ldots, d\}$, called a basis of A, such that the following hold:*

1. *$j \in T$ means that the jth column of A is a zero column,*

2. *each $C_k \times C_k$ block of A is a strictly positive block with identical rows, and*

3. *each $C_j \times C_k$, $j \neq k$, block of A is an all zero block.*

Proof. See [M2]. ∎

The set T in a basis $\{T, C_1, C_2, \ldots, C_p\}$ of A is called its T-class, and the C_k's are called its C-classes.

Two $d \times d$ idempotent stochastic matrices A and B are called essentially the same if they have the same rank p and have bases given by:

$$\begin{aligned} A &: \ \{T, C_1, C_2, \ldots, C_p\}, \\ B &: \ \{T', C_1', C_2', \ldots, C_p'\}, \end{aligned} \tag{2. 1}$$

such that for $1 \leq j, k \leq p$, the $T \times C_k$ block of A is identical to the $T' \times C_k'$ block of B (for each k), and the $C_j \times C_k$ block of A is identical to the $C_j' \times C_k'$ block of B for each pair (j, k). A and B are of the same type iff their bases as given by (2. 1) above are such that

$$|T| = |T'|, |C_1| = |C_1'|, |C_2| = |C_2'|, \ldots, |C_p| = |C_p'|,$$

and furthermore, for each $t \in T$, there is a unique $t' \in T'$ such that for each k, $1 \leq k \leq p$. the $\{t\} \times C_k$ block of A is a strictly positive block iff the $\{t'\} \times C_k'$ block of B is also so.

EXAMPLE 1 The matrices

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1/2 & 1/2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \text{ and } \begin{pmatrix} 0 & 1/3 & 2/3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

are essentially different from each other, although the second and the third matrices are of the same type. The first matrix, however, has a type different from that of the other two.

The following theorem asserts that as far as idempotent matrices are concerned, the concepts of similarity and essential sameness are equivalent.

THEOREM 2.2 *Two $d \times d$ stochastic idempotent matrices A and B are essentially the same iff there is a $d \times d$ permutation matrix P such that $PB = AP$.*

Proof. Suppose A and B are essentially the same. Consider the bases of A and B as given above in (2. 1). Define the bijection f from $\{1, 2, \ldots, d\}$ onto itself such that $f(T) = T'$ and $f(C_k) = C_k'$ for each k, and furthermore, for all $i, j \in \{1, 2, \ldots, d\}$,

$$A_{ij} = B_{f(i), f(j)}.$$

Note that this is clearly possible since A and B are essentially the same. Now define the permutation matrix P such that

$$P_{ij} = \begin{cases} 1 & \text{if } j = f(i), \\ 0 & \text{otherwise.} \end{cases}$$

Then, for all i, j,

$$(PB)_{ij} = \sum_k P_{ik} B_{kj} = P_{i,f(i)} B_{f(i),j} = B_{f(i),j}. \tag{2. 2}$$

Write $t = f^{-1}(j)$. Then

$$(AP)_{ij} = \sum_k A_{ik} P_{kj} = A_{it} P_{tj} = A_{it},$$

which is equal to the right hand side of (2. 2) since $f(t) = j$. Thus, $PB = AP$.

Conversely, every permutation matrix P defines a bijection f as above such that for each i, the element on the ith row and $f(i)$th column of P is 1. This means that the element on the ith row and jth column of P is the same as the element on the $f(i)$th row and $f(j)$th column of $P^{-1}AP$. Thus, if we define $T' = f(T)$, and $C'_k = f(C_k)$, where A has the basis $\{T, C_1, C_2, \ldots, C_p\}$, then A and $B = P^{-1}AP$ are essentially the same.

∎

COROLLARY 2.0.1 *Two $d \times d$ idempotent stochastic matrices A and B are of the same type iff there is a permutation matrix P such that the matrices B and $P^{-1}AP$ have the same bases.*

Proof. Immediate from the proof of Theorem 2.2.

∎

We note that in Corollary 2.0.1, B and $P^{-1}AP$ may not be essentially the same, despite having the same bases and the same type.

EXAMPLE 2 Let us here exhibit all possible types of idempotent 4×4 stochastic matrices of rank 2 or 3.

Rank 2

i. Basis $T = \emptyset,\ C_1 = \{1\},\ C_2 = \{2,3,4\}$:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & a & b & 1-a-b \\ 0 & a & b & 1-a-b \\ 0 & a & b & 1-a-b \end{pmatrix} \text{ where } a, b, a+b \in (0,1). \qquad (2.\ 3)$$

ii. Basis $T = \emptyset,\ C_1 = \{1,2\},\ C_2 = \{3,4\}$:

$$\begin{pmatrix} a & 1-a & 0 & 0 \\ a & 1-a & 0 & 0 \\ 0 & 0 & b & 1-b \\ 0 & 0 & b & 1-b \end{pmatrix} \text{ where } a, b \in (0,1). \qquad (2.\ 4)$$

iii. Basis $T = \{1\},\ C_1 = \{2,3\},\ C_2 = \{4\}$:

$$\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & a & 1-a & 0 \\ 0 & a & 1-a & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \text{ where } a \in (0,1). \qquad (2.\ 5)$$

iv. Basis $T = \{1,2\},\ C_1 = \{3\},\ C_2 = \{4\}$:

$$\begin{pmatrix} 0 & 0 & a & 1-a \\ 0 & 0 & b & 1-b \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \text{ where } a, b \in [0,1]. \qquad (2.\ 6)$$

Note that (2. 6) gives rise to nine different types according to whether $a, b \in (0,1)$, $a = 0$ and $b \in (0,1)$, $a = 1$ and $b \in (0,1)$, $b = 0$ and $a \in (0,1)$, $b = 1$ and $a \in (0,1)$, $(a,b) = (0,0)$, $(a,b) = (1,0)$, $(a,b) = (0,1)$ and $(a,b) = (1,1)$.

Rank 3

i. Basis $T = \emptyset,\ C_1 = \{1, 2\},\ C_2 = \{3\},\ C_3 = \{4\}$:

$$\begin{pmatrix} a & 1-a & 0 & 0 \\ a & 1-a & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \text{ where } a \in (0, 1), \tag{2. 7}$$

ii. Basis $T = \{1\},\ C_1 = \{2\},\ C_2 = \{3\},\ C_3 = \{4\}$ and none of the $T \times C_j,\ j = 1, 2, 3,$ block of a matrix of this type is 0 or 1 :

$$\begin{pmatrix} 0 & a & b & 1-a-b \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \text{ where } a, b, a+b \in (0, 1), \tag{2. 8}$$

iii. Same basis as Type (ii), but exactly one of the $T \times C_j,\ j = 1, 2, 3,$ block of a matrix of this type is 1; there are three different types:

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \text{ or } \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{2. 9}$$

iv. Same basis as Type (ii), but exactly one of the $T \times C_j,\ j = 1, 2, 3,$ block of a matrix of this type is 0; here also there are three different types:

$$\begin{pmatrix} 0 & 0 & a & 1-a \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & a & 0 & 1-a \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \text{ or } \begin{pmatrix} 0 & a & 1-a & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{2. 10}$$

where $a \in (0, 1)$.

EXAMPLE 3 Let us now consider the general form of an idempotent $d \times d$ stochastic matrix e of rank 2. Let $\{T, C_1, C_2\}$ be a basis of e. Let $k = |T|$, $c_1 = |C_1|$, and $c_2 = |C_2|$, where $k + c_1 + c_2 = d$. According to Theorem 2.1, e must have the block form

$$e = \begin{pmatrix} 0 & A' & B' \\ 0 & A & 0 \\ 0 & 0 & B \end{pmatrix}, \tag{2. 11}$$

where the top left zero block is $k \times k$, and A (resp. B) is a $c_1 \times c_1$ (resp. $c_2 \times c_2$) strictly positive stochastic matrix with identical rows, each equal to $A_1 = (a_1, a_2, \ldots, a_{c_1})$ (resp. $B_1 = (b_1, b_2, \ldots, b_{c_2})$). Since e has rank 2, each of the first k rows of e must be a linear combination of A_1 and B_1. In other words, the $k \times c_1$ matrix A' and the $k \times c_2$ matrix B' must have the form

$$A' = \begin{pmatrix} r_1 A_1 \\ r_2 A_1 \\ \vdots \\ r_k A_1 \end{pmatrix} \tag{2. 12}$$

and

$$B' = \begin{pmatrix} (1 - r_1) B_1 \\ (1 - r_2) B_1 \\ \vdots \\ (1 - r_k) B_1 \end{pmatrix}, \tag{2. 13}$$

for some constants $r_1, r_2, \ldots, r_k \in [0, 1]$.

In the sequel, we will also need the following result concerning finite groups of $d \times d$ stochastic matrices of common rank.

THEOREM 2.3 *Let G be a finite group of $d \times d$ stochastic matrices of common rank p. Let $\{T, C_1, C_2, \ldots, C_p\}$ be the basis of the identity of G. Then there an isomorphism from G to a subgroup of the group S_p of permutations on $\{1, 2, \ldots, p\}$ such that if π is the isomorphic image of $A \in G$, then the $C_j \times C_k$ block of A is an all-zero block whenever $\pi(j) \neq k$.*

Proof. See [M2].

∎

EXAMPLE 4 In Theorem 2.3, consider the case when $p = 2$ in the general $d \times d$ situation. The isomorphism from G to S_2 has e given by (2. 11) in Example 3 as the isomorphic preimage of the identity permutation. Here, we calculate the $d \times d$ stochastic matrix preimage y of the permutation (12) under this isomorphism.

Let us follow the definitions and notations in Example 3. Then, immediately from Theorem 2.3, we know that the $C_1 \times C_1$ and $C_2 \times C_2$ blocks of y must be zero blocks. Further, from the equation $ye = y$, it follows that the first k columns of y must be zero columns. Write

$$y = \begin{pmatrix} 0 & U & V \\ 0 & 0 & W \\ 0 & Z & 0 \end{pmatrix},$$

where U is $k \times c_1$, V is $k \times c_2$, W is $c_1 \times c_2$, and Z is $c_2 \times c_1$, with W and Z both stochastic. The equations $ye = y$, $ey = y$, and $y^2 = e$ then translate to the matrix equations

$$UA = U,\ VB = V,\ WB = W,\ ZA = Z, \tag{2. 14}$$

$$B'Z = U,\ A'W = V,\ AW = W,\ BZ = Z, \tag{2. 15}$$

and

$$VZ = A',\ UW = B',\ WZ = A,\ ZW = B, \tag{2. 16}$$

respectively. Since both A and B have identical rows, and therefore constant columns, the last two equations in (2. 15) imply that W and Z do, too. This information, together with the last two equations in (2. 14), completely determine W and Z :

$$W = \begin{pmatrix} B_1 \\ B_1 \\ \vdots \\ B_1 \end{pmatrix} \ (c_1\text{-many rows}) \ \text{ and } \ Z = \begin{pmatrix} A_1 \\ A_1 \\ \vdots \\ A_1 \end{pmatrix} \ (c_2\text{-many rows}).$$

The first equation in (2. 15) then implies that the entry in the jth row and lth column of U, where $j \in T$ and $l \in C_1$, is

$$\sum_{s \in C_2} (1 - r_j) b_s a_l = (1 - r_j) a_l,$$

while the second equation in (2. 15) implies that the entry in the jth row and lth column of V, where $j \in T$ and $l \in C_2$, is

$$\sum_{s \in C_1} r_j a_s b_l = r_j b_l.$$

Thus, y has the block form

$$y = \left(\begin{array}{c|c|c} & (1-r_1)A_1 & r_1 B_1 \\ 0 & (1-r_2)A_1 & r_2 B_1 \\ & \vdots & \vdots \\ & (1-r_k)A_1 & r_k B_1 \\ \hline & & B_1 \\ & & B_1 \\ 0 & 0 & \vdots \\ & & B_1 \\ \hline & A_1 & \\ & A_1 & \\ 0 & \vdots & 0 \\ & A_1 & \end{array} \right). \qquad (2.\ 17)$$

Next we present the following general theorem which forms the basis for the method presented in [DM]:

THEOREM 2.4 *Let S be a locally compact second countable Hausdorff semigroup and μ be a probability measure on the Borel subsets of S. Suppose that*

$$S = \overline{\bigcup_{n=1}^{\infty} S_\mu^n},$$

where S_μ is the support of μ.

Suppose that the sequence $\{\mu^n : n \geq 1\}$ is a tight sequence; that is, given $\varepsilon > 0$, there is a compact set K_ε such that for all $n \geq 1$, $\mu^n(K_\varepsilon) > 1 - \varepsilon$. Then the sequence $(1/n)\sum_{k=1}^{n} \mu^k$ converges weakly to a probability measure ν, where S_ν, the support of ν, is the kernel K of S. The group factor G of K (which is completely simple) is compact. The sequence μ^n converges weakly to ν iff there does not exist a subgroup H of K such that the following conditions hold:

1. *H is a normal subgroup of the group $eKe \equiv G$, where e is the identity of H,*
2. *$YX \subset H$, where Y is the set of all idempotents in Ke and X is the set of all idempotent elements in eK,*
3. *$eS_\mu e \subset gH$ for some $g \in G \setminus H$.*

Proof. See Theorem 2.1 in [LM].

■

We note that when S_μ is contained in a set of $d \times d$ stochastic matrices, which is the case of interest to us, the most important assertion in Theorem 2.4 relevant to our problem is the following: the sequence μ^n does not converge weakly if and only if there exists an idempotent e in the kernel K of S such that

$$eS_\mu e \subset gH, \tag{2. 18}$$

for some $g \in eKe \setminus H$ and some proper normal subgroup H of eKe.

We remark that Dhar and Mukherjea's solution in [DM] is essentially accomplished by translating (2. 18) into a set of conditions on S_μ by considering the possibilities for the kernel K and the proper normal subgroups of the corresponding compact group eKe.

For emphasis, we record the following information which follows when Theorem 2.4 is applied to the case under investigation. The kernel K, structurewise, is a completely simple semigroup; in other words, K is topologically isomorphic to (that is, can be identified with) the product $X \times G \times Y$, where G is a finite group and $G = eKe$, where e is some fixed idempotent matrix in K, X is a left-zero semigroup consisting of all idempotent matrices in

Ke, Y is a right-zero semigroup consisting of all idempotent matrices in eK, and $YX \subset G$. The multiplication in $X \times G \times Y$ is given by: $(x, g, y)(x', g', y') = (x, g(yx')g', y')$. Details of these are also given in [HMu].

As mentioned in Section 2.1, it is well-known that K consists of all matrices in S which have the minimal rank. Moreover, if this minimal rank is one (that is, the matrices in the kernel K have identical rows), then Lemma 2.1 below says that the sequence μ^n always converges weakly.

LEMMA 2.1 *Let μ be a probability measure with support S_μ inside a set of $d \times d$ stochastic matrices. Let K be the kernel of the closed semigroup S generated by S_μ. If the common rank of the matrices in K is equal to 1, then the sequence μ^n converges weakly.*

Proof. It is easy to verify that for any two stochastic matrices A and B in K, $AB = B$. Notice that since S is compact, the averaged sequence $(1/n)\sum_{k=1}^{n} \mu^k$ always converges to some probability measure ν, whose support S_ν is actually equal to K. Further, for any open set G containing K, $\mu^n(G) \to 1$ as $n \to \infty$. This means that if ν' is another weak limit point of (μ^n), then its support $S_{\nu'}$ is inside K. Since $\nu * \mu = \nu$, it follows that $\nu * (\mu^n) = \nu$ for each n, so $\nu * \nu' = \nu$. But $\nu * \nu' = \nu'$ since $AB = B$ for any two matrices $A, B \in K$. Thus, $\nu' = \nu$, and so μ^n converges weakly to ν.

■

Also, when the rank of the matrices in K is d (that is, when they all have full rank), then K happens to be a compact group, and in this case, $K = S$ consists of $d \times d$ invertible matrices, and as such, the question of weak convergence of μ^n can be easily resolved using well-known classical results. Thus, we will only need to look into the cases when K consists of matrices with rank r, where r is strictly between 1 and d.

For the special case when $YX = G$ in Theorem 2.4, we have the following:

COROLLARY 2.1.1 *Let μ be a probability measure on the Borel subsets of $d \times d$ stochastic matrices, and let $S = \overline{\left(\bigcup_{n\geq 1} S_\mu^n\right)}$ be the closed (multiplicative) semigroup generated by the support S_μ of μ. Let K be the kernel of S and let $X \times G \times Y$ be the product representation of K. If $YX = G$, then μ^n converges weakly.*

Proof. Let η be the identity of the group of weak limit points of μ^n. If $YX = G$, then by Theorem 2.4, the support S_η of η has the product representation $X \times G \times Y$, the same as that of S_ν, where $\nu = (\text{w}) \lim_{n\to\infty}(1/n)\sum_{k=1}^{n}\mu^k$. We also have

$$\eta * \nu = \nu * \eta = \nu, \tag{2. 19}$$

since $\mu * \nu = \nu * \mu = \nu$. By (2. 19) and Proposition 2.5 (page 74, [HMu]) it follows that for any Borel set $B \subset K$ and any $x \in K$,

$$\begin{aligned} \nu\{y : yx \in B\} &= \int \eta\{z : zyx \in B\}\, \nu(dy) \\ &= \eta\{y : yx \in B\}, \end{aligned}$$

and since $\nu = \nu * \eta$ and $\eta = \eta * \eta$, we also have

$$\begin{aligned} \nu(B) &= \int \nu\{y : yx \in B\}\, \eta(dx) \\ &= \int \eta\{y : yx \in B\}\, \eta(dx) \\ &= \eta * \eta(B) = \eta(B). \end{aligned}$$

Thus, $\nu = \eta$. It follows that

$$\mu * \eta = \eta * \mu = \eta. \tag{2. 20}$$

This means that for any weak limit point ν' of μ^n, $\nu' = \nu' * \eta = \eta * \nu' = \eta$. In other words, whenever $YX = G$, μ^n converges weakly to η.

■

Next we define cyclicity following [CR] (see page 169, at the end of the paper). Let $A_1, A_2, \ldots, A_k$ be pairwise disjoint subsets of $\{1, 2, \ldots, d\}$ so that $\cup_{i=1}^{k} A_i$ may or may not equal $\{1, 2, \ldots, d\}$. Then S_μ is called cyclic with respect to $\{A_1, A_2, \ldots, A_k\}$ if, for each x in S_μ with $A_{k+i} \equiv A_i$, $1 \le i \le k$, we have

$$\sum_{j\in A_{m+1}} x_{ij} = 1,\ i \in A_m,\ 1 \le m \le k. \tag{2. 21}$$

This definition immediately gives us the following:

LEMMA 2.2 *Let μ be a probability measure on the Borel subsets of $d \times d$ stochastic matrices, and let S_μ be the support of μ. If S_μ is cyclic, then μ^n does not converge weakly.*

Proof. Define the sets C and D as

$$C = \{x \in S \,:\, \sum_{j \in A_1} x_{ij} = 1 \text{ for each } i \in A_2\}$$

and

$$D = \{x \in S \,:\, \sum_{j \in A_2} x_{ij} = 1 \text{ for each } i \in A_1\},$$

where A_1 and A_2 are as they appear in the definition of cyclicity above, then C and D are disjoint compact subsets of S, and furthermore, for each $n \geq 1$,

$$\mu^{nk}(C) = 1 \text{ and } \mu^{nk+1}(D) = 1.$$

If μ^n converges weakly to λ as $n \to \infty$, then clearly $\lambda(C) = 1$ as well as $\lambda(D) = 1$. But this is impossible. Thus, cyclicity of S_μ implies non-weak convergence of μ^n.

■

In the next section, we explicitly solve the problem of weak convergence in 4×4 stochastic matrices following Dhar and Mukherjea's method in [DM]. As pointed out at the end of the introduction to the present chapter, one reason for this calculation is to verify whether Chakraborty and Rao's characterization, in the 3×3 case, of non-weak convergence of the sequence μ^n in terms of cyclicity of the support S_μ is still valid in the 4×4 case.

2.3 4×4 Stochastic Matrices

In this section, μ is a probability measure with support S_μ inside a set of 4×4 stochastic matrices, S is the multiplicative semigroup generated by S_μ, and K is the kernel of S. Our aim here is to use (2. 18) in Section 2.2 to find a necessary and sufficient condition on S_μ in order for the sequence μ^n not to converge weakly.

Suppose, then, that in everything that follows, μ^n does not converge weakly.

First, let the common rank of the matrices in K be 2. According to Theorem 2.3, the compact group eKe of Theorem 2.4 must then be isomorphic to the two-element symmetric group $S_2 = \{(1), (12)\}$. Here, e is one of the idempotent 4×4 stochastic matrices of rank 2 given by (2. 3), (2. 4), (2. 5), or (2. 6) in Example 2.

Suppose e has the form displayed in (2. 3). That is,

$$e = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & a & b & 1-a-b \\ 0 & a & b & 1-a-b \\ 0 & a & b & 1-a-b \end{pmatrix}$$

where $a, b, a+b \in (0,1)$. Since the C-classes of e are $C_1 = \{1\}$ and $C_2 = \{2,3,4\}$, the $C_1 \times C_1$ and $C_2 \times C_2$ blocks of the stochastic matrix $A \in eKe$ corresponding to the permutation $\pi = (12)$ are zero blocks, according to Theorem 2.3. Therefore, A must have the form

$$\begin{pmatrix} 0 & s & t & 1-s-t \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

for some $s, t \in [0,1]$. The equation $A^2 = e$ then shows that $s = a$ and $t = b$, giving

$$A = \begin{pmatrix} 0 & a & b & 1-a-b \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}. \tag{2. 22}$$

The only proper normal subgroup of $eKe = \{e, A\}$ is the trivial subgroup $\{e\}$, and consequently its coset is the singleton $\{A\}$. Thus, the condition (2. 18) translates to the equation

$$exe = A, \tag{2. 23}$$

where

$$x = \begin{pmatrix} a_1 & a_2 & a_3 & 1-a_1-a_2-a_3 \\ b_1 & b_2 & b_3 & 1-b_1-b_2-b_3 \\ c_1 & c_2 & c_3 & 1-c_1-c_2-c_3 \\ d_1 & d_2 & d_3 & 1-d_1-d_2-d_3 \end{pmatrix} \tag{2. 24}$$

is an arbitrary element of the support S_μ. We note that in (2. 24), all variables are nonnegative and at most equal to 1. (2. 23) then leads to the equations

$$a_1 = 0$$

and

$$ab_1 + bc_1 + (1 - a - b)d_1 = 1. \tag{2. 25}$$

It follows from (2. 25) that $b_1 = c_1 = d_1 = 1$, leading to

$$x = \begin{pmatrix} 0 & s & t & 1-s-t \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

We conclude that under the condition of non-weak convergence of the sequence μ^n, and with e given by (2. 3), the support S_μ must satisfy

$$S_\mu \subset \left\{ \begin{pmatrix} 0 & s & t & 1-s-t \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} : s, t, s+t \in [0, 1] \right\}. \tag{2. 26}$$

Recalling Chakraborty and Rao's definition of cyclicity, we conclude that S_μ in (2. 26) is cyclic with respect to the disjoint subsets $\{1\}$ and $\{2, 3, 4\}$ of $\{1, 2, 3, 4\}$.

In the following, when describing a condition of set inclusion that S_μ must satisfy for non-weak convergence of the sequence μ^n, as in (2. 26), if doing so will not cause confusion we will omit the description of the variables used, with the understanding that they are all non-negative and at most 1, and such that the matrix they form is stochastic.

Similar calculations show that if e is given by (2. 4), that is,

$$e = \begin{pmatrix} a & 1-a & 0 & 0 \\ a & 1-a & 0 & 0 \\ 0 & 0 & b & 1-b \\ 0 & 0 & b & 1-b \end{pmatrix},$$

where $a, b, \in (0,1)$, then S_μ must satisfy

$$S_\mu \subset \left\{ \begin{pmatrix} 0 & 0 & s & 1-s \\ 0 & 0 & t & 1-t \\ u & 1-u & 0 & 0 \\ v & 1-v & 0 & 0 \end{pmatrix} \; : \; s, t, u, v \in [0,1] \right\}, \tag{2. 27}$$

in which case S_μ is cyclic with respect to $\{1,2\}$ and $\{3,4\}$.

Let us next consider the case when e is given by (2. 5), that is,

$$e = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & a & 1-a & 0 \\ 0 & a & 1-a & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

where $a \in (0,1)$. In this case, the C-classes are $C_1 = \{2,3\}$ and $C_2 = \{4\}$, so that if $eKe = \{e, A\}$, then

$$A = \begin{pmatrix} 0 & s & 1-s & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & a & 1-a & 0 \end{pmatrix}, \tag{2. 28}$$

where $s \in [0,1]$. It then follows that elements in S_μ must be of the form

$$\begin{pmatrix} s_1 & s_2 & s_3 & 1{-}s_1{-}s_2{-}s_3 \\ t & 0 & 0 & 1-t \\ u & 0 & 0 & 1-u \\ 0 & v & 1-v & 0 \end{pmatrix}, \tag{2. 29}$$

which may be simplified further, as follows. Pick an element y of the form given by (2. 29). Then with e as above,

$$ey^2e = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & * & * & (1-(s_2{+}s_3))(at{+}(1{-}a)u) \\ 0 & * & * & (1-(s_2{+}s_3))(at{+}(1{-}a)u) \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

By inspection, this element cannot be A in (2. 28), so it must be e itself, implying that

$$(1-(s_2+s_3))\,(at+(1-a)u)=0.$$

Thus, if there exists $y \in S_\mu$ with at least one of $t, u > 0$, then $s_2 + s_3 = 1$ (since $a \in (0,1)$) and accordingly S_μ must satisfy

$$S_\mu \subset \left\{ \begin{pmatrix} 0 & s & 1-s & 0 \\ t & 0 & 0 & 1-t \\ u & 0 & 0 & 1-u \\ 0 & v & 1-v & 0 \end{pmatrix} : \begin{array}{l} s, v \in [0,1] \text{ and either } t \\ \text{or } u \text{ or both are positive} \end{array} \right\}, \tag{2. 30}$$

otherwise $t = u = 0$ and S_μ must satisfy

$$S_\mu \subset \left\{ \begin{pmatrix} s_1 & s_2 & s_3 & 1-s_1-s_2-s_3 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & v & 1-v & 0 \end{pmatrix} \right\}. \tag{2. 31}$$

Note that in (2. 30), S_μ is cyclic with respect to $\{1,4\}$ and $\{2,3\}$, and in (2. 31), S_μ is cyclic with respect to $\{2,3\}$ and $\{4\}$.

To finish the rank 2 case, we consider the situation when e is given by (2. 6), that is,

$$e = \begin{pmatrix} 0 & 0 & a & 1-a \\ 0 & 0 & b & 1-b \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

where $a, b \in [0,1]$. Here, the appropriate condition on S_μ depends on a and b. We give all possibilities below, and the corresponding condition on S_μ.

- $a, b \in (0,1)$:

$$S_\mu \subset \left\{ \begin{pmatrix} s_1 & s_2 & s_3 & 1-s_1-s_2-s_3 \\ t_1 & t_2 & t_3 & 1-t_1-t_2-t_3 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \right\}, \tag{2. 32}$$

and S_μ is cyclic with respect to $\{3\}, \{4\}$.

- $a = 0$ and $b \in (0, 1)$:

  Either S_μ satisfies

$$S_\mu \subset \left\{ \begin{pmatrix} 0 & 0 & 1 & 0 \\ t_1 & t_2 & t_3 & 1{-}t_1{-}t_2{-}t_3 \\ u & 0 & 0 & 1-u \\ 0 & 0 & 1 & 0 \end{pmatrix} \right\} \tag{2. 33}$$

  for some fixed $u > 0$, or else S_μ satisfies the same condition as (2. 32). Note that S_μ in (2. 33) is cyclic with respect to $\{1, 4\}, \{3\}$.

- $a = 1$ and $b \in (0, 1)$:

  Either S_μ satisfies

$$S_\mu \subset \left\{ \begin{pmatrix} 0 & 0 & 0 & 1 \\ t_1 & t_2 & t_3 & 1{-}t_1{-}t_2{-}t_3 \\ 0 & 0 & 0 & 1 \\ u & 0 & 1-u & 0 \end{pmatrix} \right\} \tag{2. 34}$$

  for some fixed $u > 0$, or else S_μ satisfies the same condition as (2. 32). Note that S_μ in (2. 34) is cyclic with respect to $\{1, 3\}, \{4\}$.

- $a \in (0, 1)$ and $b = 0$:

  Either S_μ satisfies

$$S_\mu \subset \left\{ \begin{pmatrix} s_1 & s_2 & s_3 & 1{-}s_1{-}s_2{-}s_3 \\ 0 & 0 & 1 & 0 \\ 0 & u & 0 & 1-u \\ 0 & 0 & 1 & 0 \end{pmatrix} \right\} \tag{2. 35}$$

  for some fixed $u > 0$, or else S_μ satisfies the same condition as (2. 32). Note that S_μ in (2. 35) is cyclic with respect to $\{2, 4\}, \{3\}$.

- $a \in (0,1)$ and $b = 1$:

  Either S_μ satisfies

$$S_\mu \subset \left\{ \begin{pmatrix} s_1 & s_2 & s_3 & 1-s_1-s_2-s_3 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & u & 1-u & 0 \end{pmatrix} \right\} \tag{2. 36}$$

  for some fixed $u > 0$, or else S_μ satisfies the same condition as (2. 32). Note that S_μ in (2. 36) is cyclic with respect to $\{2,3\}, \{4\}$.

- $a = b = 0$:

  Either S_μ satisfies

$$S_\mu \subset \left\{ \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ u & v & 0 & 1-u-v \\ 0 & 0 & 1 & 0 \end{pmatrix} \right\} \tag{2. 37}$$

  for some fixed $u, v > 0$, or else S_μ satisfies the same condition as (2. 32), (2. 33), or (2. 35). Note that S_μ in (2. 37) is cyclic with respect to $\{1,2,4\}, \{3\}$.

- $a = 1,\ b = 0$:

  Either S_μ satisfies

$$S_\mu \subset \left\{ \begin{pmatrix} 0 & s & 0 & 1-s \\ t & 0 & 1-t & 0 \\ 0 & u & 0 & 1-u \\ 0 & 0 & 1 & 0 \end{pmatrix} \right\} \tag{2. 38}$$

  for fixed $u, t > 0$, or S_μ satisfies

$$S_\mu \subset \left\{ \begin{pmatrix} 0 & s & 0 & 1-s \\ t & 0 & 1-t & 0 \\ 0 & 0 & 0 & 1 \\ v & 0 & 1-v & 0 \end{pmatrix} \right\} \tag{2. 39}$$

for fixed $s, v > 0$, or S_μ satisfies

$$S_\mu \subset \left\{ \begin{pmatrix} 0 & s & 0 & 1-s \\ t & 0 & 1-t & 0 \\ 0 & u & 0 & 1-u \\ v & 0 & 1-v & 0 \end{pmatrix} \right\}. \tag{2. 40}$$

for fixed $u, v > 0$, or else S_μ satisfies the same condition as (2. 32), (2. 34), or (2. 35). Note that S_μ in (2. 38), (2. 39), or (2. 40) is cyclic with respect to $\{1, 3\}, \{2, 4\}$.

- $a = 0,\ b = 1:$

Either S_μ satisfies

$$S_\mu \subset \left\{ \begin{pmatrix} 0 & s & 1-s & 0 \\ t & 0 & 0 & 1-t \\ u & 0 & 0 & 1-u \\ 0 & 0 & 1 & 0 \end{pmatrix} \right\} \tag{2. 41}$$

for fixed $s, u > 0$, or S_μ satisfies

$$S_\mu \subset \left\{ \begin{pmatrix} 0 & s & 1-s & 0 \\ t & 0 & 0 & 1-t \\ 0 & 0 & 0 & 1 \\ 0 & v & 1-v & 0 \end{pmatrix} \right\} \tag{2. 42}$$

for fixed $t, v > 0$, or S_μ satisfies

$$S_\mu \subset \left\{ \begin{pmatrix} 0 & s & 1-s & 0 \\ t & 0 & 0 & 1-t \\ u & 0 & 0 & 1-u \\ 0 & v & 1-v & 0 \end{pmatrix} \right\}. \tag{2. 43}$$

for fixed $u, v > 0$, or else S_μ satisfies the same condition as (2. 32), (2. 33), or (2. 36). Note that S_μ in (2. 41), (2. 42), or (2. 43) is cyclic with respect to $\{1, 4\}, \{2, 3\}$.

- $a = b = 1$:

  Either S_μ satisfies

$$S_\mu \subset \left\{ \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ u & v & 1-u-v & 0 \end{pmatrix} \right\}, \tag{2. 44}$$

  for some fixed $u, v > 0$, or else S_μ satisfies the same condition as (2. 32), (2. 34), or (2. 36). Note that S_μ in (2. 44) is cyclic with respect to $\{1, 2, 3\}, \{4\}$.

This completes the rank 2 case. We point out that in all cases, μ^n not weakly convergent implies S_μ is cyclic.

We now assume that the common rank of the matrices in the kernel K is 3. In this case, the group eKe, where e is a 4×4 idempotent stochastic matrix of rank 3, is isomorphic to a subgroup of the symmetric group S_3 of permutations on $\{1, 2, 3\}$. Note that the nontrivial proper subgroups of S_3 are the three two-element subgroups each consisting of the identity permutation and a transposition, together with the normal subgroup consisting of the even permutations. It is therefore clear that the condition which S_μ must satisfy, under the same assumption that the sequence μ^n does not converge weakly, is

$$eS_\mu e \subset \{M_{(12)}, M_{(13)}, M_{(23)}\} \tag{2. 45}$$

(assuming $eS_\mu e$ has more than one element), where M_τ is the stochastic matrix pre-image of the transposition $\tau \in \{(12), (13), (23)\}$, under the isomorphism of Theorem 2.3.

As was done in the rank 2 case, we let

$$x = \begin{pmatrix} a_1 & a_2 & a_3 & 1-a_1-a_2-a_3 \\ b_1 & b_2 & b_3 & 1-b_1-b_2-b_3 \\ c_1 & c_2 & c_3 & 1-c_1-c_2-c_3 \\ d_1 & d_2 & d_3 & 1-d_1-d_2-d_3 \end{pmatrix}$$

be an arbitrary element of the support S_μ, and solve the equation

$$exe = M_\tau \tag{2. 46}$$

arising from (2. 45), for each of the possibilities for e given by (2. 7), (2. 8), (2. 9), or (2. 10), and the corresponding stochastic matrices M_τ, $\tau = (12), (13), (23)$.

We show the details of the calculations involved using one particular type of a 4×4 idempotent stochastic matrix e of rank 3, namely the first matrix displayed in (2. 9). That is,

$$e = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

From Theorem 2.3, the stochastic matrices $M_{(12)}, M_{(13)}$, and $M_{(23)}$ in this case are given by

$$\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \text{ and } \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

respectively.

Then it follows from the equality $exe = M_{(12)}$ (after computations) that x must be of the form

$$\begin{pmatrix} s & t & u & v \\ 0 & 0 & 1 & 0 \\ w & 1-w & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

where every entry is nonnegative with each row sum equal to 1. Let us call a typical element of this form x_1. Noting that the element ex_1^2e is an element in eKe, and after computations looking at its form, it is clear that this element must be the element e, whence it follows (after computations) that x_1 must be one of the following two possible forms:

$$\begin{pmatrix} s' & t' & u' & v' \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \text{ or } \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ c & 1-c & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

where again the entries are nonnegative with each row sum 1.

Let us call a typical element of the first form y and a typical element of the second form z. Noting that if there exists a z with $c > 0$, then since $e(zy)e$ must again equal e, it follows after computations that in y, $c' = 1$ if there is a z in S_μ with $c > 0$. In other words, in the case when $exe = M_{(12)}$, the form of x must be that of z above where $c \geq 0$. By also considering the equations $exe = M_{(13)}$ and $exe = M_{(23)}$, and going through similar arguments and computations, we see that S_μ must be contained in the set

$$\left\{ \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ c & 1-c & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ d & 1-d & 0 & 0 \end{pmatrix}, \begin{pmatrix} a & 1-a & 0 & 0 \\ b & 1-b & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \right\}. \quad (2.\ 47)$$

Note that S_μ in this case cannot be cyclic if it has more than one element.

The corresponding conditions for S_μ when e is given by the other two matrices displayed in (2. 9) are

$$S_\mu \subset \left\{ \begin{pmatrix} 0 & 1 & 0 & 0 \\ f & 0 & 1-f & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} u & 0 & 1-u & 0 \\ 0 & 0 & 0 & 1 \\ g & 0 & 1-g & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ h & 0 & 1-h & 0 \end{pmatrix} \right\} \quad (2.\ 48)$$

and

$$S_\mu \subset \left\{ \begin{pmatrix} r & 0 & 0 & 1-r \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ f & 0 & 0 & 1-f \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 & 0 \\ g & 0 & 0 & 1-g \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ h & 0 & 0 & 1-h \\ 0 & 0 & 1 & 0 \end{pmatrix} \right\}, \quad (2.\ 49)$$

respectively.

Next, when e has the form given by (2. 7), that is,

$$e = \begin{pmatrix} a & 1-a & 0 & 0 \\ a & 1-a & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

where $a \in (0,1)$, then S_μ must satisfy

$$S_\mu \subset \left\{ \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ u & 1-u & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ v & 1-v & 0 & 0 \end{pmatrix}, \begin{pmatrix} s & 1-s & 0 & 0 \\ t & 1-t & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \right\} \tag{2. 50}$$

and when e is given by (2. 8), that is,

$$e = \begin{pmatrix} 0 & a & b & 1-a-b \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

where $a, b, a+b \in (0,1)$, the condition for S_μ is

$$S_\mu \subset \left\{ \begin{pmatrix} r & s & t & 1-r-s-t \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} u & v & w & 1-u-v-w \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} x & y & z & 1-x-y-z \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \right\}. \tag{2. 51}$$

Finally, when e takes each of the possible values in (2. 10), the same condition given in (2. 51) is obtained.

This finishes our analysis of the rank 2 and rank 3 cases. In view of Lemma 2.1 in Section 2.2 and the remarks following its proof, these are the only cases that need to be considered in studying problem of weak convergence in 4×4 stochastic matrices. We therefore have the following:

THEOREM 2.5 *Let μ be a probability measure on the Borel subsets of 4×4 stochastic matrices such that the minimal rank of the matrices in the kernel K of the closed semigroup generated by the support S_μ of μ is 2. Then the sequence of convolution powers μ^n does not converge weakly iff S_μ is cyclic. If the common rank of the matrices in K is 3, then cyclicity of S_μ is not necessary for non-weak convergence of μ^n.*

Proof. Follows from the preceding discussion.

■

We now present a couple of examples.

EXAMPLE 5 Consider a probability measure μ with support S_μ given by $S_\mu = \{A, B\}$, where

$$A = \begin{pmatrix} 0 & 0 & 1 & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 1 & 0 & 0 \end{pmatrix} \text{ and } B = \begin{pmatrix} 0 & 1 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 \end{pmatrix}. \quad (2.\ 52)$$

Then the semigroup S generated by S_μ has 4 elements and is given by

$$S = \{e, A, B, C\}$$

where A and B are given in (2. 52),

$$e = \begin{pmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{pmatrix} \text{ and } C = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}.$$

The multiplication table of S is given by

	e	A	B	C
e	e	C	C	C
A	C	e	e	e
B	C	e	e	e
C	C	e	e	e

It is clear from the multiplication table that the kernel K is $\{e, C\}$, which consists precisely of the matrices of the minimal rank 2. Moreover, $K = eKe$ is a two-element group, and, as such, only has the trivial subgroup $\{e\}$ as its normal subgroup. The coset $\{C\}$ is exactly the set $eS_\mu e$, and therefore the sequence μ^n does not converge weakly. Notice that S_μ is cyclic with respect to the partition $\{1, 4\}$ and $\{2, 3\}$ of $\{1, 2, 3, 4\}$.

EXAMPLE 6 Consider a two-point probability measure μ with support S_μ given by $S_\mu = \{A, B\}$, where

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \text{ and } B = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \tag{2. 53}$$

Then it can be verified that the semigroup S generated by S_μ has 8 elements and is given by

$$S = \{e, A, B, C, D, E, F, I\},$$

where A and B are given in (2. 53),

$$e = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

$$C = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad D = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

$$E = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad F = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

and I is the 4×4 identity matrix. The multiplication table of S is given by

	e	A	B	C	D	E	F	I
e	e	E	B	C	D	E	F	e
A	E	I	D	F	B	e	C	A
B	B	C	e	E	F	C	D	B
C	C	B	F	D	e	B	E	C
D	D	F	E	e	C	F	B	D
F	E	e	D	F	B	e	C	E
F	F	D	C	B	E	D	e	F
I	e	A	B	C	D	E	F	I

Except for A and I, both of which are of rank 4, all the matrices in S have rank 3. The kernel K is therefore

$$K = \{e, B, C, D, E, F\}.$$

Notice that $K = eKe$ is a group isomorphic to the symmetric group S_3 on $\{1, 2, 3\}$, and the coset $\{B, E, F\}$ of the normal subgroup $\{e, C, D\}$ contains

$$eS_\mu e = \{B, E\}.$$

Thus, the sequence μ^n does not converge weakly. Notice also that S_μ is not cyclic.

2.4 The $d \times d$ case

Let us now attempt to solve the present problem in the general $d \times d$ case. Thus, we let μ be a probability measure with support S_μ inside a set of $d \times d$ stochastic matrices, S the multiplicative semigroup generated by S_μ, and K the kernel of S.

The first step is to determine what happens when the rank of the matrices in K is 2. In this case, the group $G = eKe$, where e is a fixed idempotent in K, and $X \times G \times Y$ is a product representation of K (as described in the remarks after Theorem 2.4 of Section 2.2), has one or two elements. In case G has only one element, then since $YX \subset G$, YX must indeed equal G, and by Corollary 2.1.1, the sequence μ^n must then converge weakly.

Thus, if we assume that μ^n does not converge weakly, then the group $G = eKe$ has two elements $\{e, y\}$, where e is a $d \times d$ idempotent stochastic matrix that has the general form calculated in Example 3 of Section 2.2, and y is the $d \times d$ stochastic matrix satisfying $ey = ye = y$ and $y^2 = e$. That is, e has basis $\{T, C_1, C_2\}$, where T, C_1, and C_2 have, respectively, k, c_1, and c_2 elements with $k + c_1 + c_2 = d$, and

$$e = \begin{pmatrix} 0 & A' & B' \\ 0 & A & 0 \\ 0 & 0 & B \end{pmatrix}.$$

In this block form, the top left zero block is $k \times k$, the strictly positive rank one stochastic matrices A and B are $c_1 \times c_1$ and $c_2 \times c_2$, respectively, A' is $k \times c_1$ and B' is $k \times c_2$. If A has identical rows equal to

$$A_1 = (a_1, a_2, \ldots, a_{c_1})$$

and B has identical rows equal to

$$B_1 = (b_1, b_2, \ldots, b_{c_2}),$$

then A' and B' are given by

$$A' = \begin{pmatrix} r_1 A_1 \\ r_2 A_1 \\ \vdots \\ r_k A_1 \end{pmatrix} \text{ and } B' = \begin{pmatrix} (1-r_1)B_1 \\ (1-r_2)B_1 \\ \vdots \\ (1-r_k)B_1 \end{pmatrix},$$

for some constants $r_1, r_2, \ldots, r_k \in [0, 1]$.

We also know from Example 4 of Section 2.2 that y must then have the form

$$y = \left(\begin{array}{c|c|c} 0 & \begin{matrix} (1-r_1)A_1 \\ (1-r_2)A_1 \\ \vdots \\ (1-r_k)A_1 \end{matrix} & \begin{matrix} r_1 B_1 \\ r_2 B_1 \\ \vdots \\ r_k B_1 \end{matrix} \\ \hline 0 & 0 & \begin{matrix} B_1 \\ B_1 \\ \vdots \\ B_1 \end{matrix} \\ \hline 0 & \begin{matrix} A_1 \\ A_1 \\ \vdots \\ A_1 \end{matrix} & 0 \end{array}\right),$$

where the diagonal blocks are zero blocks of size $k \times k$, $c_1 \times c_1$, and $c_2 \times c_2$, respectively.

We now proceed to determine the block structure of an arbitrary stochastic matrix

$$x = \begin{pmatrix} X_{11} & X_{12} & X_{13} \\ X_{21} & X_{22} & X_{23} \\ X_{31} & X_{32} & X_{33} \end{pmatrix} \tag{2. 54}$$

in S_μ. Here, the diagonal blocks X_{jj}, $j = 1, 2, 3$, are of size $k \times k$, $c_1 \times c_1$, and $c_2 \times c_2$, respectively. (The blocks corresponding to x' in S_μ will accordingly be referred to as X'_{ij}, those for x'' will be referred to as X''_{ij}, etc.)

Since we are assuming non-weak convergence of μ^n, exe must be equal to the non-identity element in the group eKe. In other words,

$$eS_\mu e = \{y\}, \tag{2. 55}$$

which implies, in particular, that

$$A(X_{21}A' + X_{22}A) = 0 \tag{2. 56}$$

and

$$B(X_{31}B' + X_{33}B) = 0. \tag{2. 57}$$

Since both A and B are strictly positive, it follows from (2. 56) and (2. 57) that

$$X_{22} = 0 \tag{2. 58}$$

and

$$X_{33} = 0, \tag{2. 59}$$

which consequently give

$$AX_{21}A' = 0 \tag{2. 60}$$

and

$$BX_{31}B' = 0. \tag{2. 61}$$

Now define the sets

$$\left.\begin{aligned} T_1 &= \{t \in T \,:\, r_t = 0\} \\ T_2 &= \{t \in T \,:\, r_t = 1\} \\ T_3 &= \{t \in T \,:\, r_t \in (0,1)\}. \end{aligned}\right\} \tag{2. 62}$$

For $s, t \in C_1$, the (s,t)-entry of $AX_{21}A'$ in (2. 60) is 0. This entry is

$$\begin{aligned} (AX_{21}A')_{st} &= \sum_{u \in C_1} \sum_{w \in T} a_u (X_{21})_{uw} r_w a_t \\ &= a_t \sum_{u \in C_1} \sum_{w \in T_2 \cup T_3} a_u (X_{21})_{uw} r_w, \end{aligned}$$

implying that $(X_{21})_{uw} = 0$ for all $u \in C_1$ and $w \in T_2 \cup T_3$. Thus,

$$X_{21}|_{C_1 \times (T_2 \cup T_3)} = 0. \tag{2. 63}$$

Similarly, from (2. 61) we obtain

$$X_{31}|_{C_2 \times (T_1 \cup T_3)} = 0. \tag{2. 64}$$

Next, we know that ex^2e must be either e or y. In the 4×4 rank 2 case, as we saw in the preceding section, this is easy to determine by looking at the form of ex^2e and comparing with e or y, finding that in all cases, $ex^2e = e$. In the general case, it is still true that $ex^2e = e$, and, in fact, we have following:

LEMMA 2.3 *Let μ be a probability measure with support S_μ inside a set of $d \times d$ stochastic matrices, S the multiplicative semigroup generated by S_μ, and K the kernel of S. Let the rank of the matrices in K be 2, and let e be a fixed idempotent in K. Assume that μ^n does not converge weakly. Then*

$$\begin{aligned} eS_\mu^{2n}e &= \{e\}, \qquad n \geq 1, \\ eS_\mu^{2n+1}e &= \{y\}, \qquad n \geq 0. \end{aligned} \tag{2. 65}$$

Proof. Let x_1, x_2 be in S_μ. Then the element ex_1 <u>cannot</u> be an idempotent; for if it is, then ex_1, being an element in eK, belongs to Y, which is a right-zero semigroup, and as such, $(ex_1)e = e$, contradicting (2. 55). Similarly, the element x_2e <u>cannot</u> be an idempotent; for if it is, then $e(x_2e) = e$, since x_2e will then be an element in X, a left-zero semigroup. Thus, considering the product representation $X \times G \times Y$ of K, where $YX = \{e\}$ and $G = \{e, y\}$, the elements ex_1 and x_2e, considered as elements in $X \times G \times Y$, have representations given by

$$\begin{aligned} ex_1 &= (x', y, y'), \quad x' \in X, y' \in Y, \\ x_2e &= (x'', y, y''), \quad x'' \in X, y'' \in Y. \end{aligned}$$

Notice that an element of the form (x''', e, y'''), where $x''' \in X$ and $y''' \in Y$, is an idempotent since $YX = \{e\}$. Thus, we have

$$\begin{aligned} ex_1x_2e &= (x', y, y')(x'', y, y'') \\ &= (x', y(y'x'')y, y'') \\ &= (x', e, y''), \end{aligned}$$

which is an idempotent. Since $ex_1x_2e \in eKe = G$, $ex_1x_2e = e$. This proves that $eS_\mu^2e = \{e\}$.

Let us now observe that each element in the set S_μ^2e is an idempotent, since for any x_1, x_2 in S_μ,

$$(x_1x_2e)(x_1x_2e) = x_1x_2(ex_1x_2e) = x_1x_2e,$$

by the previous step. Also, any element in the set yS_μ is an idempotent, since for any $x_3 \in S_\mu$,

$$(yx_3)(yx_3) = (yex_3)(eyx_3) = y(ex_3e)yx_3 = y^3x_3 = yx_3,$$

using (2. 55). Thus, $yx_3 = eyx_3 \in Y$, and $x_1x_2e \in X$. Since $YX = \{e\}$, it follows that for $x_1, x_2, x_3 \in S_\mu$, we have

$$(yx_3)(x_1x_2e) = e,$$

or

$$ex_3x_1x_2e = y.$$

Thus, we have proven:

$$eS_\mu^3e = \{y\}.$$

Now suppose we have proven: $eS_\mu^{2k}e = e$ for $k \leq n$. Then since $(S_\mu^{2n}e)(S_\mu^{2n}e) = S_\mu^{2n}(eS_\mu^{2n}e) = S_\mu^{2n}e$, every element in $S_\mu^{2n}e$ is an idempotent, and therefore, in X. Also, since $(yS_\mu)(yS_\mu) = (ye)S_\mu(ey)S_\mu = y(eS_\mu e)yS_\mu = y^3S_\mu = yS_\mu$, using (2. 55), every element in $yS_\mu = eyS_\mu$ is an idempotent, and therefore, in Y. Hence,

$$(yS_\mu)(S_\mu^{2n}e) \subset YX = \{e\},$$

or

$$eS_\mu^{2n+1}e = \{y\}.$$

Similarly, since $(eS_\mu^2)(eS_\mu^2) = (eS_\mu^2 e)S_\mu^2 = eS_\mu^2$, every element in $eS_\mu^2 \in Y$, and this means that $eS_\mu^{2n+2}e = (eS_\mu^2)(S_\mu^{2n}e) \subset YX = \{e\}$. This proves (2. 65) completely.

■

Now we will use the equation $exx'e = e$ for any two arbitrary elements $x, x' \in S_\mu$ (see (2. 65)). This gives us the equations

$$A(X'_{21}X_{11} + X'_{23}X_{31})B' + AX'_{21}X_{13}B = 0 \tag{2. 66}$$

and

$$B(X'_{31}X_{11} + X'_{32}X_{21})A' + BX'_{31}X_{12}A = 0. \tag{2. 67}$$

Every term in (2. 66) and (2. 67) is separately zero since we are dealing with stochastic matrices. Thus, since A and B are strictly positive, we have

$$X'_{21}X_{13} = 0 \text{ and } X'_{31}X_{12} = 0; \tag{2. 68}$$

$$AX'_{21}X_{11}B' = 0 \text{ and } AX'_{23}X_{31}B' = 0; \tag{2. 69}$$

$$BX'_{31}X_{11}A' \text{ and } BX'_{32}X_{21}A' = 0. \tag{2. 70}$$

Note that in all the equations above the blocks X_{ij} correspond to the element x and the blocks X'_{ij} to x', where x and x' are two arbitrary elements in S_μ.

Let us now define the sets $T_1^* \subset T_1$ and $T_2^* \subset T_2$ as follows:

$$T_1^* = \{t \in T_1 : \text{ there exists } x \in S_\mu \text{ such that } (X_{21})_{jt} > 0 \text{ for some } j \in C_1\},$$

$$T_2^* = \{t \in T_2 : \text{ there exists } x \in S_\mu \text{ such that } (X_{31})_{jt} > 0 \text{ for some } j \in C_2\}.$$

Note that T_1^* and T_2^* are defined independent of any particular element in S_μ. Then it follows from (2. 68) that

$$X_{13}|_{T_1^* \times C_2} = 0; \qquad X_{21}|_{C_1 \times (T_1 \setminus T_1^*)} = 0; \tag{2. 71}$$

also,

$$X_{12}|_{T_2^* \times C_1} = 0; \qquad X_{31}|_{C_2 \times (T_2 \setminus T_2^*)} = 0. \tag{2. 72}$$

It also follows from (2. 69) that

$$X_{11}|_{T_1^* \times (T_1 \cup T_3)} = 0, \tag{2. 73}$$

and from (2. 70) that

$$X_{11}|_{T_2^* \times (T_2 \cup T_3)} = 0. \tag{2. 74}$$

Now we have some information about the elements x in S_μ with respect to the partition

$$\{C_1 \cup T_2^*, C_2 \cup T_1^*\} \tag{2. 75}$$

from equations (2. 59), (2. 71), (2. 72), (2. 73), and (2. 74). If every element of S_μ is not cyclic with respect to the partition (2. 75), then we can form a new partition from (2. 75) by the addition of new subsets of T_1 and T_2 disjoint from T_1^* and T_2^*, respectively. This we do as follows:

We define T_1^{**} and T_2^{**} by

$$T_1^{**} = \{t \in T_1 \setminus T_1^* \ : \ \text{there exists } x \in S_\mu \text{ such that } (X_{11})_{ut} > 0 \text{ for some } u \in T_2^*\},$$
$$T_2^{**} = \{t \in T_2 \setminus T_2^* \ : \ \text{there exists } x \in S_\mu \text{ such that } (X_{11})_{ut} > 0 \text{ for some } u \in T_1^*\}.$$

Then we use (2. 65) again and the equation $eS_\mu^3 e = \{y\}$. Thus, for arbitrary elements x'', x', and x in S_μ, we have

$$X''_{21}X'_{11}X_{12} = 0, \qquad X''_{31}X'_{11}X_{13} = 0, \tag{2. 76}$$

since $y|_{C_1 \times C_1}$ and $y|_{C_2 \times C_2}$ are both zero blocks. It follows from (2. 76) that

$$X_{13}|_{T_1^{**} \times C_2} = 0, \qquad X_{12}|_{T_2^{**} \times C_1} = 0. \tag{2. 77}$$

From the equation $eS_\mu^3 e = \{y\}$, we also have, besides equation (2. 76), the following equations:

$$AX''_{21}X'_{11}X_{11}A' = 0, \tag{2. 78}$$

and

$$BX''_{31}X'_{11}X_{11}B' = 0. \tag{2. 79}$$

Notice that given any $w \in T_2^{**}$ and any $t \in T_2 \cup T_3$, thre exist $v \in T_1^*$, $x' \in S_m u$, $x'' \in S_m u$, and $u \in C_1$, such that

$$(X''_{21})_{uv} > 0, \quad (X'_{11})_{vw} > 0, \quad \text{and} \quad A'_{ts} > 0$$

for any s in C_1. Then (2. 78) implies that

$$X_{11}|_{T_2^{**} \times (T_2 \cup T_3)} = 0. \tag{2. 80}$$

Similarly, (2. 79) implies that

$$X_{11}|_{T_1^{**} \times (T_1 \cup T_3)} = 0. \tag{2. 81}$$

It is now clear that equations (2. 77), (2. 80), and (2. 81) take us closer to the cyclicity partition we have been looking for, and this induction process can be continued, if necessary.

Thus, we now consider the partition

$$\{C_1 \cup T_2^* \cup T_2^{**}, C_2 \cup T_1^* \cup T_1^{**}\} \tag{2. 82}$$

If there still remains an element x in S_μ which is not cyclic with respect to the partition (2. 82), we consider (2. 65) again and the equation $eS_\mu^4 e = \{e\}$ to continue the process. As before, for arbitrary elements x''', x'', x', and x in S_μ, we obtain the following equations:

$$\begin{aligned} X'''_{21} X''_{11} X'_{11} X_{13} &= 0; & (2.\ 83)\\ X'''_{31} X''_{11} X'_{11} X_{12} &= 0; & (2.\ 84)\\ A X'''_{21} X''_{11} X'_{11} X_{11} B' &= 0; & (2.\ 85)\\ B X'''_{31} X''_{11} X'_{11} X_{11} A' &= 0. & (2.\ 86) \end{aligned}$$

From (2. 83), (2. 84), (2. 85), and (2. 86), and from the definitions

$$T_1^{***} = \{t \in T_1 \backslash (T_1^* \cup T_1^{**}) \ : \ \text{there exists } x \in S_\mu \text{ such that } (X_{11})_{ut} > 0 \text{ for some } u \in T_2^{**}\}$$

and

$$T_2^{***} = \{t \in T_2 \backslash (T_2^* \cup T_2^{**}) \ : \ \text{there exists } x \in S_\mu \text{ such that } (X_{11})_{ut} > 0 \text{ for some } u \in T_1^{**}\},$$

it again follows as before that

$$X_{13}|_{T_1^{***}\times C_2} = 0, \qquad X_{12}|_{T_2^{***}\times C_1} = 0, \tag{2. 87}$$

$$X_{11}|_{T_1^{***}\times (T_1\cup T_3)} = 0, \qquad X_{11}|_{T_2^{***}\times (T_2\cup T_3)} = 0. \tag{2. 88}$$

We continue this process if still there remains an element in S_μ which is not cyclic with respect to the partition

$$\{C_1 \cup T_2^* \cup T_2^{**} \cup T_2^{***}, C_2 \cup T_1^* \cup T_1^{**} \cup T_1^{***}\}. \tag{2. 89}$$

This process must terminate after some steps, since both T_1 and T_2 are finite, and we will then have a partition with respect to which <u>all</u> the elements in S_μ are cyclic.

Finally, let us comment that if there is an idempotent e in K with no zero columns and of rank 2, then we will carry out the above analysis using this particular e, and in this case, the analysis is extremely simple. If we write e in block form as

$$e = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}, \tag{2. 90}$$

where A is a $c_1 \times c_1$ strictly positive rank one stochastic matrix indexed by C_1, B is a $c_2 \times c_2$ strictly positive rank one stochastic matrix indexed by C_2, $c_1 = |C_1|$, $c_2 = |C_2|$, and $c_1 + c_2 = d$. Then, the element y is of the form

$$y = \begin{pmatrix} 0 & W \\ Z & 0 \end{pmatrix}, \tag{2. 91}$$

where W is $c_1 \times c_2$, Z is $c_2 \times c_1$, each row of W is the same as any row of B, and each row of Z is the same as any row of A.

If we write $x \in S_\mu$ in the same block form as in (2. 90), that is,

$$x = \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix}, \tag{2. 92}$$

then equation (2. 55), namely that $eS_\mu e = \{y\}$, immediately implies that

$$X_{11} = 0, \qquad X_{22} = 0. \tag{2. 93}$$

This means that S_μ is cyclic with respect to the partition $\{C_1, C_2\}$.

Thus, we have now proven

THEOREM 2.6 *Let μ be a probability measure on the Borel subsets of $d \times d$ stochastic matrices such that the minimal rank of the matrices in the closed semigroup generated by the support of μ is 2. Then the sequence of convolution powers μ^n does not converge weakly iff the support of μ is cyclic.*

EXAMPLE 7 Consider a two-point probability measure μ on 7×7 stochastic matrices such that $S_\mu = \{x, y\}$, where

$$x = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \end{pmatrix}, \quad y = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

Note that K consists of rank two matrices and K contains an idempotent e given by

$$e = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

It can be verified easily that $eS_\mu e = \{y\}$, where $eKe = \{e, y\}$. Obviously the sequence μ^n does not converge weakly, and S_μ here is cyclic with respect to the partition given by:

$$\{C_1 \cup T_2^* \cup T_2^{**}, C_2 \cup T_1^*\},$$

where

$$C_1 = \{6\},\ C_2 = \{7\},\ T_1 = \{1,2,3\},\ T_2 = \{4,5\},\ T_1^* = \{3\},\ T_2^* = \{4\},\ T_2^{**} = \{5\}.$$

■

We now continue with the general $d \times d$ case, still under the assumption of non-weak convergence of the sequence μ^n, but this time we consider the case when the common rank of the matrices in the kernel K is greater than 2.

Consider Example 6 in Section 2.3. We will use this example to produce a probability measure μ on $d \times d$ stochastic matrices (d is any given integer > 3), where the minimal rank of the matrices in S is 3, such that μ^n does not converge weakly, and yet, S_μ is not cyclic.

First, let A be a 4×4 stochastic matrix. Let A_4 be the last row of A. Construct the $(d-4) \times 4$ stochastic matrix A' with identical rows, each row being $A_4 = (a_1, a_2, a_3, a_4)$ Then the $d \times d$ matrix A'' defined by

$$A'' = \left(\begin{array}{c|c} A & 0 \\ \hline A' & 0 \end{array} \right) \qquad (2.\ 94)$$

whose last $d-4$ columns are all zero columns, is a stochastic matrix with the same rank as A. If $B = (b_{ij})$ is another 4×4 stochastic matrix, similarly construct B' and

$$B'' = \left(\begin{array}{c|c} B & 0 \\ \hline B' & 0 \end{array} \right).$$

Then

$$A''B'' = \left(\begin{array}{c|c} AB & 0 \\ \hline A'B & 0 \end{array} \right).$$

Since A' has identical rows, so does $A'B$, and each row of $A'B$ has the entry

$$\sum_{s=1}^{4} a_s b_{sj}$$

in its jth column, which is the jth entry in the last row of AB. Hence $A'B = (AB)'$, and therefore

$$A''B'' = (AB)'',$$

which implies that $A \mapsto A''$ is an isomorphism (into). Thus, Example 6 can be recast, through the correspondence $A \to A''$, to a similar example in the context of $d \times d$ stochastic matrices, $d \geq 4$.

Expanding on this notion, let us once again start from a 4×4 stochastic matrix A. Append $d-4$ zeros to each row of A and then add $d-4$ identical rows of length d, each with zero entries except in the 5th position, which is 1, to create a $d \times d$ stochastic matrix A^* of rank 1 more than that of A. That is,

$$A^* = \left(\begin{array}{c|c} A & 0 \\ \hline 0 & \begin{array}{ccccc} 1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ & & \vdots & & \\ 1 & 0 & 0 & \cdots & 0 \end{array} \end{array}\right).$$

If B is another 4×4 stochastic matrix, similarly construct B^* and calculate

$$\begin{aligned} A^*B^* &= \left(\begin{array}{c|c} A & 0 \\ \hline 0 & \begin{array}{ccccc} 1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ & & \vdots & & \\ 1 & 0 & 0 & \cdots & 0 \end{array} \end{array}\right) \left(\begin{array}{c|c} B & 0 \\ \hline 0 & \begin{array}{ccccc} 1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ & & \vdots & & \\ 1 & 0 & 0 & \cdots & 0 \end{array} \end{array}\right) \\ &= \left(\begin{array}{c|c} AB & 0 \\ \hline 0 & \begin{array}{ccccc} 1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ & & \vdots & & \\ 1 & 0 & 0 & \cdots & 0 \end{array} \end{array}\right) \end{aligned}$$

to deduce that $A \mapsto A^*$ is an isomorphism (into). Thus, we can again recast Example 6 to show that when the rank of the matrices in K is 4, non-weak convergence of μ^n may not imply that S_μ is cyclic.

More generally, we may consider any rank p, $4 \leq p < d$, by constructing the $d \times d$ stochastic matrix $A^*(p)$ from any 4×4 stochastic matrix of rank 3 as follows:

$$A^*(p) = \left(\begin{array}{c|ccccc|c} A & & & 0 & & & 0 \\ \hline 0 & & & I_{p-3} & & & 0 \\ \hline & 0 & 0 & 0 & \cdots & 1 & \\ & 0 & 0 & 0 & \cdots & 1 & \\ 0 & & & \vdots & & & 0 \\ & 0 & 0 & 0 & \cdots & 1 & \end{array}\right),$$

where I_{p-3} is the $(p-3) \times (p-3)$ identity matrix, and where the bottom blocks are of size $(d-p-1) \times 4$, $(d-p-1) \times (p-3)$, and $(d-p-1) \times (d-p-1)$, respectively. It is not difficult to see that, once again, $A \mapsto A^*(p)$ is an isomorphism. Consequently, we conclude that S_μ need not be cyclic in the case of $d \times d$ stochastic matrices when the rank in case is p, $4 \leq p < d$.

Chapter 3

Weak Convergence in Circulant Matrices

3.1 $d \times d$ Circulant Matrices

By a $d \times d$ circulant matrix we mean a matrix of the form

$$\begin{pmatrix} x_0 & x_1 & \cdots & x_{d-1} \\ x_{d-1} & x_0 & \cdots & x_{d-2} \\ \vdots & \vdots & & \vdots \\ x_1 & x_2 & \cdots & x_0 \end{pmatrix}$$

where each row after the first is just the previous row cyclically shifted to the right by one position. Thus, a $d \times d$ matrix $x = (x_{jk}),\ j, k = 0, 1, \ldots, d-1$, is a circulant matrix if and only if

$$x_{jk} = x_{0,k-j}, \tag{3. 1}$$

where the subscripts are taken modulo d, and the entries x_{jk} are all reals. Clearly, a circulant matrix is determined completely by its first row (or column), and we shall, for brevity, denote by $circ(x_0, x_1, \ldots, x_{d-1})$ the $d \times d$ circulant matrix with first row elements $x_0, x_1, \ldots, x_{d-1}$. The permutation matrix $circ(0, 1, 0, \ldots, 0)$ is denoted by P.

Circulant matrices are well-studied in the literature (see, for example, [D]). The following results are well-known:

1. Let x^* denote the conjugate transpose of the $d \times d$ matrix x. Then the following are equivalent:

   i. x is circulant;

ii. x^* is circulant;

iii. $Px = xP$, where the permutation matrix $circ(0, 1, 0, \ldots, 0)$;

iv. $x = f(P)$ for some polynomial f of degree less than d.

2. Let $x = circ(x_0, x_1, \ldots, x_{d-1})$. Consider the $d \times d$ matrix F whose jth column, $1 \leq j \leq d$, is the column with entries $\frac{1}{\sqrt{d}}(1, \omega^{j-1}, \omega^{2(j-1)}, \ldots, \omega^{(d-1)(j-1)})$, where $\omega = \exp(2\pi i/d)$. Define $\lambda_j = x_0 + x_1\omega^{j-1} + x_2\omega^{2(j-1)} + \cdots + x_{d-1}\omega^{(d-1)(j-1)}$. Then the circulant matrix x has the following spectral representation:

$$x = FD_xF^*, \tag{3. 2}$$

where F^* is the conjugate transpose of the unitary matrix F, and D_x is the $d \times d$ diagonal matrix diag$(\lambda_1, \lambda_2, \ldots, \lambda_d)$. Note that the unitary matrix F is independent of x. If the rank of x is r, then exactly $d - r$ diagonal entries of D_x are zeros.

Let μ be a probability measure on the Borel sets of $d \times d$ *real* circulant matrices and let S be the closed (multiplicative) semigroup generated by the support S_μ of μ, so that

$$S = \overline{\bigcup_{n \geq 1} S_\mu^n}. \tag{3. 3}$$

We are interested in studying the problem of weak convergence of the convolution sequence $(\mu^n)_{n \geq 1}$, where, as usual,

$$\mu^{n+1}(B) = \int \mu^n\{y \,:\, yx \in B\}\, \mu(dx),$$

for any Borel set $B \subset S$. Or, equivalently, if $X_1, X_2, \ldots$ are i.i.d. random matrices in S such that $P(X_1 \in B) = \mu(B)$, then $P(X_1X_2 \cdots X_n \in B) = \mu^n(B)$. The following result is well-known (see [HMu]).

LEMMA 3.1 *Assume that the convolution sequence $(\mu^n)_{n \geq 1}$ is tight. That is, given $\epsilon > 0$, we can find a compact subset K_ϵ of S such that $\mu^n(K_\epsilon) > 1 - \epsilon$ for all $n \geq 1$. Then, the sequence μ_n, defined by $\mu_n = (1/n)\sum_{k=1}^{n} \mu^k$, converges weakly to some idempotent probability measure λ such that*

$$\lambda = \mu * \lambda = \lambda * \mu.$$

The support S_λ is the minimal ideal of S, and as such, consists of all those matrices in S which have the minimal rank. In other words,

$$S_\lambda = \{y \in S \,:\, \text{rank } y \leq \text{rank } x \text{ for any } x \in S\}.$$

■

In what follows, throughout this paper, we assume that the sequence (μ^n) is tight so that the assertions in Lemma 3.1 above hold. Since S in (3. 3) is a commutative semigroup, the minimal ideal S_λ in Lemma 3.1 is a compact abelian group and λ is the Haar measure of this group (see [HMu]).

LEMMA 3.2 *Let K be a compact abelian group of $d \times d$ matrices of rank r, where $0 < r \leq d$. Then K is topologically isomorphic to a compact abelian group of $r \times r$ invertible matrices with determinant ± 1.*

Proof. Let e be the identity of K. Then there exists an invertible $d \times d$ matrix y such that

$$y^{-1}ey = \begin{pmatrix} 0 & 0 \\ 0 & I_r \end{pmatrix}, \tag{3. 4}$$

where I_r is the $r \times r$ identity matrix.

For $x \in K$, write

$$y^{-1}xy = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \tag{3. 5}$$

where D is $r \times r$, A is $(d-r) \times (d-r)$, B is $(d-r) \times r$, and C is $r \times (d-r)$.

Since $y^{-1}xy = (y^{-1}ey)(y^{-1}xy) = (y^{-1}xy)(y^{-1}ey)$, we have

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ C & D \end{pmatrix} = \begin{pmatrix} 0 & B \\ 0 & D \end{pmatrix}$$

and consequently $A = 0, B = 0$, and $C = 0$. It is clear that K is topologically isomorphic (under the map $x \mapsto y^{-1}xy$) to the group G given by

$$G = \left\{ D \,:\, y^{-1}xy = \begin{pmatrix} 0 & 0 \\ 0 & D \end{pmatrix} \text{ for some } x \in K \right\}. \tag{3. 6}$$

Since G is a compact abelian group and the determinant map $\det : G \to \mathbb{R}$ is a continuous homomorphism, the range of this map, being a compact group must be either $\{1\}$ or $\{1, -1\}$.

∎

We will need our next result, the structure of an idempotent $d \times d$ real circulant matrix $e = circ(e_0, e_1, \ldots, e_{d-1})$ of rank r, where $0 < r < d$, when we discuss the problem of weak convergence in circulant matrices especially in the next two sections.

Recall that e is idempotent if $e^2 = e$. If $\lambda_0, \lambda_1, \ldots, \lambda_{d-1}$ are the eigenvalues of e, then from (3. 2), $e = F\Lambda F^*$, where $\Lambda = \text{diag}(\lambda_0, \lambda_1, \ldots, \lambda_{d-1})$. Thus, the equation $e^2 = e$ translates to

$$F\Lambda^2 F^* = F\Lambda F^*,$$

or

$$\text{diag}(\lambda_0^2, \lambda_1^2, \ldots, \lambda_{d-1}^2) = \Lambda^2 = \Lambda = \text{diag}(\lambda_0, \lambda_1, \ldots, \lambda_{d-1}).$$

In other words, $\lambda_j = 0$ or 1, $j = 0, 1, \ldots, d-1$.

Moreover, the 0-1 eigenvalues $\lambda_0, \lambda_1, \ldots, \lambda_{d-1}$ satisfy

$$\lambda_j = \lambda_{d-j}, \quad j = 1, 2, \ldots, \lfloor (d-1)/2 \rfloor.$$

Call a sequence $a_1, a_2, \ldots, a_{d-1}$ of real numbers *palindromic* if $a_j = a_{d-j}$, $j = 1, 2, \ldots, \lfloor (d-1)/2 \rfloor$.

LEMMA 3.3 *Let* $e = circ(e_0, e_1, \ldots, e_{d-1})$ *be a* $d \times d$ *idempotent circulant matrix with eigenvalues* $\lambda_0, \lambda_1, \ldots, \lambda_{d-1}$. *Then*

i. each λ_j *is either* 0 *or* 1;

ii. If e *has rank* r, *where* $0 < r < d$, *then* $e_0 = r/d$, *and exactly* r *of the* λ_j*s are equal to* 1. *In particular, if* $r = d-1$ *and* d *is even, there are two possible idempotents:*

$$e = circ((d-1)/d, -1/d, -1/d, \ldots, -1/d). \tag{3. 7}$$

or

$$e = circ((d-1)/d, 1/d, -1/d, 1/d, \ldots, -1/d, 1/d). \tag{3. 8}$$

If $r = d - 1$ *and* d *is odd, then there is only one idempotent and it is given by (3. 7).*

If $r = 1$ *and* d *is even, there are two possible idempotents:*

$$e = circ(1/d, 1/d, \ldots, 1/d, 1/d) \tag{3. 9}$$

or

$$e = circ(1/d, -1/d, \ldots, 1/d, -1/d). \tag{3. 10}$$

If $r = 1$ *and* d *is odd, then there is only one idempotent and it is given by (3. 9).*

Proof.

i. Clear from the discussion preceeding the lemma.

ii. Recall that the rank of a diagonalizable matrix is precisely the number of its nonzero eigenvalues, and that the trace of a matrix is the sum of its eigenvalues. Thus, for the idempotent circulant matrix e, the trace and the rank coincide, leading to the equation $de_0 = r$.

If $r = d - 1$ and d is even, then either $\lambda_0 = 0$ and $\lambda_1 = \lambda_2 = \cdots = \lambda_{d-1} = 1$, or $\lambda_{d/2} = 0$ and the rest of the λ_j's are equal to 1. (3. 7) and (3. 8) now follow from (3. 2).

If $r = d - 1$ and d is odd, then $\lambda_0 = 0$ and $\lambda_1 = \lambda_2 = \cdots = \lambda_{d-1} = 1$, and the idempotent is uniquely given by (3. 7).

If $r = 1$ and d is even, then either $\lambda_0 = 1$ and $\lambda_j = 0$ for $j = 1, 2, \ldots, d-1$, or $\lambda_{d/2} = 1$ and the rest of the λ_j's are zero. Thus, by (3. 2), these two possibilities correspond to (3. 9) and (3. 10), respectively.

If $r = 1$ and d is odd, then $\lambda_0 = 1$ and $\lambda_j = 0$ for $j = 1, 2, \ldots, d - 1$. So there is a unique idempotent, and it is given by (3. 9).

■

Next, we will characterize the structure of a compact abelian group of $d \times d$ circulant matrices.

LEMMA 3.4 *Let K be a compact abelian group of $d \times d$ circulant matrices. Suppose that the rank of the matrices in K is r, $0 < r \leq d$. Then K is isomorphic to a compact group H of invertible r by r matrices, where each matrix in H has the same diagonal block form as follows: for any $x \in H$, the first k_1 elements along the diagonal are 1 or -1, and there are k_2 diagonal blocks, where each block is a 2×2 rotation matrix of the form*

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix},$$

where $k_1 + 2k_2 = r$, and the numbers k_1 and k_2 are independent of $x \in H$.

Proof. We use the spectral representation (3. 2) of elements in K. For $x \in K$, write

$$x = FD_xF^*,$$

where F and D_x are as in (3. 2). Since $x \in K$ has rank r, exactly $d - r$ diagonal entries of D_x are zeros.

First, we claim that for $x \in K$, if $(D_x)_{ii} = 0$, then $(D_y)_{ii} = 0$ for all $y \in K$. In other words, the zero entries occur in the same diagonal positions for any D_x, $x \in K$. To see this, note that if this is not true for some x and y in K, then since

$$xy = (FD_xF^*)(FD_yF^*) = FD_xD_yF^* = FD_{xy}F^*,$$

so that $D_xD_y = D_{xy}$, it will follow that the number of zeros on the diagonal of D_{xy} is strictly greater than that for either D_x or D_y, which contradicts that x, y, and xy have the same rank.

Second, if for any $x \in K$, $\alpha + i\beta$ is a nonzero entry on the diagonal of D_x, then since $\{x^k : k \geq 1\}$ is bounded, $|\alpha + i\beta|^k$ cannot go to infinity or zero, and this means that

$|\alpha + i\beta| = 1$, or $\alpha^2 + \beta^2 = 1$. Thus, each diagonal entry of D_x is zero or has absolute value 1. [Notice that if $\lim_{k\to\infty} x^{n_k} = y$, $x, y \in K$, then if $0 < |\alpha + i\beta| < 1$, where $\alpha + i\beta$ is a diagonal entry of x, then $D_y = \lim_{k\to\infty} D_x^{n_k} = \lim_{k\to\infty} F^* x^{n_k} F = F^* y F$, so that the number of zeros on the diagonal of D_y will exceed that for D_x. This contradicts the fact that x and y have the same rank.]

Third, let the jth diagonal entry of both D_x and D_y, $x, y \in K$, be nonzero, and be respectively $\alpha + i\beta$ and $\gamma + i\delta$. Thus,

$$\begin{aligned} \lambda_j &= x_0 + x_1\omega^{j-1} + x_2\omega^{2(j-1)} + \cdots + x_{d-1}\omega^{(d-1)(j-1)} \\ &= \alpha + i\beta, \; j \geq 2, \end{aligned}$$

so that the $(d-j+2)$th diagonal entry of D_x is

$$\begin{aligned} \lambda_{d-j+2} &= x_0 + x_1\omega^{d-j+1} + x_2\omega^{2(d-j+1)} + \cdots + x_{d-1}\omega^{(d-1)(d-j+1)} \\ &= \alpha - i\beta. \end{aligned}$$

It is then clear that the $(d-j+2)$th diagonal entry of D_y will also be $\gamma - i\delta$, since $\omega^{k(d-j+1)}$ is the conjugate of $\omega^{k(j-1)}$, for $1 \leq k \leq d-1$. Since $\lambda_j = \overline{\lambda_{d-j+2}}$, $1 \leq j \leq d$, it is clear that λ_1 is real, λ_2 and λ_d are conjugate, λ_3 and λ_{d-1} are conjugate, and so on.

It follows from the above analysis that we can find a finite number of permutation matrices $P_1, P_2, \ldots, P_m$ such that for each $x \in K$, the matrix

$$P_m^{-1} P_{m-1}^{-1} \cdots P_1^{-1} D_x P_1 P_2 \cdots P_m$$

is of the form

$$\begin{pmatrix} 0 &&&&&&&&&&&& \\ & \ddots &&&&&&&&&&& \\ && 0 &&&&&&&&&& \\ &&& 1 &&&&&&&&& \\ &&&& \ddots &&&&&&&& \\ &&&&& 1 &&&&&&& \\ &&&&&& -1 &&&&&& \\ &&&&&&& \ddots &&&&& \\ &&&&&&&& -1 &&&& \\ &&&&&&&&& \alpha_1+i\beta_1 & 0 && \\ &&&&&&&&& 0 & \alpha_1-i\beta_1 && \\ &&&&&&&&&&& \alpha_2+i\beta_2 & 0 \\ &&&&&&&&&&& 0 & \alpha_2-i\beta_2 \\ &&&&&&&&&&&&& \ddots \end{pmatrix}$$

where $\alpha_j^2 + \beta_j^2 = 1$ for each j, and the conjugate eigenvalues are adjacent entries on the diagonal. Notice also that the map

$$\begin{pmatrix} \alpha + i\beta & 0 \\ 0 & \alpha - i\beta \end{pmatrix} \mapsto \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$$

is an isomorphism. The lemma is now clear.

■

In the next two sections, we consider the special case of 3×3 and 4×4 circulant matrices.

3.2 3×3 Circulant Matrices

Let K be the kernel of the closed semigroup S generated by the support S_μ of a probability measure μ on 3×3 circulant matrices. As before, we assume that (μ^n) is tight, so that K is a compact abelian group. Note that the matrices in K all have the same rank.

We first try to determine how the matrices in K look like. We discuss separately the cases when the matrices in K have rank $1, 2$ or 3.

(i) Let the rank of the matrices in K be 1. Then K is topologically isomorphic to a subgroup of the multiplicative group $\{1, -1\}$. Hence, either $K = \{e\}$ or $K = \{e, -e\}$, where $e = circ(1/3, 1/3, 1/3)$ is the unique idempotent of rank 1.

(ii) Suppose the matrices in K all have rank 2. If $x = circ(x_0, x_1, x_2) \in K$, then the eigenvalues of x are

$$\begin{aligned} \lambda_0 &= x_0 + x_1 + x_2 \\ \lambda_1 &= x_0 + x_1\omega + x_2\omega^2, \\ \lambda_2 &= x_0 + x_1\omega^2 + x_2\omega, \end{aligned}$$

where $\omega = (-1 + \sqrt{3}i)/2$.

For x to have rank 2, exactly one of λ_0, λ_1 and λ_2 must be equal to 0. Note that

$$\begin{aligned} \lambda_1\lambda_2 &= x_0^2 + x_1^2 + x_2^2 - (x_0x_1 + x_0x_2 + x_1x_2) \\ &= ((x_0 - x_1)^2 + (x_0 - x_2)^2 + (x_1 - x_2)^2)/2. \end{aligned}$$

Thus, if either $\lambda_1 = 0$ or $\lambda_2 = 0$, then $x_0 = x_1 = x_2$ and consequently x will have rank 1. Thus, it must be the case that $\lambda_0 = 0$, and therefore every matrix in K must be of the form

$$circ(x_0, x_1, -(x_0 + x_1)), \tag{3. 11}$$

with x_0 and x_1 not both zero.

We show that if K is infinite then K is topologically isomorphic to the circle group T of complex numbers $z = e^{i\theta}$, $\theta \in [0, 2\pi)$.

By Lemma 3.3, $e = circ(2/3, -1/3, -1/3)$ is the identity of K. From the proof of Lemma 3.2, we know that there exists an invertible 3×3 matrix y satisfying (3. 6). In the present case, we may take the matrix y to be the orthogonal matrix

$$y = \begin{pmatrix} 1/\sqrt{3} & -1/\sqrt{6} & -1/\sqrt{2} \\ 1/\sqrt{3} & -1/\sqrt{6} & 1/\sqrt{2} \\ 1/\sqrt{3} & 2/\sqrt{6} & 0 \end{pmatrix}. \tag{3. 12}$$

With y given by (3. 12) and $x \in K$ as in (3. 11), we find that

$$y^{-1}xy = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{3}{2}x_0 & \frac{\sqrt{3}}{2}(x_0 + 2x_1) \\ 0 & -\frac{\sqrt{3}}{2}(x_0 + 2x_1) & \frac{3}{2}x_0 \end{pmatrix}.$$

Thus, we conclude that the compact abelian group G of Lemma 3.2, to which K is topologically isomorphic, satisfies

$$G \subseteq \left\{ \begin{pmatrix} \frac{3}{2}x_0 & \frac{\sqrt{3}}{2}(x_0 + 2x_1) \\ -\frac{\sqrt{3}}{2}(x_0 + 2x_1) & \frac{3}{2}x_0 \end{pmatrix} \; : \; x_0, x_1 \text{ not both } 0 \right\}. \tag{3. 13}$$

Note that the determinant of the matrices in G must be ± 1, since G is compact. The determinant of an element of G is

$$\frac{9}{4}x_0^2 + \frac{3}{4}(x_0 + 2x_1)^2 = 3(x_0^2 + x_0x_1 + x_1^2).$$

Since this quantity is always positive, we conclude that all matrices in G have determinant 1. This leads to the equation

$$x_0^2 + x_0x_1 + x_1^2 = 1/3,$$

so

$$x_1 = \frac{-3x_0 \pm \sqrt{12 - 27x_0^2}}{6} \tag{3. 14}$$

with $x_0 \in [-2/3, 2/3]$. In view of (3. 11), we deduce that K is the union of the sets

$$\{circ(x_0, f(x_0), g(x_0)) \; : x_0 \in [-2/3, 2/3]\} \tag{3. 15}$$

and

$$\{circ(x_0, g(x_0), f(x_0)) \; : x_0 \in [-2/3, 2/3]\} \tag{3. 16}$$

where

$$f(x) = \frac{-3x - \sqrt{12 - 27x^2}}{6} \tag{3. 17}$$

and

$$g(x) = \frac{-3x + \sqrt{12 - 27x^2}}{6}. \tag{3. 18}$$

Also, (3. 14) implies

$$\frac{\sqrt{3}}{2}(x_0 + 2x_1) = \pm\frac{1}{2}\sqrt{4 - 9x_0^2}, \tag{3. 19}$$

and substituting (3. 19) in (3. 13) gives a clearer description of the compact abelian group G :

$$\begin{aligned} G \;\subseteq\; & \left\{ \begin{pmatrix} \frac{3}{2}x_0 & \frac{1}{2}\sqrt{4-9x_0^2} \\ -\frac{1}{2}\sqrt{4-9x_0^2} & \frac{3}{2}x_0 \end{pmatrix} \;:\; x_0 \in [-2/3, 2/3] \right\} \\ & \bigcup \left\{ \begin{pmatrix} \frac{3}{2}x_0 & -\frac{1}{2}\sqrt{4-9x_0^2} \\ \frac{1}{2}\sqrt{4-9x_0^2} & \frac{3}{2}x_0 \end{pmatrix} \;:\; x_0 \in [-2/3, 2/3] \right\}. \end{aligned}$$

Now write $x_0 = \frac{2}{3}\cos\theta$. Then by direct computation, one sees that the map

$$e^{i\theta} \leftrightarrow \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \leftrightarrow \begin{cases} circ(x_0, f(x_0), g(x_0)) & \text{if } \sin\theta \geq 0 \\ circ(x_0, g(x_0), f(x_0)) & \text{if } \sin\theta < 0 \end{cases} \tag{3. 20}$$

is a topological isomorphism between the circle group T and K. Here, f and g are given by (3. 17) and (3. 18), respectively.

(iii) Suppose the matrices in K all have rank 3. Then K must coincide with the whole of S, since the identity matrix $e = circ(1, 0, 0)$ is in K, and therefore, $S = eS \subseteq K$. This means S itself must be a compact abelian group of 3×3 circulant matrices which have determinant ± 1.

Let $H = \{x = circ(x_0, x_1, x_2) \in S \;:\; \det x = 1\}$. If $y \in S$ is such that $\det y = -1$, then $S = H \cup yH$.

Now for all $x \in H$,

$$\begin{aligned} \det x &= x_0^3 + x_1^3 + x_2^3 - 3x_0x_1x_2 \\ &= (x_0 + x_1 + x_2)((x_0 - x_1)^2 + (x_0 - x_2)^2 + (x_1 - x_2)^2)/2 \\ &= 1, \end{aligned} \tag{3. 21}$$

so $x_0 + x_1 + x_2 > 0$. Moreover, it is easy to establish (by induction, say) that the entries in each row of x^n sum to $(x_0 + x_1 + x_2)^n$, for any $n \geq 1$. Since H is a closed subgroup of the compact group S, H, too, is compact. This forces

$$x_0 + x_1 + x_2 = 1, \tag{3. 22}$$

so every element of H is of the form

$$circ(x_0, x_1, 1 - (x_0 + x_1)). \tag{3. 23}$$

(3. 21) and (3. 22) then imply

$$(x_0 - x_1)^2 + (2x_0 + x_1 - 1)^2 + (2x_1 + x_0 - 1)^2 = 2,$$

or

$$x_0^2 + x_1^2 + x_0x_1 = x_0 + x_1. \tag{3. 24}$$

Thus,

$$x_1 = \frac{1 - x \pm \sqrt{(1 + 3x_0)(1 - x_0)}}{2} \tag{3. 25}$$

with $x_0 \in [-1/3, 1]$. In view of (3. 23), we conclude that H is the union of the sets

$$\{circ(x_0, f(x_0), g(x_0)) \ : x_0 \in [-1/3, 1]\}$$

and

$$\{circ(x_0, g(x_0), f(x_0)) \ : x_0 \in [-1/3, 1]\}$$

where

$$f(x) = \frac{1 - x - \sqrt{(1 + 3x)(1 - x)}}{2} \tag{3. 26}$$

and

$$g(x) = \frac{1 - x + \sqrt{(1+3x)(1-x)}}{2}. \tag{3. 27}$$

Consider now the orthogonal matrix s given by

$$s = \begin{pmatrix} 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3} \\ -2/\sqrt{6} & 0 & 1/\sqrt{3} \end{pmatrix}.$$

For $x = circ(x_0, x_1, 1 - (x_0 + x_1)) \in H$, we have, using (3. 25),

$$s^{-1}xs = \begin{pmatrix} \frac{1}{2}(3x_0 - 1) & \pm\frac{\sqrt{3}}{2}\sqrt{(1+3x_0)(1-x_0)} & 0 \\ \mp\frac{\sqrt{3}}{2}\sqrt{(1+3x_0)(1-x_0)} & \frac{1}{2}(3x_0 - 1) & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Thus, writing $x_0 = (2\cos\theta + 1)/3$, direct computation yields the topological isomorphism

$$e^{i\theta} \leftrightarrow \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \leftrightarrow \begin{cases} circ(x_0, f(x_0), g(x_0)) & \text{if } \sin\theta \geq 0 \\ circ(x_0, g(x_0), f(x_0)) & \text{if } \sin\theta < 0 \end{cases} \tag{3. 28}$$

between the circle group and H. Here, f and g are given by (3. 26) and (3. 27), respectively.

It follows that S is topologically isomorphic either to the circle group T or to the direct product of T and $\{1, -1\}$.

Next, we give a well-known characterization of the proper closed (compact) subgroups of the circle group T.

LEMMA 3.5 *Every proper closed subgroup of the circle group T is finite cyclic.*

Proof. Observe that T may be identified with $U = [0, 1)$ under addition mod 1 with its usual topology at every point, except that an open neighborhood of 0 is $[0, a) \cup (b, 1)$, where $a, b \in (0, 1)$.

Let H be a proper compact subgroup of U. Then the identity $0 \in H$, and we claim that 0 and every point in H must be isolated. If 0 is not isolated, then given $\varepsilon > 0$, there must be $x \in H$ such that $x \in (0, \varepsilon)$. Now given any $y \in U$, there exists an integer n such that $y \leq nx \leq y + 2\varepsilon$. Since $nx \in H$, this means that H is dense in U. It follows by compactness of H that $H = U$. Therefore, H cannot be proper.

We conclude that the points of a proper compact subgroup H are all isolated, so that such a subgroup must be finite. Let $H = \{h_0, h_1, \ldots, h_{n-1}\}$, where $0 = h_0 < h_1 < \cdots < h_{n-1}$ and n is the cardinality of H. Thus h_1 is the smallest positive element in H.

We claim that $h_k = kh_1$ for all $k = 1, 2, \ldots, n-1$. If this is not the case, let $s > 1$ be the first integer such that $h_s \neq sh_1$. Since $(s-1)h_1 = h_{s-1} < h_s$ and h_s is the smallest element in H greater than h_{s-1}, we have

$$(s-1)h_1 < h_s < sh_1.$$

But this implies

$$0 < h_s - (s-1)h_1 < sh_1 - (s-1)h_1 = h_1. \tag{3. 29}$$

Since $h_s - (s-1)h_1 \in H$, (3. 29) is a contradiction to the minimality of h_1. We conclude that H is finite cyclic, as desired.

■

Before we state the main result of this section, we state without proof a well-known necessary and sufficient condition for the convergence of (μ^n), under the condition of tightness. The version we present here is taken from part (iii) of Theorem 2.13 in [HMu], and is adapted to our present situation.

LEMMA 3.6 *Let μ be a probability measure on a commutative semigroup S of $d \times d$ real matrices such that the support S_μ of μ generates S. Suppose that the sequence $(\mu^n)_{n\geq 1}$ is tight. Then K, the kernel (that is, the smallest ideal) of S, is a compact abelian group. Furthermore, (μ^n) converges weakly if and only if there does not exist a closed subgroup H of the compact abelian group K such that*

$$eS_\mu \subset gH \tag{3. 30}$$

for some $g \in K \setminus H$. Here, e is the identity of K.

∎

In view of the above discussion and Lemma 3.6, we have the following

THEOREM 3.1 *Let μ be a probability measure on 3×3 real circulant matrices. Let S be the closed commutative semigroup generated by the support S_μ of μ, and let K be the kernel of S. Suppose $(\mu^n)_{n\geq 1}$ is tight. Then*

(i) If K consists of rank 1 matrices, then either

(a) $K = \{e\}$ and $\mu^n \xrightarrow{w} \delta_e$, or

(b) $K = \{e, -e\}$ and (μ^n) converges weakly if and only if

$$S_\mu \not\subseteq \{circ(x_0, x_1, -1 - x_0 - x_1) \,:\, x_0, x_1 \in \mathbb{R}\}.$$

Here, $e = circ(1/3, 1/3, 1/3)$ and δ_e is the point mass at e.

(ii) If K consists of rank 2 matrices, then either

(a) $K = \{e\}$ and $\mu^n \xrightarrow{w} \delta_e$,

(b) $K = \{e, -e\}$ and (μ^n) converges weakly if and only if

$$S_\mu \not\subseteq \{circ(x_0, x_0 + 1, x_0 + 1) \,:\, x_0 \in \mathbb{R}\}, \qquad \textit{or}$$

(c) K is topologically isomorphic to the circle group T (or a finite subgroup T_0 of T) and (μ^n) converges weakly if and only if the image of S_μ under the topological isomorphism given by (3. 20) is not contained in a coset of a finite subgroup of T (respectively, a proper subgroup of T_0).

Here, $e = circ(2/3, -1/3, -1/3)$.

(iii) If K consists of rank 3 matrices, then either

(a) $K = \{I_3\}$ *and* $\mu^n \xrightarrow{w} \delta_{I_3}$,

(b) $K = \{I_3, -I_3\}$ *and* (μ^n) *converges weakly if and only if*

$$S_\mu \neq \{-I_3\}, \qquad \textit{or}$$

(c) $K = S$ *is topologically isomorphic to the circle group* T *or to* $T \times \{1, -1\}$ *(or a finite subgroup* T_0 *of* T *or* $T \times \{1, -1\}$*) and* (μ^n) *converges weakly if and only if the image of* S_μ *under the topological isomorphism given by (3. 28) is not contained in a coset of a finite subgroup of* T *or* $T \times \{1, -1\}$ *(respectively, a proper subgroup of* T_0*).*

■

Let us consider two examples.

EXAMPLE 8 Suppose that

$$S_\mu \subset \{circ(x, x-1, x-1) \,:\, x \in \mathbb{R}\}.$$

Suppose also that for some positive integer n,

$$\mu^n(e) > 0,$$

where e is the idempotent circulant matrix $circ(2/3, -1/3, -1/3)$. It is easily verified that for any matrix $y \in S_\mu$, $ye = ey = e$. In other words, the kernel K of $S = \overline{\cup_{k\geq 1} S_\mu^k}$ is $\{e\}$. By Lemma 2.19 (page 106 in [HMu]), it follows that

$$\lim_{n\to\infty} \mu^n(e) = 1$$

so that μ^n converges weakly to δ_e.

EXAMPLE 9 Suppose that

$$S_\mu \subset \{circ(x_0, x_1, x_2) \,:\, x_0^2 + x_1^2 + x_2^2 - x_0x_1 - x_1x_2 - x_0x_2 = 1\}.$$

Let us denote the set on the right by M. Suppose also that for some positive integer n_0,

$$\mu^{n_0}(A) > 0,$$

where the set A is given by

$$A = \left\{ circ\left(x_0, -\frac{1}{2}x_0 + \frac{\sqrt{3}}{2}\sqrt{\frac{4}{9} - x_0^2}, -\frac{1}{2}x_0 - \frac{\sqrt{3}}{2}\sqrt{\frac{4}{9} - x_0^2} \right) \, : \, x_0^2 \leq \frac{4}{9} \right\}.$$

It can also be verified that M is a semigroup. Indeed, if

$$(x_0 - x_1)^2 + (x_1 - x_2)^2 + (x_0 - x_2)^2 = 2$$

and

$$(y_0 - y_1)^2 + (y_1 - y_2)^2 + (y_0 - y_2)^2 = 2,$$

then

$$circ(x_0, x_1, x_2) \cdot circ(y_0, y_1, y_2) = circ(z_0, z_1, z_2),$$

and

$$\begin{aligned}
(z_0 - z_1)^2 + (z_1 - z_2)^2 + (z_0 - z_2)^2 & \\
&= [x_0(y_0 - y_1) + x_1(y_2 - y_0) + x_2(y_1 - y_2)]^2 \\
&\quad + [x_0(y_1 - y_2) + x_1(y_0 - y_1) + x_2(y_2 - y_0)]^2 \\
&\quad + [x_0(y_0 - y_2) + x_1(y_2 - y_1) + x_2(y_1 - y_0)]^2 \\
&= 2(x_0^2 + x_1^2 + x_2^2) \\
&\quad -2(x_0x_1 + x_0x_2 + x_1x_2)(y_0^2 + y_1^2 + y_2^2 - y_0y_1 - y_1y_2 - y_0y_2) \\
&= 2.
\end{aligned}$$

Also, notice that $y \in M$ and $\det y = 0$ imply that $y \in A$, and all matrices in A have rank 2. It follows that $A \cap S$, where $S = \overline{\cup_{n \geq 1} S_\mu^n}$, is a compact abelian group, and the kernel of S. By Proposition 2.19 (page 106 in [HMu]), it also follows

$$\lim_{n \to \infty} \mu^n(A) = 1$$

so that (μ^n) is tight.

It is also easily seen that if y is the orthogonal matrix given by

$$y = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{6} & -1/\sqrt{2} \\ 1/\sqrt{3} & 1/\sqrt{6} & 1/\sqrt{2} \\ 1/\sqrt{3} & -2/\sqrt{6} & 0 \end{pmatrix},$$

then $y^{-1}Ay$ is the circle group.

REMARK 1 (Personal communication from Karl H. Hofmann of Darmstadt, Germany) Any closed infinite abelian subgroup C of $SO(3)$ is a circle group. The reason is, briefly, the following: C is a Lie group, and the identity component C_0 of C must, thus, be open in C. It follows that C_0 is a torus. In $SO(3)$, however, each torus is a circle group, and is maximal abelian, implying that C is the circle group.

In the case of 4×4 circulant matrices (considered next), we will observe the same type of behavior with the matrices in the kernel group K.

3.3 4×4 Circulant Matrices

Let K be the kernel of the closed semigroup S generated by the support S_μ of a probability measure μ on 4×4 circulant matrices. As before, we assume that (μ^n) is tight, so that K is a compact abelian group.

We first try to determine how the matrices in K look like. Note that the eigenvalues of $x = circ(x_0, x_1, x_2, x_3) \in K$ are given by

$$\lambda_0 = x_0 + x_1 + x_2 + x_3 \qquad (3.\ 31)$$

$$\lambda_1 = (x_0 - x_2) + (x_1 - x_3)i, \qquad (3.\ 32)$$

$$\lambda_2 = x_0 - x_1 + x_2 - x_3, \qquad (3.\ 33)$$

$$\lambda_3 = \overline{\lambda_1} = (x_0 - x_2) - (x_1 - x_3)i. \qquad (3.\ 34)$$

Also, if F is the 4×4 Fourier matrix of section 3.1, then we have the spectral representation (see (3. 2))

$$x = F\,\mathrm{diag}(\lambda_0, \lambda_1, \lambda_2, \lambda_3)\,F^*. \tag{3. 35}$$

For notational convenience, we will denote by D_x the diagonal matrix on the right side of (3. 35).

Note from (3. 35) that the rank of x coincides with that of D_x, so it is exactly the number of nonzero λ_j.

It is also clear from (3. 35) that for every $k \geq 1$,

$$x^k = F\,D_x^k\,F^*.$$

Now let $\lambda_j \neq 0$. Then since $(x^k)_{k\geq 1}$ is bounded, $|\lambda_j|^k$ cannot go to infinity, as $k \to \infty$. Thus, $|\lambda_j| \leq 1$. Suppose $|\lambda_j| < 1$. Take a subsequence (x^{n_k}) of (x^k) converging to $y = FD_yF^* \in K$. Then

$$y = \lim_{k\to\infty} x^{n_k} = \lim_{k\to\infty} F\,D_x^{n_k}\,F^* = F\left(\lim_{k\to\infty} D_x^{n_k}\right)F^*,$$

so $D_y = \lim_{k\to\infty} D_x^{n_k}$. Since $|\lambda_j| < 1$, the (j, j) entry of D_y must be zero. This means D_y has rank strictly less than that of x, contradicting $y \in K$. We conclude that if $\lambda_j \neq 0$ then $|\lambda_j| = 1$.

Now we discuss separately the cases when the matrices in K have rank $1, 2, 3$ or 4. To facilitate the discussion, we will make use of the vector

$$z = (\lambda_0, \lambda_2, |\lambda_1|^2) = (x_0 + x_1 + x_2 + x_3, x_0 - x_1 + x_2 - x_3, (x_0 - x_2)^2 + (x_1 - x_3)^2).$$

(i) Let the rank of the matrices in K be 1. Then either $z = (\pm 1, 0, 0)$ or $z = (0, \pm 1, 0)$. In either case we have

$$(x_0 - x_2)^2 + (x_1 - x_3)^2 = 0,$$

so

$$x_0 = x_2,\ x_1 = x_3.$$

If $z = (\pm 1, 0, 0)$, then we have

$$\begin{aligned} x_0 + x_1 + x_2 + x_3 &= 2(x_0 + x_1) = \pm 1, \\ x_0 - x_1 + x_2 - x_3 &= 2(x_0 - x_1) = 0. \end{aligned}$$

Hence

$$\pm 1/2 = x_0 + x_1,\ x_0 = x_1,$$

so $x_0 = \pm 1/4$. This gives the circulants

$$e_{11} = circ(1/4, 1/4, 1/4, 1/4),\ -e_{11} = circ(-1/4, -1/4, -1/4, -1/4).$$

On the other hand, if $z = (0, \pm 1, 0)$, then we get

$$\pm 1/2 = x_0 - x_1,\ x_0 = -x_1,$$

which give the circulants

$$e_{12} = circ(1/4, -1/4, 1/4, -1/4),\ -e_{12} = circ(-1/4, 1/4, -1/4, 1/4).$$

Observe that e_{11} and e_{12} are idempotent rank 1 circulants.

We conclude that either $K = \{e_{11}\}$ or $K = \{e_{11}, -e_{11}\}$, or $K = \{e_{12}\}$ or $K = \{e_{12}, -e_{12}\}$.

(ii) Suppose the matrices in K all have rank 2. Then either $z = (\pm 1, \pm 1, 0)$ or $z = (0, 0, 1)$.

The first possibility forces $x_0 = x_2,\ x_1 = x_3$. When $z = (1, 1, 0)$, we have the equations

$$\begin{aligned} 2(x_0 + x_1) &= 1, \\ 2(x_0 - x_1) &= 1, \end{aligned}$$

so $x_0 = 1/2$ and $x_1 = 0$. This gives the idempotent rank 2 circulant

$$e_{21} = circ(1/2, 0, 1/2, 0).$$

Similarly, when $z = (-1, -1, 0)$, we get the circulant

$$-e_{21} = circ(-1/2, 0, -1/2, 0).$$

When $z = (1, -1, 0)$ we have

$$f = circ(0, 1/2, 0, 1/2),$$

and when $z = (-1, 1, 0)$ we have

$$-f = circ(0, -1/2, 0, -1/2).$$

A quick calculation tells us that $\{e_{21}, -e_{21}, f, -f\}$ is isomorphic to the direct product of the multiplicative group $\{1, -1\}$ with itself. Thus, in the first possibility, when $z = (\pm 1, \pm 1, 0)$, we have that K is a subgroup of $\{e_{21}, -e_{21}, f, -f\}$.

The second possibility, when $z = (0, 0, 1)$, leads to

$$x_0 + x_2 = x_1 + x_3 = 0$$

and

$$(x_0 - x_2)^2 + (x_1 - x_3)^2 = 1,$$

so

$$(2x_0)^2 + (2x_1)^2 = 1,$$

so

$$x_1 = \pm\sqrt{1/4 - x_0^2}.$$

This means K consists of circulants of the form

$$\begin{aligned} x &= circ(x_0, x_1, -x_0, -x_1) \\ &= circ\left(x_0, \pm\sqrt{1/4 - x_0^2}, -x_0, \mp\sqrt{1/4 - x_0^2}\right),\ x_0 \in [-1/2, 1/2]. \end{aligned}$$

Note that if

$$D = \frac{1}{2}\begin{pmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{pmatrix}, \tag{3. 36}$$

then D is orthogonal and

$$DxD^{-1} = \begin{pmatrix} 2x_0 & -2x_1 & 0 & 0 \\ 2x_1 & 2x_0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

where $x_0 = \frac{1}{2}\cos\theta$ and $x_0 \in [-1/2, 1/2]$. Thus, in this case

$$K = \left\{ circ\left(x_0, \pm\sqrt{1/4 - x_0^2}, -x_0, \mp\sqrt{1/4 - x_0^2} \right) \;:\; x_0 \in [-1/2, 1/2] \right\}$$

is a compact group topologically isomorphic to the circle group. Observe that the identity of K is

$$e_{22} = circ(1/2, 0, -1/2, 0),$$

the other idempotent rank 2 circulant.

(iii) Suppose the matrices in K all have rank 3. Then either $z = (\pm 1, 0, 1)$ or $z = (0, \pm 1, 1)$.

When $z = (1, 0, 1)$, we get the equations

$$1/2 = x_0 + x_2 = x_1 + x_3,$$

and

$$(2x_0 - 1/2)^2 + (2x_1 - 1/2)^2 = 1.$$

Thus,

$$x_1 = \frac{1}{4} \pm \frac{1}{2}\sqrt{(3 - 4x_0)(1 + 4x_0)},$$

and we deduce that K consists of circulants of the form

$$\begin{aligned} x &= circ(x_0, x_1, 1/2 - x_0, 1/2 - x_1) \\ &= circ\Bigg(x_0, \frac{1}{4} \pm \frac{1}{2}\sqrt{(3 - 4x_0)(1 + 4x_0)}, 1/2 - x_0, \\ &\qquad \frac{1}{4} \mp \frac{1}{2}\sqrt{(3 - 4x_0)(1 + 4x_0)} \Bigg) \end{aligned}$$

with $x_0 \in [-1/4, 3/4]$. Let D_1 be the orthogonal matrix

$$D_1 = \frac{1}{2}\begin{pmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \end{pmatrix}. \tag{3. 37}$$

Then

$$D_1 x D_1^{-1} = \begin{pmatrix} 2x_0 - 1/2 & -(2x_1 - 1/2) & 0 & 0 \\ 2x_1 - 1/2 & 2x_0 - 1/2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

where $x_0 = \frac{1}{4} + \frac{1}{2}\cos\theta$ and $x_0 \in [-1/4, 3/4]$. Hence, in this case,

$$\begin{aligned} K \quad = \quad & \Big\{ circ\Big(x_0, \frac{1}{4} \pm \frac{1}{2}\sqrt{(3-4x_0)(1+4x_0)}, 1/2 - x_0, \\ & \frac{1}{4} \mp \frac{1}{2}\sqrt{(3-4x_0)(1+4x_0)} \Big) \Big\}, \end{aligned}$$

with $x_0 \in [-1/4, 3/4]$, is a compact group topologically isomorphic to the circle group. Note that the identity of K is

$$e_{31} = circ(3/4, 1/4, -1/4, 1/4),$$

one of two idempotent rank 3 circulants.

A similar situation happens when $z = (0, 1, 1)$. We get the equations

$$1/2 = x_0 + x_2 = -(x_1 + x_3),$$

and

$$(2x_0 - 1/2)^2 + (2x_1 + 1/2)^2 = 1,$$

so

$$x_1 = -\frac{1}{4} \pm \frac{1}{2}\sqrt{(3-4x_0)(1+4x_0)}$$

and K therefore consists of circulants of the form

$$\begin{aligned} x &= circ(x_0, x_1, 1/2 - x_0, -1/2 - x_1) \\ &= circ\Bigg(x_0, -\frac{1}{4} \pm \frac{1}{2}\sqrt{(3-4x_0)(1+4x_0)}, 1/2 - x_0, \\ &\qquad -\frac{1}{4} \mp \frac{1}{2}\sqrt{(3-4x_0)(1+4x_0)}\Bigg) \end{aligned}$$

with $x_0 \in [-1/4, 3/4]$. Let D be the orthogonal matrix given by (3. 36). Then

$$DxD^{-1} = \begin{pmatrix} 2x_0 - 1/2 & -(2x_1 + 1/2) & 0 & 0 \\ 2x_1 + 1/2 & 2x_0 - 1/2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

where $x_0 = \frac{1}{4} + \frac{1}{2}\cos\theta$ and $x_0 \in [-1/4, 3/4]$. Again, in this case

$$\begin{aligned} K &= \Bigg\{circ\Bigg(x_0, -\frac{1}{4} \pm \frac{1}{2}\sqrt{(3-4x_0)(1+4x_0)}, 1/2 - x_0, \\ &\qquad -\frac{1}{4} \mp \frac{1}{2}\sqrt{(3-4x_0)(1+4x_0)}\Bigg)\Bigg\}, \end{aligned}$$

with $x_0 \in [-1/4, 3/4]$, is a compact group topologically isomorphic to the circle group. The identity of K is

$$e_{32} = circ(3/4, -1/4, -1/4, -1/4),$$

the other idempotent rank 3 circulant.

Now note that when $z = (-1, 0, 1)$ we get the equations

$$-1/2 = x_0 + x_2 = x_1 + x_3,$$

and

$$(2x_0 + 1/2)^2 + (2x_1 + 1/2)^2 = 1.$$

If K is to exist in this case, then its identity must have $x_0 = 3/4$. The last equation, however, leads to

$$4 + (2x_1 + 1/2)^2 = 1,$$

which does not have a real solution for x_1. We conclude that we cannot have $z = (-1, 0, 1)$.

The same argument may be used to show the nonexistence of K when $z = (0, -1, 1)$.

(iv) Suppose the matrices in K all have rank 4. Then $z = (\pm 1, \pm 1, 1)$.

When $z = (1, 1, 1)$, we get the equations

$$1 = x_0 + x_2,\ x_3 = -x_1,$$

$$1 = (2x_0 - 1)^2 + (2x_1)^2.$$

Thus,

$$x_1 = \pm\sqrt{x_0(1 - x_0)}.$$

Therefore, K consists of circulants of the form

$$\begin{aligned} x &= circ(x_0, x_1, 1 - x_0, -x_1) \\ &= circ\left(x_0, \pm\sqrt{x_0(1-x_0)}, 1 - x_0, \mp\sqrt{x_0(1-x_0)}\right),\ x_0 \in [0, 1]. \end{aligned}$$

Let D be the orthogonal matrix given by (3. 36). Then

$$DxD^{-1} = \begin{pmatrix} 2x_0 - 1 & -2x_1 & 0 & 0 \\ 2x_1 & 2x_0 - 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

where $x_0 = \frac{1+\cos\theta}{2} = \cos^2(\theta/2)$ and $x_0 \in [0, 1]$. Once again, in this case

$$K = \left\{ circ\left(x_0, \pm\sqrt{x_0(1-x_0)}, 1 - x_0, \mp\sqrt{x_0(1-x_0)}\right) \,:\, x_0 \in [0, 1] \right\}$$

is a compact group topologically isomorphic to the circle group. The identity of K is $I_4 = circ(1, 0, 0, 0)$.

An argument similar to the one given above for the rank 3 case may be used to deduce that the cases when z is equal to either $(1, -1, 1)$, $(-1, 1, 1)$ or $(-1, -1, 1)$ are not possible.

The preceding discussion leads to the following

THEOREM 3.2 *Let μ be a probability measure on 4×4 real circulant matrices. Let S be the closed commutative semigroup generated by the support S_μ of μ, and let K be the kernel of S. Suppose $(\mu^n)_{n\geq 1}$ is tight. Then*

(i) If K consists of rank 1 matrices, then either

(a) $K = \{e_{11}\}$ and $\mu^n \xrightarrow{w} \delta_{e_{11}}$, or

(b) $K = \{e_{11}, -e_{11}\}$ and (μ^n) converges weakly if and only if

$$S_\mu \not\subseteq \{circ(x_0, x_1, x_2, -1 - x_0 - x_1 - x_2) : x_0, x_1, x_2 \in \mathbb{R}\}, \quad \textit{or}$$

(c) $K = \{e_2\}$ and $\mu^n \xrightarrow{w} \delta_{e_{12}}$, or

(d) $K = \{e_{12}, -e_{12}\}$ and (μ^n) converges weakly if and only if

$$S_\mu \not\subseteq \{circ(x_0, x_1, x_2, 1 + x_0 - x_1 + x_2) : x_0, x_1, x_2 \in \mathbb{R}\}$$

Here, $e_{11} = circ(1/4, 1/4, 1/4, 1/4)$ and $e_{12} = circ(1/4, -1/4, 1/4, -1/4)$. δ_x is the point mass at x.

(ii) If K consists of rank 2 matrices, then either

(a) $K = \{e_{21}\}$ and $\mu^n \xrightarrow{w} \delta_{e_{21}}$, or

(b) K is isomorphic to a nontrivial subgroup of $\{1, -1\}^2$, and (μ^n) converges weakly if and only if

$$S_\mu \not\subseteq \{circ(x_0, x_1, -1 - x_0, -x_1) : x_0, x_1 \in \mathbb{R}\} \cup$$
$$\{circ(x_0, x_1, -x_0, \pm 1 - x_1) : x_0, x_1 \in \mathbb{R}\}, \qquad \textit{or}$$

(c) K is topologically isomorphic to the circle group T (or a finite subgroup T_0 of T) and (μ^n) converges weakly if and only if $e_{22}S_\mu$ is not contained in a proper coset of a finite subgroup of K.

Here, $e_{21} = circ(1/2, 0, 1/2, 0)$ *and* $e_{22} = circ(1/2, 0, -1/2, 0)$.

(iii) If K *consists of rank* 3 *matrices, then either*

(a) $K = \{e_{31}\}$ *and* $\mu^n \xrightarrow{w} \delta_{e_{31}}$*, or*

(b) $K = \{e_{31}, -e_{31}\}$ *and* (μ^n) *converges weakly if and only if*

$$S_\mu \nsubseteq \{circ(x_0, -1 - x_0, 1 + x_0, -1 - x_0) \,:\, x_0 \in \mathbb{R}\}, \qquad \text{or}$$

(c) $K = \{e_{32}\}$ *and* $\mu^n \xrightarrow{w} \delta_{e_{32}}$*, or*

(d) $K = \{e_{32}, -e_{32}\}$ *and* (μ^n) *converges weakly if and only if*

$$S_\mu \nsubseteq \{circ(x_0, 1 + x_0, 1 + x_0, 1 + x_0) \,:\, x_0 \in \mathbb{R}\}, \qquad \text{or}$$

(e) K *is topologically isomorphic to the circle group* T *(or a finite subgroup* T_0 *of* T*) and* (μ^n) *converges weakly if and only if neither* $e_{31}S_\mu$ *nor* $e_{32}S_\mu$ *is contained in a proper coset of a finite subgroup of* K.

Here, $e_{31} = circ(3/4, 1/4, -1/4, 1/4)$ *and* $e_{32} = circ(3/4, -1/4, -1/4, -1/4)$.

(iv) If K *consists of rank* 4 *matrices, then either*

(a) $K = \{I_4\}$ *and* $\mu^n \xrightarrow{w} \delta_{I_4}$*, or*

(b) $K = \{I_4, -I_4\}$ *and* (μ^n) *converges weakly if and only if*

$$S_\mu \neq \{-I_4\}, \qquad \text{or}$$

(c) $K = S$ *topologically isomorphic to the circle group* T *(or a finite subgroup* T_0 *of* T*) and* (μ^n) *converges weakly if and only if* S_μ *is not contained in a proper coset of a finite subgroup of* S.

3.4 $d \times d$ Toeplitz Matrices

In this section, we extend the investigation of the weak convergence problem to a wider class of certain Toeplitz matrices, which are matrices whose elements along a diagonal are constant.

A $d \times d$ Toeplitz matrix $x = (x_{ij}), i, j = 0, 1, \ldots, d-1$, is completely determined by $2d-1$ constants $x_k,\ k = 0, \pm 1, \pm 2, \ldots, \pm(d-1)$ satisfying

$$x_{ij} = x_{j-i}.$$

Thus, if we index the diagonals of x, from the lower left to the upper right, using the integers from $-(d-1)$ to $d-1$, then the common value of the entries in the kth diagonal is precisely x_k, $k = -(d-1), \ldots, d-1$.

Here, we consider the class $\mathcal{S}_d(t)$ of $d \times d$ Toeplitz matrices satisfying

$$x_{-k} = x_{d-k}\, t \qquad (k = 1, 2, \ldots, d-1)$$

where t is a fixed parameter. That is,

$$\mathcal{S}_d(t) = \left\{ \begin{pmatrix} x_0 & x_1 & x_2 & \cdots & x_{d-2} & x_{d-1} \\ x_{d-1}t & x_0 & x_1 & \cdots & x_{d-3} & x_{d-2} \\ x_{d-2}t & x_{d-1}t & x_0 & \cdots & x_{d-4} & x_{d-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ x_2 t & x_3 t & x_4 t & \cdots & x_0 & x_1 \\ x_1 t & x_2 t & x_3 t & \cdots & x_{d-1}t & x_0 \end{pmatrix} : x_0, x_1, \ldots, x_{d-1} \in \mathbb{R} \right\}.$$

Note that $\mathcal{S}_d(1)$ is the class of circulant matrices. Note also that $\mathcal{S}_d(t)$ is a commutative semigroup under ordinary matrix multiplication.

Next we discuss the spectral representation of a matrix x in $\mathcal{S}_d(t)$. For brevity, we shall identify x using its first row elements, and write

$$x = toep(x_0, x_1, \ldots, x_{d-1})$$

to mean

$$x = \begin{pmatrix} x_0 & x_1 & x_2 & \cdots & x_{d-2} & x_{d-1} \\ x_{d-1}t & x_0 & x_1 & \cdots & x_{d-3} & x_{d-2} \\ x_{d-2}t & x_{d-1}t & x_0 & \cdots & x_{d-4} & x_{d-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ x_2t & x_3t & x_4t & \cdots & x_0 & x_1 \\ x_1t & x_2t & x_3t & \cdots & x_{d-1}t & x_0 \end{pmatrix}.$$

Consider the $d \times d$ matrix P whose jth column, $1 \leq j \leq d$, is the column with entries

$$(1, s\omega^{j-1}, s^2\omega^{2(j-1)}, \ldots, s^{d-1}\omega^{(d-1)(j-1)}), \tag{3. 38}$$

where $\omega = \exp(2\pi i/d)$ and $s = \sqrt[d]{t}$. Observe that the columns of P are linearly independent. Now define

$$\lambda_j = \sum_{k=0}^{d-1} x_k s^k \omega^{k(j-1)} \qquad (j = 1, 2, \ldots, d). \tag{3. 39}$$

Then x has the following spectral representation:

$$x = PD_xP^{-1}, \tag{3. 40}$$

where D_x is the $d \times d$ diagonal matrix diag$(\lambda_1, \lambda_2, \ldots, \lambda_d)$.

In what follows, we assume that the sequence (μ^n), where μ is a probability measure on the Borel sets of $\mathcal{S}_d(t)$, is tight, so that the assertions in Lemma 3.1 again hold. If S_μ is the support of μ and

$$S = \overline{\bigcup_{n \geq 1} S_\mu^n}, \tag{3. 41}$$

then, as before, S is a commutative semigroup, the minimal ideal S_λ in Lemma 3.1 is a compact abelian group, λ is the Haar measure of this group, and, as before, we have Lemma 3.2.

We will also need

LEMMA 3.7 *Let* $e = toep(e_0, e_1, \ldots, e_{d-1})$ *be a* $d \times d$ *idempotent matrix in* $\mathcal{S}_d(t)$ *with eigenvalues* $\lambda_0, \lambda_1, \ldots, \lambda_{d-1}$. *Write* $s = \sqrt[d]{t}$. *Then* *i. each* λ_j *is either* 0 *or* 1;

ii. If e *has rank* r, *where* $0 < r < d$, *then* $e_0 = r/d$, *and exactly* r *of the* λ_j*s are equal to* 1. *In particular, if* $r = d-1$ *and* d *is even, there are two possible idempotents:*

$$e = toep((d-1)/d, -s^{-1}/d, -s^{-2}/d, \ldots, -s^{-(d-1)}/d). \tag{3. 42}$$

or

$$e = toep((d-1)/d, s^{-1}/d, -s^{-2}/d, s^{-3}/d, \ldots, s^{(d-2)}/d, -s^{-(d-1)}/d). \tag{3. 43}$$

If $r = d-1$ *and* d *is odd, then there is only one idempotent and it is given by (3. 42).*

If $r = 1$ *and* d *is even, there are two possible idempotents:*

$$e = toep(1/d, s^{-1}/d, s^{-2}/d, \ldots, s^{-d-2}/d, s^{-(d-1)}/d) \tag{3. 44}$$

or

$$e = toep(1/d, -s^{-1}/d, s^{-2}/d, \ldots, s^{-(d-2)}/d, -s^{-(d-1)}/d). \tag{3. 45}$$

If $r = 1$ *and* d *is odd, then there is only one idempotent and it is given by (3. 44).*

Using the spectral representation (3. 40), we observe that Lemma 3.4 still holds for a compact abelian group in $\mathcal{S}_d(t)$.

Following the same procedures as in earlier sections, we again will have analogous results for $S_d(t)$. We omit the details.

Chapter 4
Numerical Calculation of Lyapunov Exponents for Some Random Fibonacci Recurrences

4.1 Random Matrix Products and Random Fibonacci Sequences

We introduce here some concepts from the theory of random matrix products that will be useful in the sequel, specializing in the case of matrices in $M = M(2, \mathbb{R})$, the space of 2×2 real matrices.

For a (column) vector $x = \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2$ and $Y \in M$ we let

$$\begin{aligned} \|x\| &= \sqrt{a^2 + b^2}\,, \\ \|Y\| &= \sup\{\|Yx\| \,:\, x \in \mathbb{R}^2,\ \|x\| = 1\}. \end{aligned}$$

Let $A_1, A_2, \ldots$ be a sequence of i.i.d. random matrices in M with common distribution μ such that the expectation $E(\log^+ \|A_1\|)$, where $f^+ := \sup(f, 0)$, is finite. It is then clear, from the subadditivity of the sequence $\{E(\log \|A_n A_{n-1} \cdots A_1\|) \,:\, n \geq 1\}$, that the limit

$$\lim_{n\to\infty} \frac{E(\log \|A_n A_{n-1} \cdots A_1\|)}{n}$$

exists in $\mathbb{R} \cup \{-\infty\}$. This limit is called the <u>(upper) Lyapunov exponent</u> associated with μ (or with the random sequence $\{A_n \,:\, n \geq 1\}$), and is denoted by γ. A well-known result by Furstenberg and Kesten [FK] gives the stronger result

$$\gamma = \lim_{n\to\infty} \frac{1}{n} \log \|A_n A_{n-1} \cdots A_1\| \tag{4. 1}$$

almost surely.

As an immediate application of Furstenberg and Kesten's result, consider the random sequence

$$\left\{ x_n = \begin{pmatrix} a_n \\ b_n \end{pmatrix} \ : \ n \geq 1 \right\}$$

of column vectors in $\mathbb{R}^2$ such that

$$x_n = A_n x_{n-1}, \qquad n \geq 1,$$

with $x_0 = \begin{pmatrix} a_0 \\ b_0 \end{pmatrix}$ fixed, and $A_1, A_2, \ldots$ is a sequence of i.i.d. random matrices in M with associated Lyapunov exponent γ. Then

$$x_n = A_n A_{n-1} \cdots A_1 x_0,$$

and therefore, almost surely,

$$\limsup_{n\to\infty} |a_n|^{1/n} \leq \limsup_{n\to\infty} \|x_n\|^{1/n} \leq \limsup_{n\to\infty} \|A_n A_{n-1} \cdots A_1\|^{1/n} \left(\sqrt{a_0^2 + b_0^2}\right)^{1/n} = e^{\gamma},$$

so that

$$\limsup_{n\to\infty} \frac{\log |a_n|}{n} \leq \gamma \tag{4. 2}$$

almost surely. The inequality

$$\gamma \leq \liminf_{n\to\infty} \frac{\log |a_n|}{n} \tag{4. 3}$$

holds almost surely as well, although it is not straightforward to prove (see [BL]). Combining (4. 2) and (4. 3) then yields

$$\gamma = \lim_{n\to\infty} \frac{\log |a_n|}{n}, \tag{4. 4}$$

almost surely.

Next, we introduce Furstenberg's Theorem, one of whose conclusions is a nice integral formula for the Lyapunov exponent γ associated with certain random matrix sequences.

Two nonzero vectors $x, y \in \mathbb{R}^2$ <u>have the same direction</u> if for some $\lambda \in \mathbb{R}$, $x = \lambda y$. This defines an equivalence relation $\sim$ on $\mathbb{R}^2 \setminus \{0\}$, and each equivalence class under $\sim$ is called a <u>direction</u>. The set of directions is called the <u>projective space</u>, denoted by $P = P(\mathbb{R}^2)$. For $x \in \mathbb{R}^2 \setminus \{0\}$, let $\overline{x} \in P$ denote its direction.

Let $G = GL(2, \mathbb{R})$ be the group of nonsingular matrices in M. Observe that G acts on P via $Y \cdot \overline{x} = \overline{Yx}$, where $Y \in G$ and and $x \in \mathbb{R}^2 \setminus \{0\}$. This allows us to define a "convolution" product $\mu * \nu$ between a probability measure μ on G and a probability measure ν on P, defined as the distribution on P satisfying

$$\int_P f(\overline{x})\, d(\mu * \nu)(\overline{x}) = \int_P \int_G f(Y \cdot \overline{x})\, d\mu(Y)\, d\nu(\overline{x})$$

for any bounded Borel function f on P. ν is said to be μ-<u>invariant</u> if

$$\mu * \nu = \nu.$$

ν is said to be <u>continuous</u> if $\nu(\{\overline{x}\}) = 0$ for all $\overline{x} \in P(\mathbb{R}^2)$.

We now state without proof Furstenberg's Theorem (see, for example, [BL], pp. 53-54):

THEOREM 4.1 *Let μ be a probability measure on G, the group of 2×2 invertible real matrices. Let G_μ be the smallest closed subgroup of G which contains the support of μ. Suppose that the following hold:*

(i) for $Y \in G_\mu$, $|\det Y| = 1$,

(ii) G_μ is not compact, and

(iii) for any $x \in \mathbb{R}^2 \setminus \{0\}$, $|\{\overline{Yx} : Y \in G_\mu\}| \geq 2$.

Then if $\{Y_1, Y_2, \ldots\}$ is a sequence of independent μ-distributed random matrices in G with $E(\log^+ \|Y_1\|) < \infty$ and γ as its associated Lyapunov exponent,

$$\gamma = \lim_{n\to\infty} \frac{1}{n} \log \|Y_n \cdots Y_1 x\| > 0$$

for any $x \in \mathbb{R}^2 \setminus \{0\}$, Moreover, there exists a unique continuous μ-invariant measure ν on $P = P(\mathbb{R}^2)$ such that

$$\gamma = \int_P \int_{G_\mu} \log \frac{\|Yx\|}{\|x\|}\, d\mu(Y)\, d\nu(\overline{x}). \tag{4. 5}$$

Now we recall the three types of random Fibonacci sequences we introduced in Chapter 1:

$$x_{n+1} = x_n \pm x_{n-1}, \qquad (4.\ 6)$$

$$x_{n+1} = \pm x_n + x_{n-1}, \qquad (4.\ 7)$$

and

$$x_{n+1} = x_n \pm \beta x_{n-1}, \qquad (4.\ 8)$$

where $x_0 = x_1 = 1$, each $\pm$ sign is chosen independently and $+$ and $-$ occur with probabilities p and $q := 1 - p$, respectively, with $0 < p < 1$. In (4. 8), $\beta > 0$ is a fixed "growth parameter." We refer to (4. 6) and (4. 7) as "Viswanath-type" and to (4. 8) as "Embree-Trefethen-type" random Fibonacci sequences.

For a fixed value of p, let us denote the Lyapunov exponents of the random sequences (4. 6), (4. 7), and (4. 8) by γ_1, γ_2, and γ_β, respectively. It is immediate from Furstenberg and Kesten's result, as we shall see below, that, in fact, γ_1, γ_2, and γ_β are equal to the almost sure limit

$$\lim_{n\to\infty} \frac{\log|x_n|}{n}$$

with x_n given by (4. 6), (4. 7), and (4. 8), respectively. Our aim in this chapter is to numerically investigate the behavior of γ_1, γ_2, and γ_β as a function of p. We follow [Vi] for the theoretical part and [ET] for the numerical part of our solution.

We begin with the Viswanath-type recurrences (4. 6) and (4. 7) and express these equations using matrices. For (4. 6) we have the matrix equation

$$\begin{aligned}\begin{pmatrix} x_n \\ x_{n+1} \end{pmatrix} &= \begin{pmatrix} 0 & 1 \\ \pm 1 & 1 \end{pmatrix} \begin{pmatrix} x_{n-1} \\ x_n \end{pmatrix} \\ &= M_n^{(1)} M_{n-1}^{(1)} \cdots M_1^{(1)} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \end{aligned} \qquad (4.\ 9)$$

where $M_1^{(1)}, M_2^{(1)}, \ldots$ in (4. 9) are i.i.d. matrices such that $M_1^{(1)}$ is either

$$A_+ = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$

with probability p, or

$$A_- = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}$$

with probability q. For (4. 7), the corresponding matrix formulation is given by

$$\begin{aligned} \begin{pmatrix} x_n \\ x_{n+1} \end{pmatrix} &= \begin{pmatrix} 0 & 1 \\ 1 & \pm 1 \end{pmatrix} \begin{pmatrix} x_{n-1} \\ x_n \end{pmatrix} \\ &= M_n^{(2)} M_{n-1}^{(2)} \cdots M_1^{(2)} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \end{aligned} \tag{4. 10}$$

where $M_1^{(2)}, M_2^{(2)}, \ldots$ in (4. 10) are i.i.d. matrices such that $M_1^{(2)}$ is either

$$B_+ = A_+$$

with probability p, or

$$B_- = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}$$

with probability q.

Since the matrix products appearing in (4. 9) and (4. 10) are products of i.i.d. random matrices, it follows from Furstenberg and Kesten's result that γ_1 (resp. γ_2) is, in fact, equal to the (upper) Lyapunov exponent associated with the random sequence $M_1^{(1)}, M_2^{(1)}, \ldots$ (resp. $M_1^{(2)}, M_2^{(2)}, \ldots$). [BL]. And since Furstenberg's theorem applies in this case, we obtain, from equation (4. 5), the integral formulas

$$\gamma_1 = \int_{-\infty}^{\infty} a(m, p, 1)\, d\nu_1(m), \tag{4. 11}$$

and

$$\gamma_2 = \int_{-\infty}^{\infty} a(m, p, 1)\, d\nu_2(m), \tag{4. 12}$$

where

$$a(m, p, t) = \frac{1}{2} \left(p \log \frac{m^2 + (m+t)^2}{m^2 + (m-t)^2} + \log \frac{m^2 + (m-t)^2}{1 + m^2} \right), \tag{4. 13}$$

and where ν_1 and ν_2 are the unique invariant continuous probability measure for the random walk on directions $\overline{x}$ in the plane (parametrized either by slopes $m \in (-\infty, \infty]$ or by angles

$\theta \in (-\pi/2, \pi/2]$) induced respectively by the probability distributions μ_1 and μ_2 on 2×2 (real) matrices satisfying

$$\mu_1(A_+) = p, \qquad \mu_1(A_-) = q, \tag{4. 14}$$

and

$$\mu_2(B_+) = p, \qquad \mu_2(B_-) = q. \tag{4. 15}$$

The problem, however, is that we do not have a closed form expression for the invariant measures ν_1 and ν_2 to evaluate the integrals (4. 11) and (4. 12) when $p \neq 1/2$. Thus, we proceed numerically as follows.

Since ν_1 and ν_2 satisfy

$$\nu_1 = \mu_1 * \nu_1$$

and

$$\nu_2 = \mu_2 * \nu_2,$$

and both are measures defined on Borel subsets of $\mathbb{R}$ (as slopes) or $(-\pi/2, \pi/2]$ (as angles), we have the invariance equations

$$\nu_1([a,b]) \;=\; p\,\nu_1\left(\frac{1}{[a,b]-1}\right) + q\,\nu_1\left(\frac{1}{-[a,b]+1}\right), \tag{4. 16}$$

and

$$\nu_2([a,b]) \;=\; p\,\nu_2\left(\frac{1}{[a,b]-1}\right) + q\,\nu_2\left(\frac{1}{[a,b]+1}\right), \tag{4. 17}$$

for any slope interval $[a,b]$ with $\pm 1 \notin (a,b)$, or the corresponding equations

$$\nu_1(I) \;=\; p\,\nu_1\left(\tan^{-1}\frac{1}{(\tan I)-1}\right) + q\,\nu_1\left(\tan^{-1}\frac{1}{-(\tan I)+1}\right), \tag{4. 18}$$

and

$$\nu_2(I) \;=\; p\,\nu_2\left(\tan^{-1}\frac{1}{(\tan I)-1}\right) + q\,\nu_2\left(\tan^{-1}\frac{1}{(\tan I)+1}\right), \tag{4. 19}$$

for any angular interval $I \subset (-\pi/2, \pi/2]$ with $\pm\frac{\pi}{4} \notin I$. The bulk of the numerical calculation is done by "discretizing" equations (4. 18) and (4. 19). We describe the details of this process in the next section.

The same theoretical calculations may be done for random recurrences (4. 8) of Embree-Trefethen type. We skip the details and just present the relevant analogous quantities. With x_n as in (4. 8), and the same assumption of independence of choice of $+$ and $-$ signs with the same respective probabilities p and q as before, and for fixed $\beta > 0$, we obtain ν_β as the invariant measure satisfying

$$\nu_\beta(I) \;=\; p\,\nu_\beta\left(\tan^{-1}\frac{\beta}{(\tan I)-1}\right)+q\,\nu_\beta\left(\tan^{-1}\frac{\beta}{-(\tan I)+1}\right) \qquad (4.\ 20)$$

for any angular interval $I \subset (-\pi/2, \pi/2]$ with $\pm\frac{\pi}{4} \notin I$. As before, we use Furstenberg's Theorem to arrive at the integral formula

$$\gamma_\beta = \int_{-\infty}^{\infty} a(m,p,\beta)\, d\nu_\beta(m), \qquad (4.\ 21)$$

with $a(m,p,\beta)$ as in (4. 13).

In the next section we describe the numerical approximation of γ_1, γ_2, and γ_β for $\beta > 0$.

4.2 Numerical Calculation of the Lyapunov Exponents

The first step in the numerical calculation of the Lyapunov exponents γ_1, γ_2, and γ_β, is the numerical approximation of the corresponding invariant measures ν_1, ν_2, and ν_β. We follow the ideas in [ET] for this numerical approximation.

We subdivide the interval $[-\pi/2, \pi/2]$ into $N = 2^n$ equally spaced angular intervals $I_1, I_2, \ldots, I_N$, each of length $\Delta = \pi/N$. We then approximate each of the angular invariance equations (4. 18), (4. 19), and (4. 20) on the discrete set of intervals $I_j = [-\frac{\pi}{2}+(j-1)\Delta, -\frac{\pi}{2}+j\Delta]$, $j = 1, \ldots, N$, as follows.

Let g be the map

$$g(x,\beta) = \tan^{-1}\frac{\beta}{\tan x - 1}.$$

Observe $g(x,\beta)$ is continuous for all $x \in (-\pi/2, \pi/2]$ except at $x = \pi/4$.We can then write (4. 18), (4. 19), and (4. 20) as

$$\nu_1(I) \;=\; p\,\nu_1(g(I,1)) + q\,\nu_1(-g(I,1)), \qquad (4.\ 22)$$

$$\nu_2(I) \;=\; p\,\nu_2(g(I,1)) + q\,\nu_2(-g(-I,1)) \tag{4. 23}$$

and

$$\nu_\beta(I) \;=\; p\,\nu_\beta(g(I,\beta)) + q\,\nu_\beta(-g(I,\beta)), \tag{4. 24}$$

respectively. Let $\ell(I)$ denote the length of the angular interval $I \subset (-\pi/2, \pi/2]$. Then, writing $\nu_1^{(j)}, \nu_2^{(j)}$, and $\nu_\beta^{(j)}$ to denote the discrete approximations to $\nu_1(I_j)$, $\nu_2(I_j)$, and $\nu_\beta(I_j)$, respectively, we have the equations

$$\nu_1^{(j)} = \frac{1}{\Delta}\sum_{k=1}^{N}\Big(p\,\ell\big(I_k \cap g(I_j,1)\big) + q\,\ell\big(I_k \cap -g(I_j,1)\big)\Big)\nu_1^{(k)}, \tag{4. 25}$$

$$\nu_2^{(j)} = \frac{1}{\Delta}\sum_{k=1}^{N}\Big(p\,\ell\big(I_k \cap g(I_j,1)\big) + q\,\ell\big(I_k \cap -g(-I_j,1)\big)\Big)\nu_2^{(k)}, \tag{4. 26}$$

and

$$\nu_\beta^{(j)} = \frac{1}{\Delta}\sum_{k=1}^{N}\Big(p\,\ell\big(I_k \cap g(I_j,\beta)\big) + q\,\ell\big(I_k \cap -g(I_j,\beta)\big)\Big)\nu_\beta^{(k)} \tag{4. 27}$$

for $j = 1, 2, \ldots, N$. In the system (4. 25) of N linear equations in the N unknowns $\nu_1^{(k)}$, $k = 1, 2, \ldots, N$, the quantity in parentheses found on the right side of the jth equation, $j = 1, 2, \ldots, N$, represents the amount of overlap between I_k and $g(I_j, 1)$ and between I_k and $-g(I_j, 1)$. The corresponding quantities in systems (4. 26) and (4. 27) are explained similarly.

All three $N \times N$ linear systems (4. 25), (4. 26), and (4. 27) are of rank $N-1$, which can be made consistent by replacing the Nth equation by the respective conservation conditions

$$\sum_{j=1}^{N}\nu_1^{(j)} = 1, \qquad \sum_{j=1}^{N}\nu_2^{(j)} = 1, \qquad \sum_{j=1}^{N}\nu_\beta^{(j)} = 1. \tag{4. 28}$$

We also see that these linear systems are sparse: the length $\ell(g(I_j, \beta))$ of the image of the jth subinterval I_j is at most Δ times the maximum value of $|g'(x, \beta)|$ for $x \in [-\pi/2, \pi/2]$, which is

$$\frac{2 + \beta^2 + \sqrt{4 + \beta^4}}{2\beta}. \tag{4. 29}$$

Therefore, the number of nonzero coefficients in (4. 27) is $O(\beta)$ if $\beta > 1$, and $O(1/\beta)$ if $\beta < 1$. If $\beta = 1$, then (4. 29) evaluates to $1 + g$, where $g = \frac{1+\sqrt{5}}{2}$ is the golden ratio, and consequently there are at most 4 nonzero coefficients in (4. 25) or (4. 26).

Thus, in the end, our numerical approximation to the invariant measures ν_1, ν_2, and ν_β consists of solving the sparse linear systems (4. 25) to (4. 27) together with their respective conservation conditions displayed in (4. 28).

We solved these linear systems using Mathematica 5.2's built-in sparse systems solver, taking N to 2^{21} when $p = 1/2$ to compare our results with the numerical values reported by Embree and Trefethen in [ET]. For this part of the computational process, we would like to acknowledge the use of the services provided by Research Computing, University of South Florida.

Once we found that our results were consistent with Embree and Trefethen's, we lowered the value of N to 256 for all other values of p in the calculations that produced the figures we report here.

Figures 1 and 2 show the respective graphs – actually, histograms – of the invariant measures ν_1 and ν_2 computed in the manner just described for $p = 0.1, 0.2, \ldots, 0.9$. Figures 3 to 6 display the same information for invariant measures ν_β for $\beta = 1/2, 3/4, 2$, and 8, respectively. These histograms are based on $N = 256$ equally sized subdivisions of $[-\pi/2, \pi/2]$.

The vertical axis in each graph represents the approximated value of the invariant measures on the subintervals I_k, $k = 1, 2, \ldots, N$.

Several interesting observations may be derived from these graphs. First, it is evident from Figure 1 that ν_1 exhibits a reflection property with respect to $p = 1/2$:

$$\nu_1(I_k, p) = \nu_1(-I_k, 1 - p). \tag{4. 30}$$

Also, Figure 2 suggests that the smoothness of the invariant measure ν_2 appears to decrease as p increases. This behaviour is similarly observed for ν_β for a fixed β and, moreover, does not seem to depend on β.

On the other hand, the opposite situation seems to occur when p is fixed and β is allowed to vary. Figures 7 to 10 show how ν_β behaves for different values of $\beta < 1$ when p is fixed at $0.2, 0.4, 0.6$, and 0.8, respectively. Figures 11 to 14 display the same information, but for values of $\beta > 1$.

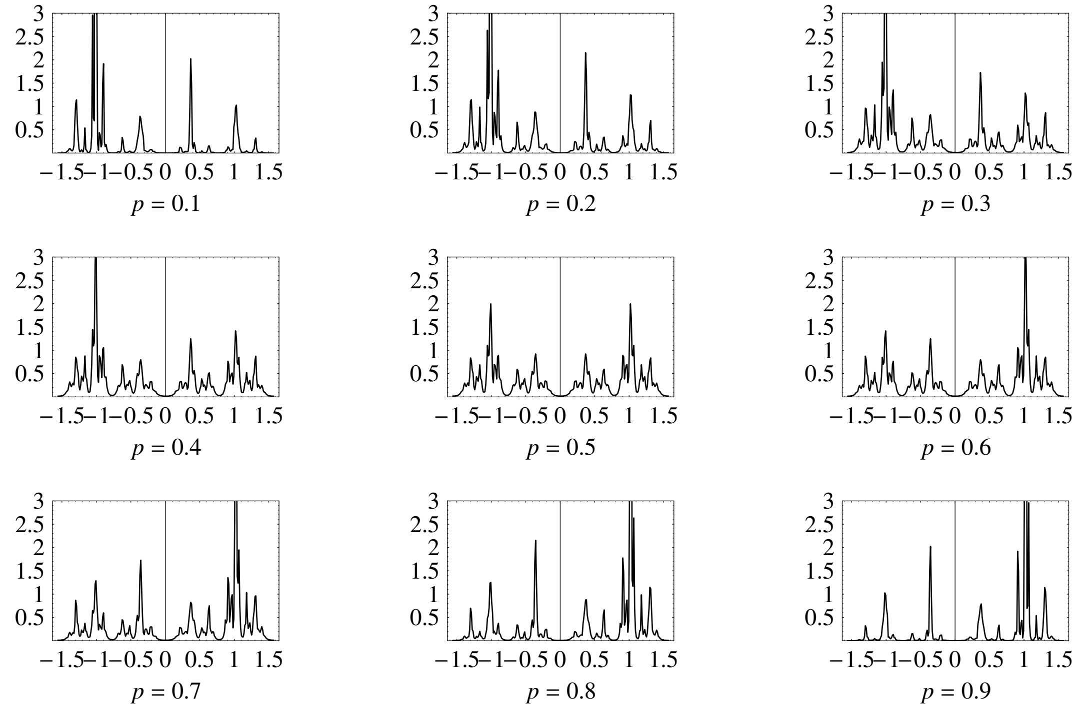

Figure 1: $\nu_1(p)$ for different values of p

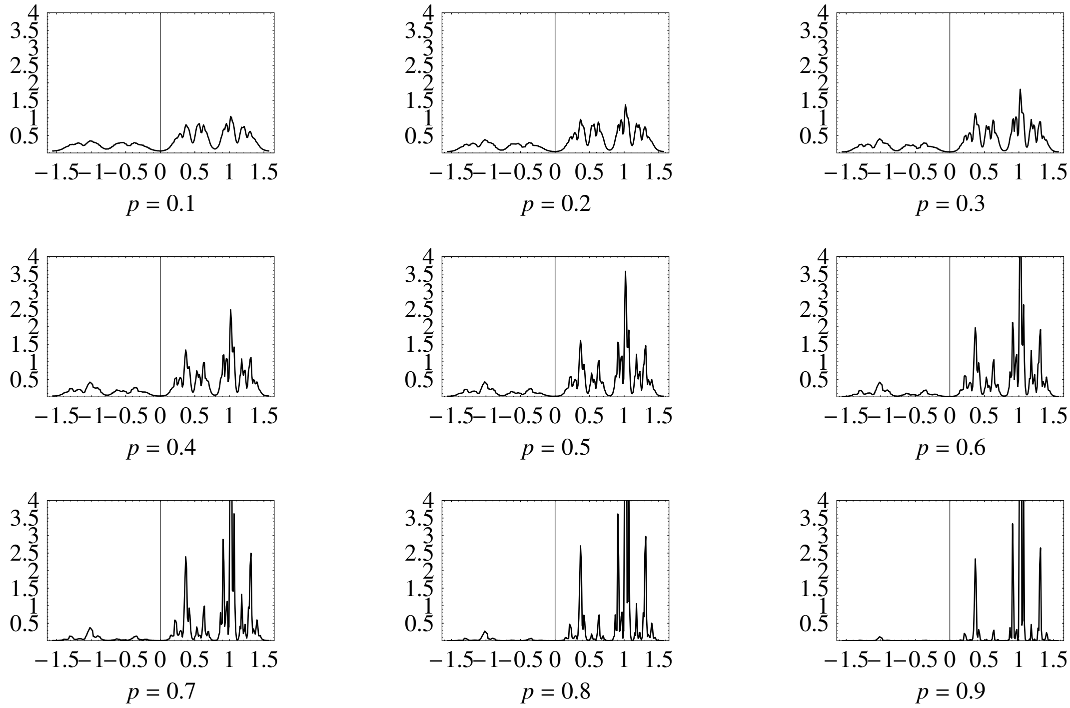

Figure 2: $\nu_2(p)$ for different values of p

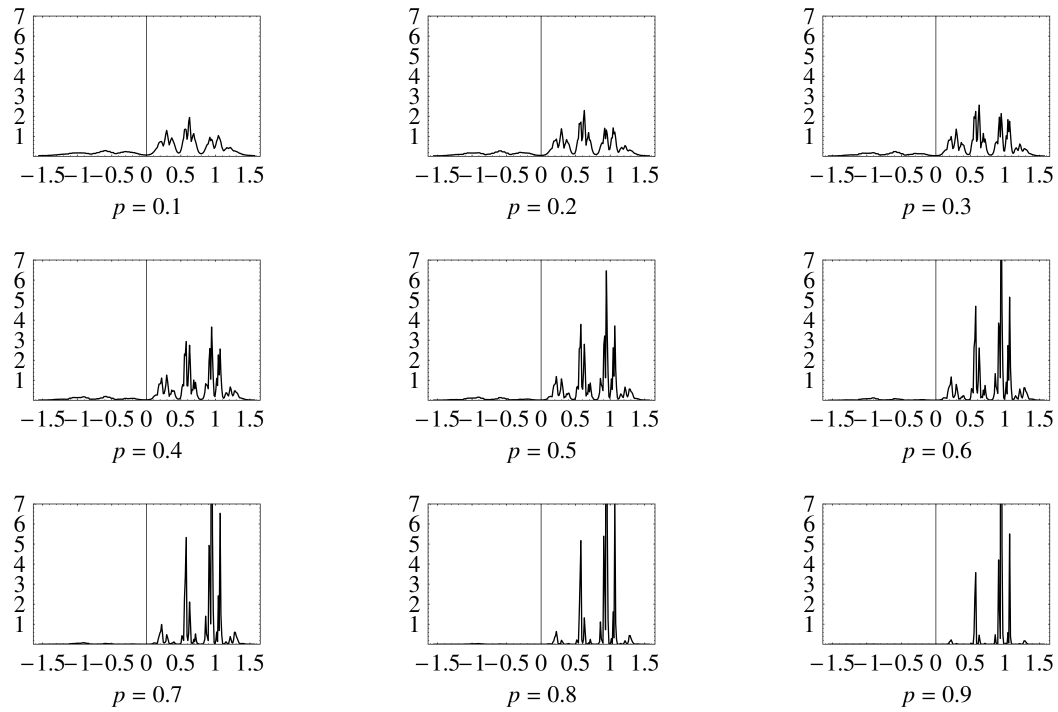

Figure 3: $\nu_\beta(p)$, where $\beta = 1/2$, for different values of p

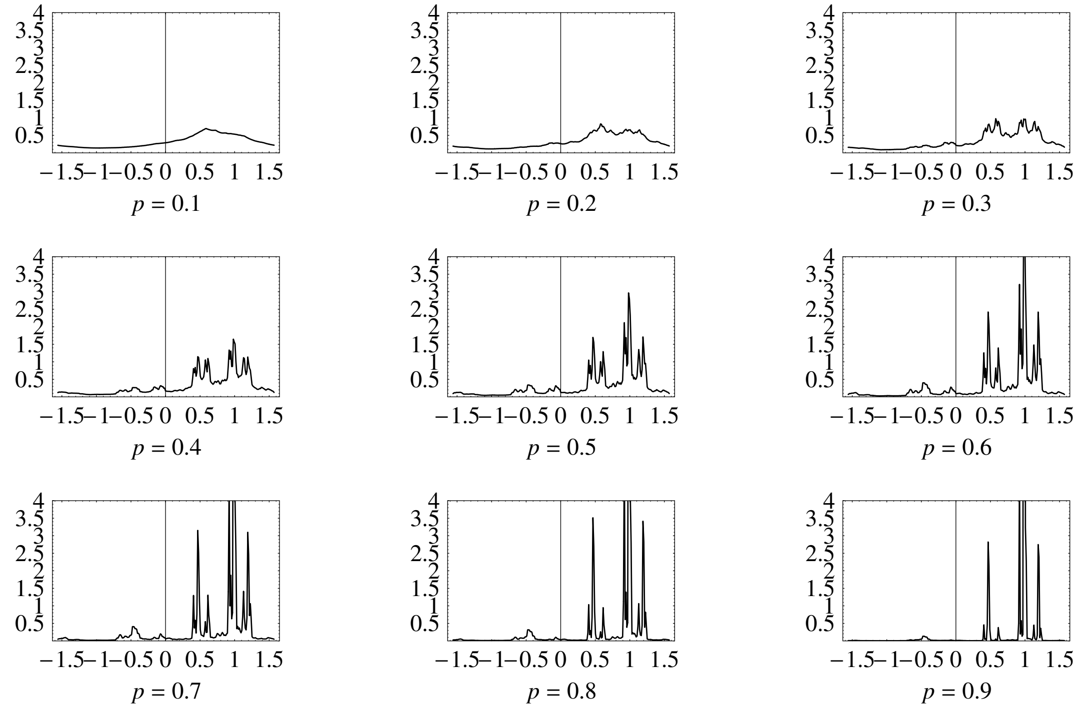

Figure 4: $\nu_\beta(p)$, where $\beta = 3/4$, for different values of p

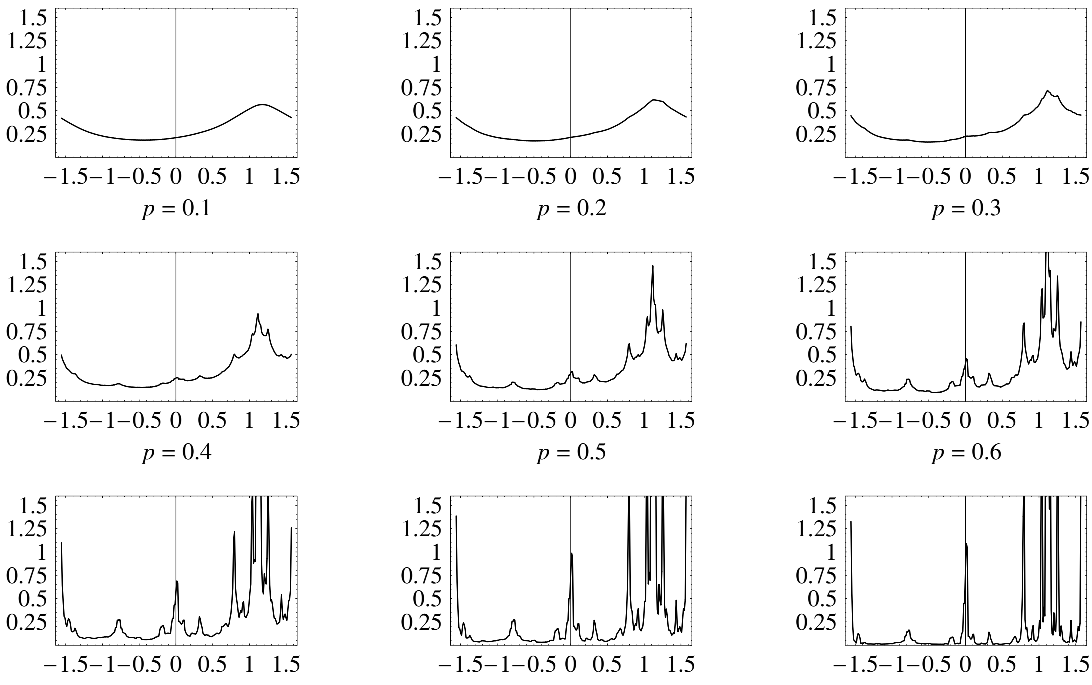

Figure 5: $\nu_\beta(p)$, where $\beta = 2$, for different values of p

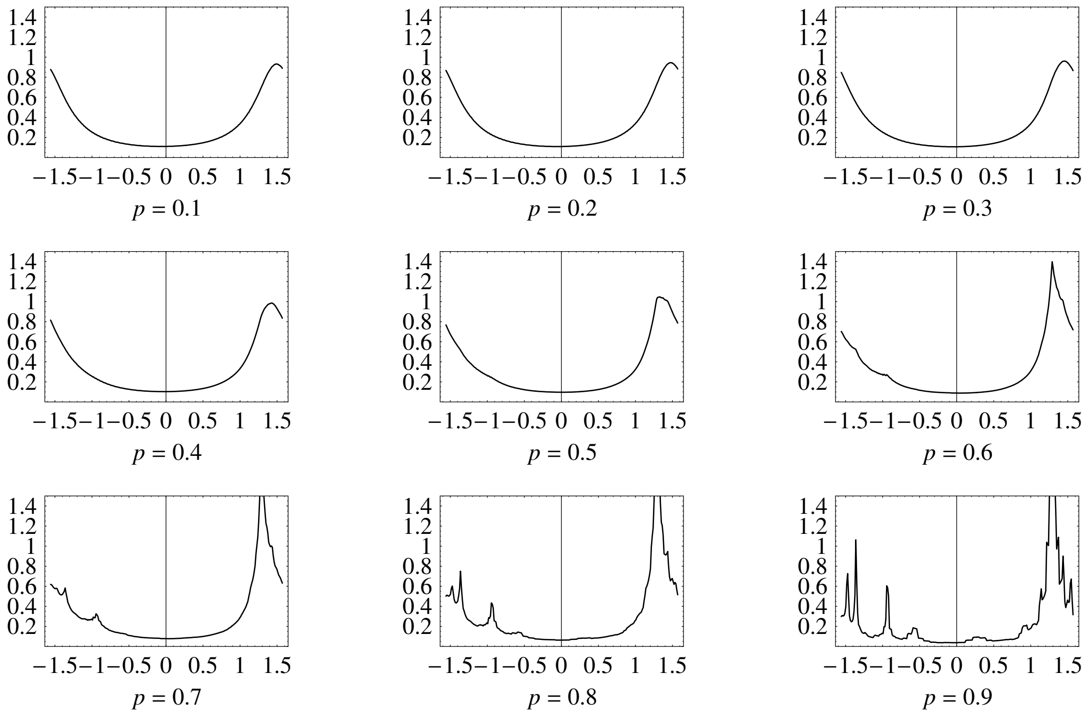

Figure 6: $\nu_\beta(p)$, where $\beta = 8$, for different values of p

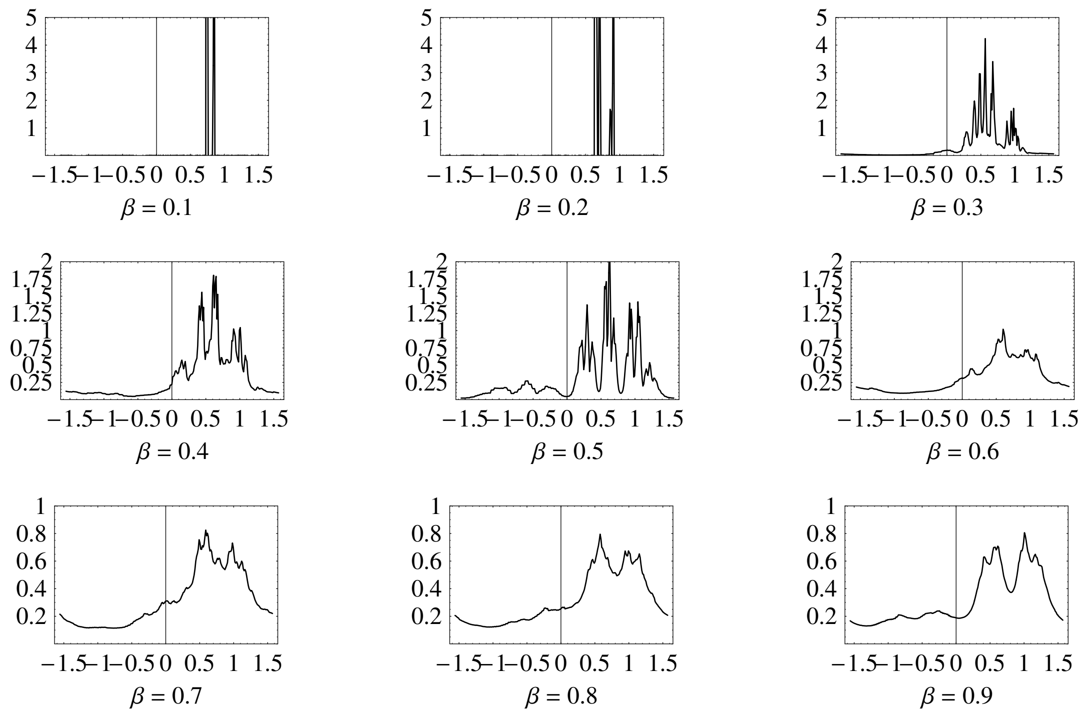

Figure 7: $\nu_\beta(p)$ for fixed $p = 0.2$ and different values of $\beta < 1$

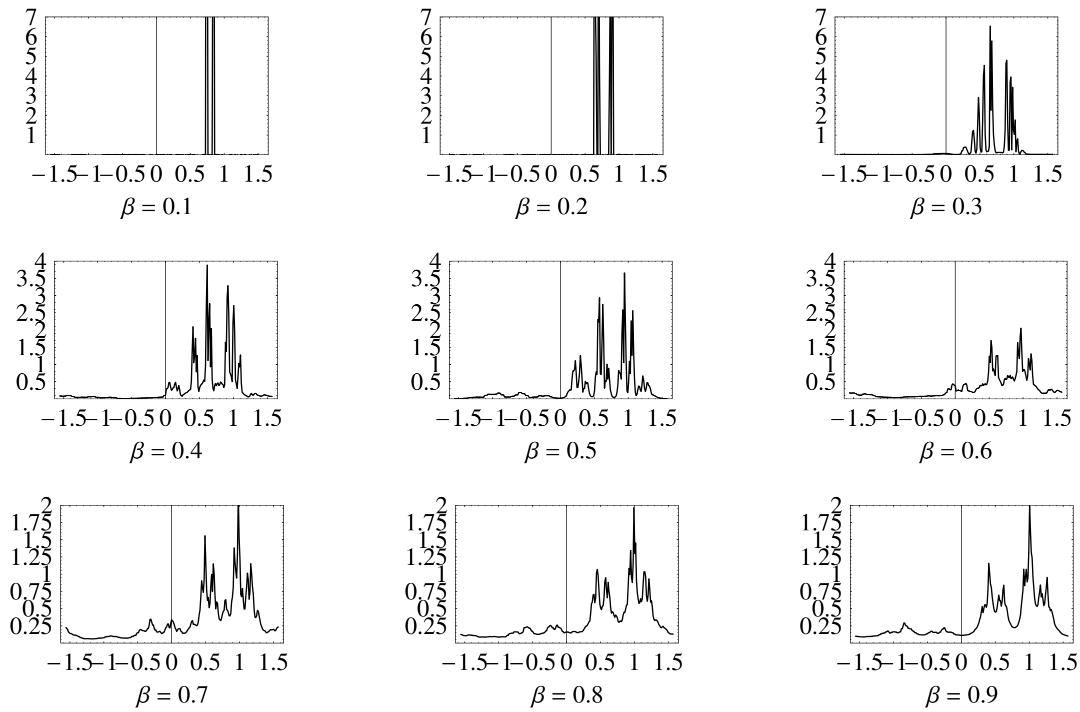

Figure 8: $\nu_\beta(p)$ for fixed $p = 0.4$ and different values of $\beta < 1$

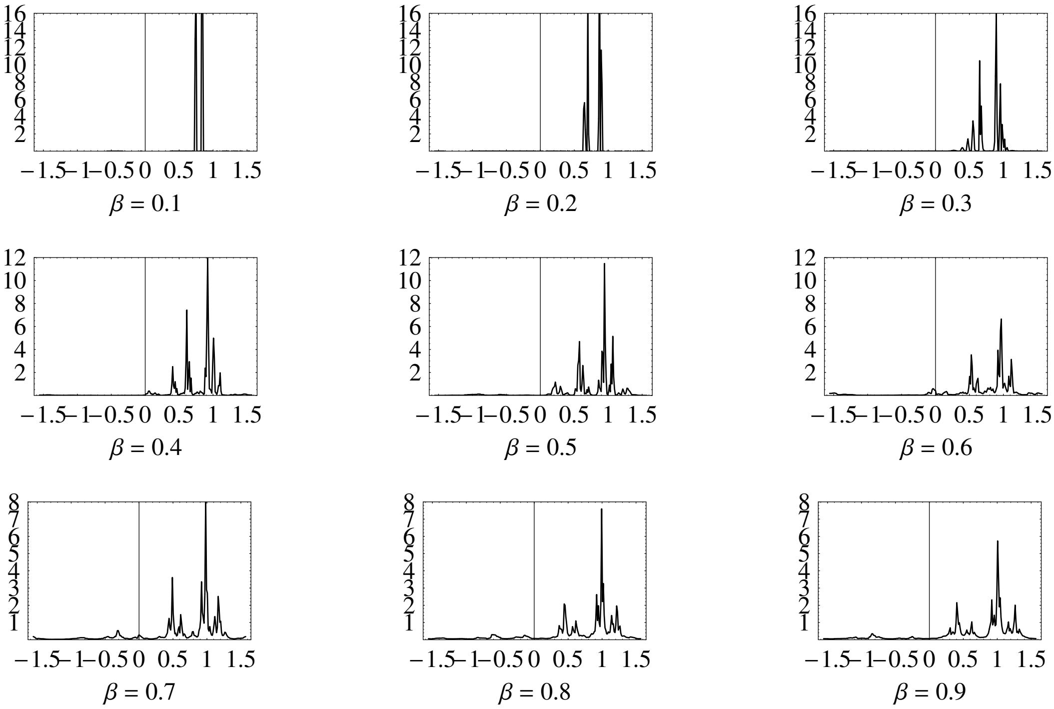

Figure 9: $\nu_\beta(p)$ for fixed $p = 0.6$ and different values of $\beta < 1$

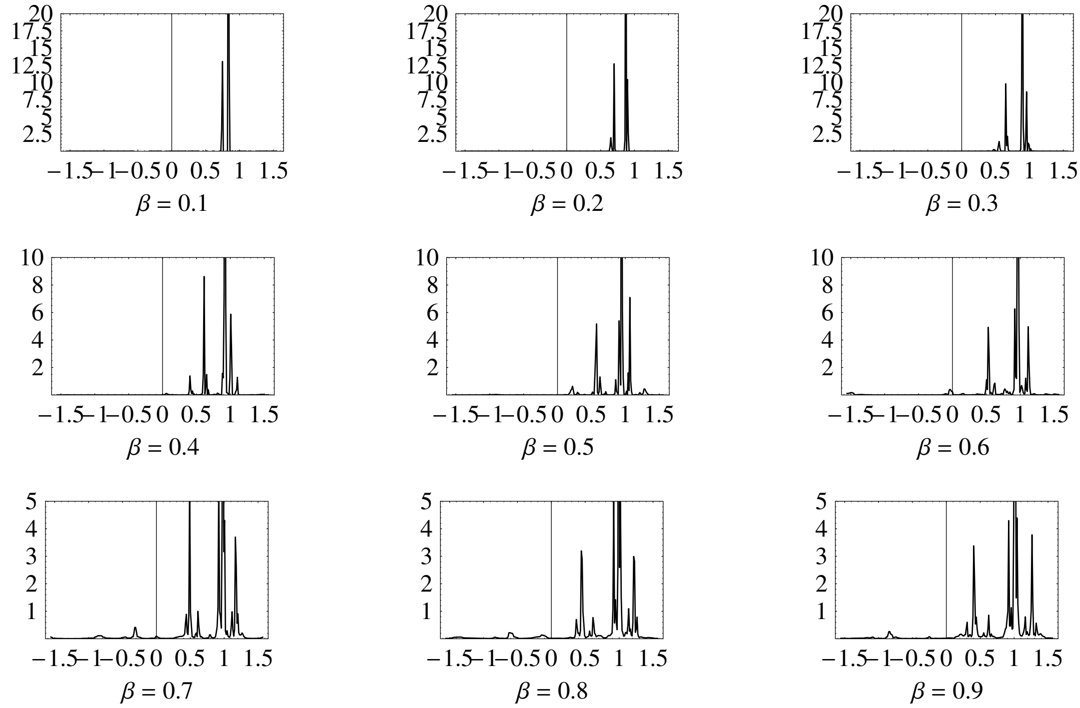

Figure 10: $\nu_\beta(p)$ for fixed $p = 0.8$ and different values of $\beta < 1$

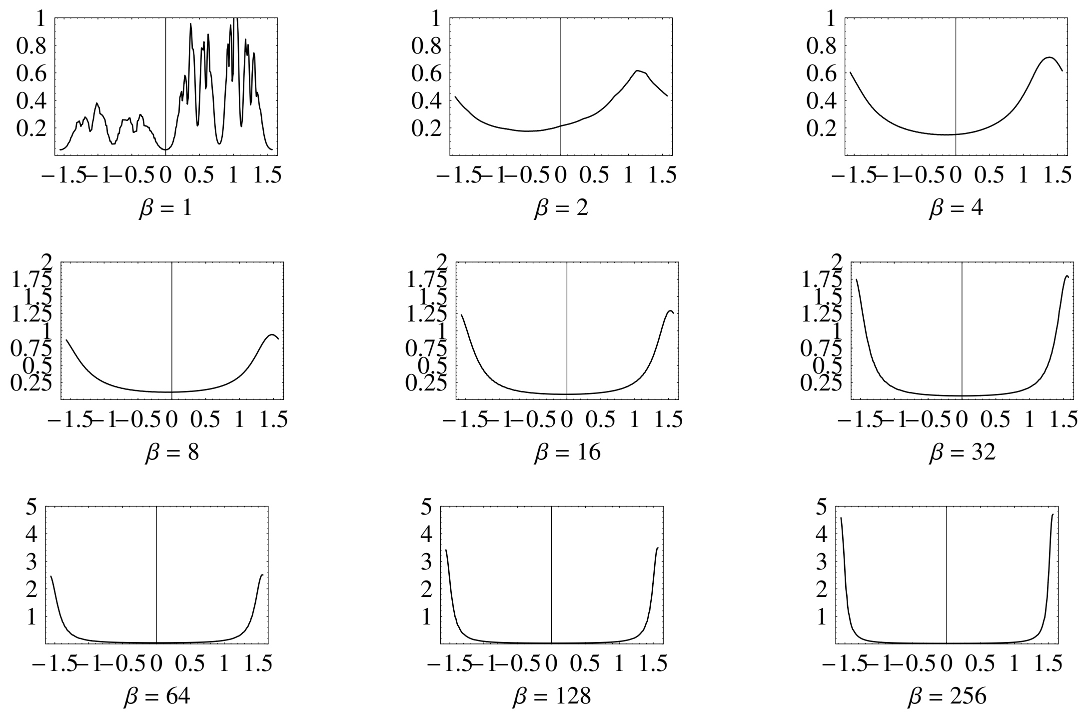

Figure 11: $\nu_\beta(p)$ for fixed $p = 0.2$ and different values of $\beta > 1$

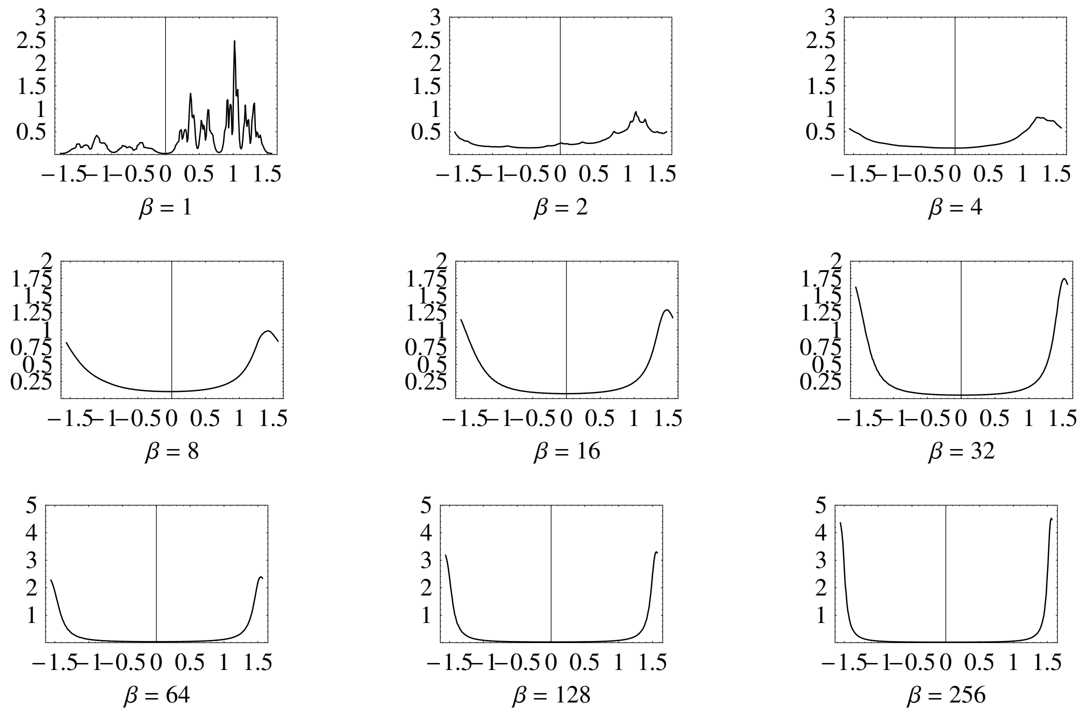

Figure 12: $\nu_\beta(p)$ for fixed $p = 0.4$ and different values of $\beta > 1$

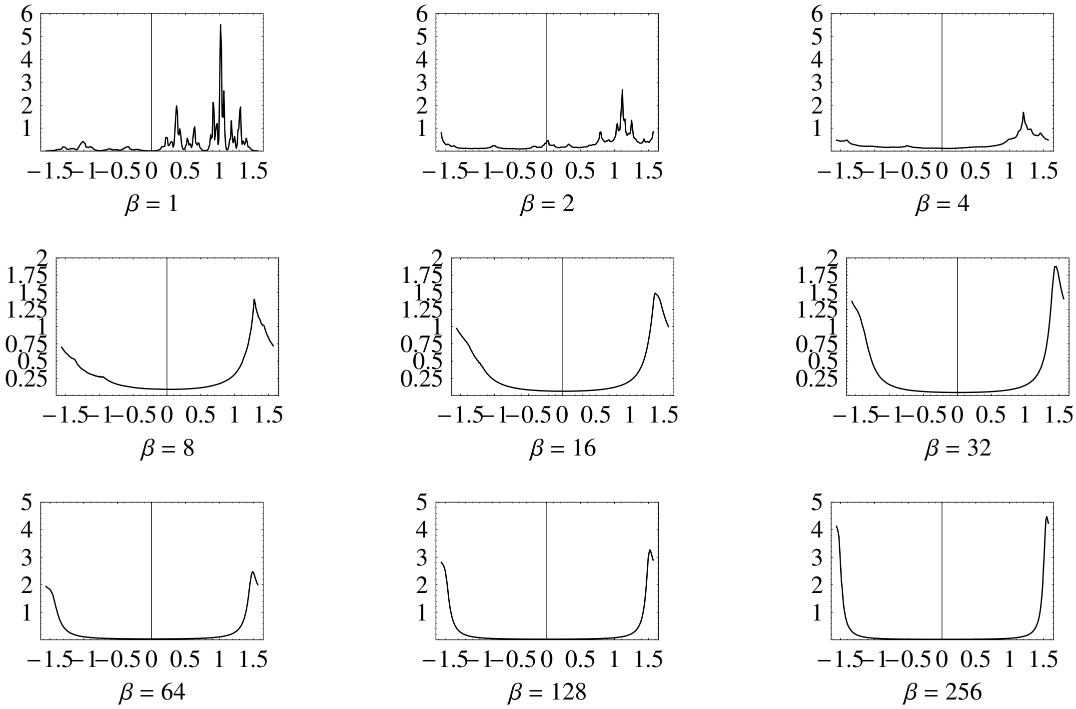

Figure 13: $\nu_\beta(p)$ for fixed $p = 0.6$ and different values of $\beta > 1$

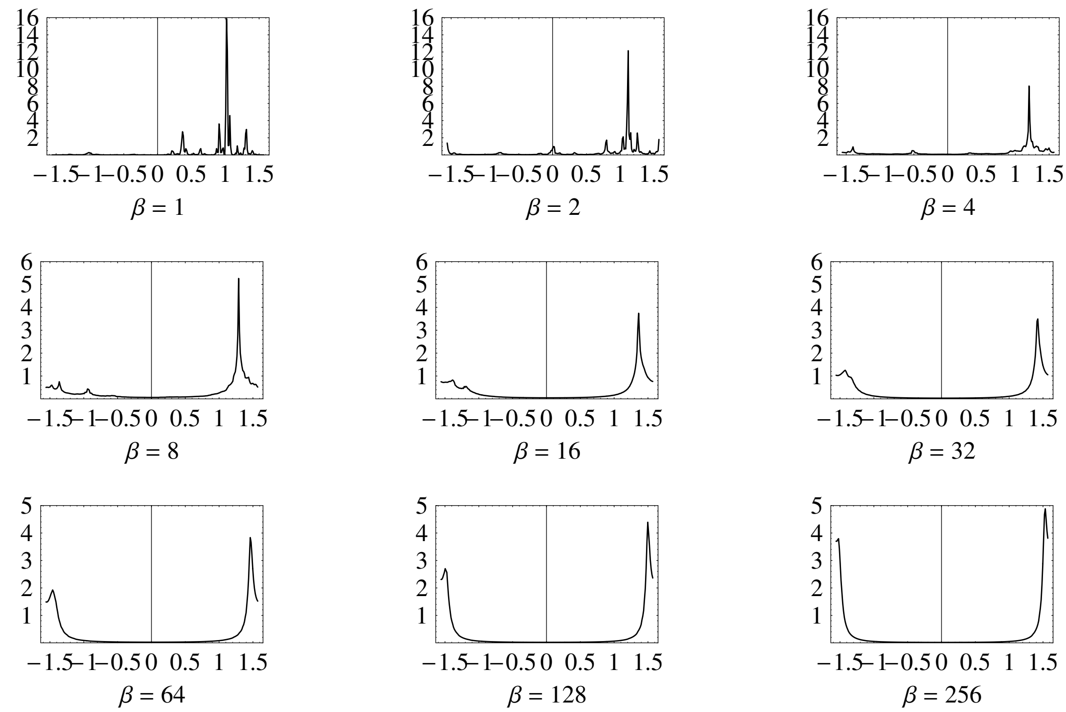

Figure 14: $\nu_\beta(p)$ for fixed $p = 0.8$ and different values of $\beta > 1$

Finally, once we have the approximation to the invariant measures ν_1, ν_2, and ν_β, the corresponding Lyapunov exponents γ_1, γ_2, and γ_β may be calculated by numerical integration applied to (4. 11), (4. 12), and (4. 21), respectively.

Figures 15 and 16 show the Lyapunov exponents γ_1 and γ_2 vs. p for 200 values of p between 0 and 1. Note how γ_2 increases as p increases, whereas γ_1 exhibits symmetry with respect to $p = 1/2$.

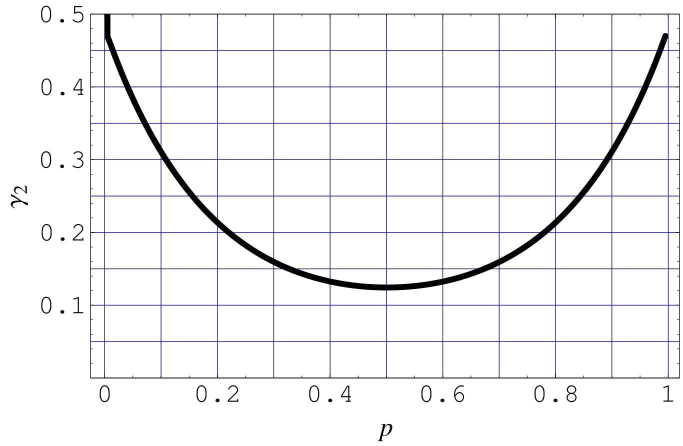

Figure 15: γ_1 vs. p

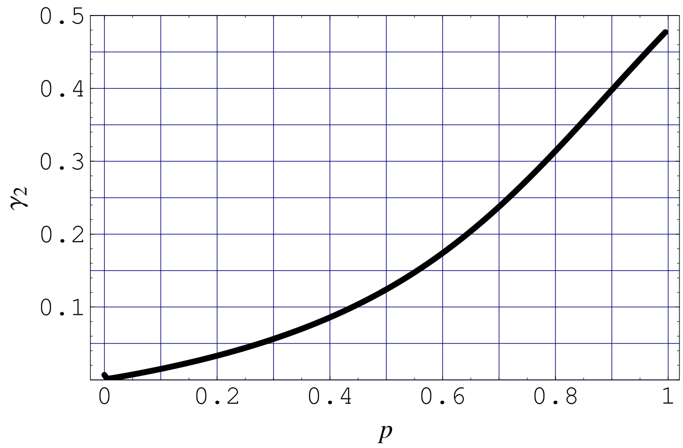

Figure 16: γ_2 vs. p

Figure 17 shows γ_β for $\beta = 0.5, 0.75, 2,$ and 8. In the figure, darker curves correspond to smaller values of β. It appears that for values of $\beta \geq 1$, $\gamma_\beta > 0$ (hence the corresponding random recurrence (4. 8) grows exponentially) no matter what p is. On the other hand, for values of $\beta < 1$, the phenomenon observed by Embree and Trefethen in [ET] appears to have an analogue: there exists some $p^* = p^*(\beta)$ for which $\gamma_\beta(p^*) = 0$, which means the corresponding random recurrence (4. 8) neither grows nor decays.

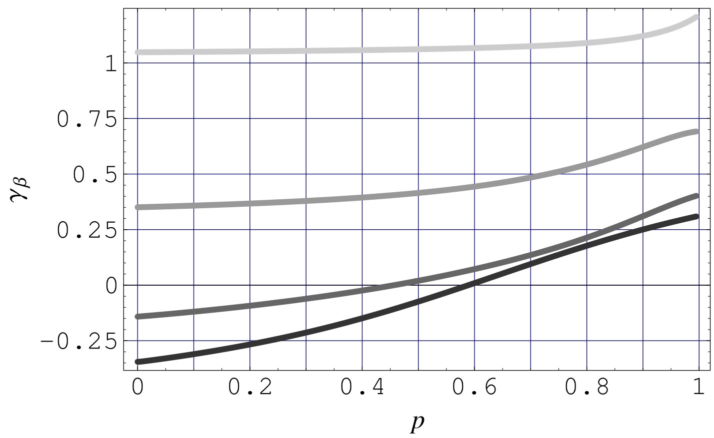

Figure 17: γ_β vs. p for $\beta = 0.5, 0.75, 2,$ and 8

References

[BL] P. Bougerol and J. Lacroix, Random Products of Matrices with Applications to Infinite-Dimensional Schrödinger Operators, Birkhäuser, Basel, 1984.

[CR] S. Chakraborty and B. V. Rao, "Convolution powers of probabilities on stochastic matrices of order 3," Sankhyā: The Indian Journal of Statistics, Vol. 60, Series A, Pt. 2, 151–170, 1998.

[D] P. J. Davis, *Circulant Matrices*, John Wiley & Sons, New York, 1979.

[Di] P. Diaconis, "Patterned Matrices," Proceedings of Symposia in Applied Mathematics 40, 37–58, 1990.

[DM] S. Dhar and A. Mukherjea, "Convergence in Distribution of Products of I.I.D. Nonnegative Matrices," Journal of Theoretical Probability, Vol. 10, No. 2, 375–393, 1997.

[ET] M. Embree and L. Trefethen, "Growth and Decay of Random Fibonacci Sequences," Proc. Roy. Soc. London Ser. A, Math. Phys. Eng. Sci., Volume 455, 2471-2485, 1999.

[FK] H. Furstenberg and H. Kesten, "Products of Random Matrices," Ann. Math. Stat. 31, 457–469, 1960.

[HMu] G. Högnäs and A. Mukherjea, *Probability Measures on Semigroups*, Plenum Press, New York, 1995.

[HMi] F. Y. Hunt, and W. M. Miller, "On the approximation of invariant measures," J. Stat. Phys., Vol. 66, No. 1/2, 535-548, 1992.

[LM] C. C. Lo and A. Mukherjea, “Convergence in Distribution of Products of d by d Random Matrices,” Journal of Mathematical Analysis and Applications, Vol. 162, No. 1, 71–91, 1991.

[M1] A. Mukherjea, “Limit theorems: stochastic matrices, ergodic markov chains and measures on semigroups,” Probabilistic Analysis and Related Topics, Vol. 2, 143-203, 1979.

[M2] A. Mukherjea, “Completely Simple Semigroups of Matrices,” Semigroup Forum, Vol. 33, 405-429, 1986.

[Pe] Y. Peres, “Analytic dependence of Lyapunov exponents on transition probabilities,” Lyapunov Exponents, Lecture Notes in Math. 1486, Springer-Verlag, Berlin, 1986.

[MR1] M. Rosenblatt, *Markov Processes: Structure and Asymptotic Behavior*, Springer-Verlag, 1971.

[MR2] M. Rosenblatt, “Products of i.i.d. stochastic matrices,” Journal of Mathematical Analysis and Applications, Vol. 11, 1–10, 1965.

[Vi] D. Viswanath, “Random Fibonacci Sequences and the Number $1.13198824...$,” Mathematics of Computation, Volume 69, Number 231, 1131–1155, 2000.

About the Author

Edgardo Cureg is a native of the Philippines. In 1984 he entered the University of the Philippines, initially majoring in Computer Science. In 1986, after taking a course in Abstract Algebra under the illustrious Dr. Aurora Trance, he changed his major to Mathematics. He went on to study Information Engineering in Japan from 1988 to 1996 under scholarships offered by Japan's *Monbusho* (now called the Ministry of Education, Culture, Sports, Science and Technology, or MEXT) and IBM Asia. He went back to the Philippines in 1997 and taught Mathematics at De La Salle University in Manila. In 2000, he relocated to the United States and began his graduate education in Mathematics at the University of South Florida, where, in August 2001, he was awarded a Master of Arts degree. That same year he was awarded a three-year Graduate Research Fellowship by the National Science Foundation. Since his admission to doctoral candidacy in 2003 he has been conducting research on products of random matrices under the supervision of Dr. Arunava Mukherjea.