The Annals of Probability
1994, Vol. 22, No. 4, 2223–2234

TIGHTNESS OF PRODUCTS OF I.I.D. RANDOM MATRICES. II

BY ARUNAVA MUKHERJEA

University of South Florida

In this paper, we prove a necessary and sufficient condition for tightness of products of i.i.d. finite-dimensional random real matrices. We give an example illustrating the use of our theorem and treat in more details the case of 2×2 real matrices.

1. Introduction. In this paper we continue our earlier study of tightness of products of i.i.d. random $d \times d$ matrices in [9]. In [9], we proved a necessary and sufficient condition for tightness of products of i.i.d. random nonnegative matrices. The methods in [9] cannot be extended to the general situation of real matrices; however, we show in this paper that it is not difficult to understand to a considerable extent what tightness means for i.i.d. random real matrices and to provide a characterization of tightness for general real matrices. Here we treat the most general situation and assume *no* condition on the matrices or on the support of the distribution involved. Our result here, we feel, is complete, and we do not expect any additional insights without imposing additional conditions.

It is relevant to mention here the pioneering paper of Kesten and Spitzer [5] in the context of tightness and convergence in distribution of products of i.i.d. nonnegative matrices that did set the stage for the next series of papers in this area, those of Bougerol (the paper [1] as well as a number of other papers not mentioned here) and others of the present author (see the references). In [1], Bougerol considered $d \times d$ random *real* matrices and considered the same problem as addressed in this paper. He provided two main theorems, one giving a necessary condition [in the case of general finite d, under various conditions imposed on the support $S(\mu)$ of the distribution μ of the i.i.d. sequence, the condition considered in one case, for example, requiring that the invertible $d \times d$ matrices have μ-measure 1], whereas the other gives a sufficient condition based on some moment conditions on μ as well as certain requirements on certain upper Lyapounov exponents. In [8], the author, in an effort to extend and better understand some of the results of Kesten and Spitzer [5], presented a number of results on weak convergence (the same as convergence in distribution in the present situation) for products of i.i.d. *real* matrices. The methods used in [8, 9] and in the present paper are semigroup methods and very different from the ones employed in [5] or [1]. Even though the methods used here are not the everyday tools of most probabilists, they are extremely simple and easy to follow for any probabilist. The key concept one needs to understand in order to understand this paper is the concept and properties of a completely simple

Received October 1992.
AMS 1991 *subject classifications.* 60B10, 60J15.
Key words and phrases. Random matrices, weak convergence, tight sequence of measures, completely simple semigroup.

semigroup. For this and other semigroup results in the context of probability, we refer the reader to [11].

Let $X_1, X_2, \ldots$ be a sequence of $d \times d$ random i.i.d. *real* matrices with distribution μ with support $S(\mu)$. Then $W_n \equiv X_n X_{n-1} \cdots X_1$, the product being the usual matrix product, has distribution μ^n, the nth convolution power of μ. Recall that for $n \geq 1$,

$$\mu^{n+1}(\mathcal{B}) = \int \mu^n\{z: zx \in \mathcal{B}\}\mu(dx)$$

and

$$S(\mu^n) = \text{closure}\{x_1 x_2 \cdots x_n \mid x_i \in S(\mu) \text{ for each } i\},$$

with usual topology for matrices. Notice that $P(W_n \notin S \text{ for some } n \geq 1) = 0$, if $S = \text{cl}(\bigcup_{n=1}^{\infty} S(\mu^n))$. The sequence (μ^n) is called tight if given $\varepsilon > 0$, there is a compact set $\mathcal{K} \subset S$ such that for each $n \geq 1$, $\mu^n(\mathcal{K}) > 1 - \varepsilon$. The problem of tightness comes up in problems of weak convergence (see [5]) as well as in numerous other contexts, for example, to decide when there is a μ-invariant distribution or to obtain laws of large numbers for the random walk (W_n); see [10] for details.

We present our results in Section 2.

2. Results.

THEOREM 1. *Let μ be a (Borel) probability measure on $d \times d$ real matrices with usual topology, with support* $\text{supp}(\mu)$ *or in short, $S(\mu)$. Let S be the closed multiplicative semigroup generated by $S(\mu)$. Let $m(S)$ be the set defined by*

$$m(S) = \{x \in S: \mathit{rank}\ x \leq \mathit{rank}\ y \mathit{\ for\ all\ } y \in S\}.$$

Let a be the rank of the matrices in $m(S)$. Then the following results hold:

(i) *Suppose that $a = d$. Then the sequence of convolution powers (μ^n) is tight iff S is a compact group.*

(ii) *Suppose that $a = 0$. Then the sequence (μ^n) is tight iff μ^n converges weakly to δ_0, the unit mass at the zero matrix.*

(iii) *Suppose that $0 < a < d$. Then the sequence (μ^n) is tight iff the following two conditions hold*: (*a*) *there is a compact group $\mathcal{G}$ of invertible $a \times a$ matrices and an invertible $d \times d$ matrix y such that for any x in S, the matrix $y^{-1}xy$ can be uniquely represented in the form*

$$y^{-1}xy = \begin{pmatrix} A & BD \\ DC & D \end{pmatrix}, \tag{1}$$

where D is an element of $\mathcal{G}$, C is an $a \times (d-a)$ matrix, B is a $(d-a) \times a$ matrix and A is a $(d-a) \times (d-a)$ matrix, and (*b*) *for any open set V containing the set* *of matrices given by*

$$\mathcal{M} \equiv \left\{ \begin{pmatrix} BDC & BD \\ DC & D \end{pmatrix} \middle| \text{ there exists an } A \text{ such that the matrix in (1) is an element of } y^{-1}Sy \right\},$$

$$\lim_{n\to\infty} \mu^n\big(yVy^{-1}\big) = 1.$$

In the "only if" case, the set $\mathcal{M}$ coincides with the set $m(y^{-1}Sy)$. Therefore, when the sequence (μ^n) is tight, then (μ^n) converges weakly iff there does not exist a proper normal subgroup $\mathcal{H}$ of $\mathcal{G}_0$, where $\mathcal{G}_0$ is the compact group with identity e_0 of $d \times d$ matrices given by $\{(\begin{smallmatrix} 0 & 0 \\ 0 & D\end{smallmatrix})$: $D \in \mathcal{G}\}$ with $\mathcal{G}$ as in condition (1), such that $E(e_0 \cdot y^{-1}Sy)E(y^{-1}Sy \cdot e_0) \subset \mathcal{H}$ and $e_0(y^{-1}S(\mu)y)e_0 \subset g\mathcal{H}$ for some $g \in \mathcal{G}_0 - \mathcal{H}$. [Here, $E(W)$ stands for the set of all idempotent elements in the set W.]

PROOF. Let us prove (i). The "if" part is obvious. For the "only if" part, notice that following the same proof given on page 32 in [10], it follows that

$$\frac{1}{n}\sum_{k=1}^{n} \mu^k \text{ converges weakly to a probability measure } \lambda,$$

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

This implies that the support $S(\lambda)$ of λ is the minimal (completely simple) ideal of S (see Proposition 3.13 and Theorem 3.15 in [10]) and that if e is an idempotent matrix in $S(\lambda)$ (being completely simple, it will contain at least one such matrix), then $e \cdot S(\lambda) \cdot e$ is a compact group (see page 25 in [10]). Because the matrices in S are assumed to be of full rank and because $S(\lambda)$ is an ideal of S, it follows that e is the usual identity matrix and $S = S \cdot e \subset S(\lambda) = e \cdot S(\lambda) \cdot e$.

Let us prove (ii). The "if" part is obvious. For the "only if" part, notice that when $0 \in S$, the support $S(\lambda)$, being the minimal ideal of S, must be the singleton $\{0\}$ so that $\lambda = \delta_0$. Then the weak convergence of μ^n to λ follows from Theorem 2.1 in [6].

For the "if" part of (iii), let us then suppose that conditions (a) and (b) hold. If for each positive integer k, $\mu^k(m(S)) = 0$, then it follows from Lemma 7.7 in [4] that the sequence (μ^n) is tight. Note that condition (b) is used here. Let us then suppose that there is a positive integer k such that $\mu^k(m(S)) > 0$. Because $m(S)$ is an ideal of S, it follows that

$$\lim_{n\to\infty} \mu^n\big(m(S)\big) = 1. \tag{2}$$

By condition (a), there is an invertible matrix y such that for each x in S, $y^{-1}xy$ has representation given in (1). Let us define the set S_1 and the measure λ as

$$S_1 = \{y^{-1}xy \mid x \in S\},$$
$$\lambda(y^{-1}\mathcal{B}y) = \mu(\mathcal{B}), \quad \mathcal{B} \subset S. \tag{3}$$

It is easily verified that for $n \geq 1$,

$$\lambda^n\left(y^{-1}\mathcal{B}y\right) = \mu^n(\mathcal{B})$$

and also

$$m(S_1) = y^{-1} \cdot m(S) \cdot y$$

so that it follows from (2) that

$$\lim_{n\to\infty} \lambda^n\left(m(S_1)\right) = 1. \tag{4}$$

Also, it is clear that the sequence (μ^n) is tight iff the sequence (λ^n) is tight.

Let $\varepsilon > 0$. Then it follows from (3) that there is a positive integer N and a compact subset $\mathcal{A} \subset m(S_1)$ such that

$$\lambda^N(\mathcal{A}) > \sqrt[3]{1-\varepsilon}. \tag{5}$$

By condition (a) every matrix in S_1 has the form as in (1), that is, it looks like

$$\begin{pmatrix} A & BD \\ DC & D \end{pmatrix}, \tag{6}$$

where D is an element of the compact group $\mathcal{G}$ of $a \times a$ matrices. Note that the matrix D itself has rank a. Thus, if the matrix in (6) belongs to $m(S_1)$, then it has rank a and, therefore, it must look like

$$\begin{pmatrix} BDC & BD \\ DC & D \end{pmatrix}, \tag{7}$$

where $D \in \mathcal{G}$. In fact, every matrix in $m(S_1)$ is of the form (7). Now we claim that the set $\mathcal{A} \cdot m(S_1) \cdot \mathcal{A} \equiv \{x_1 \cdot x_2 \cdot x_3\colon x_1$ and x_3 are in $\mathcal{A}$ and $x_2 \in m(S_1)\}$ is a compact subset of S_1. To prove this claim, let us consider

$$\begin{pmatrix} B_1D_1C_1 & B_1D_1 \\ D_1C_1 & D_1 \end{pmatrix} \in \mathcal{A}, \qquad \begin{pmatrix} B_2D_2C_2 & B_2D_2 \\ D_2C_2 & D_2 \end{pmatrix} \in \mathcal{A}$$

and

$$\begin{pmatrix} BDC & BD \\ DC & D \end{pmatrix} \in m(S_1).$$

Notice that

$$\begin{aligned} &\begin{pmatrix} B_1D_1C_1 & B_1D_1 \\ D_1C_1 & D_1 \end{pmatrix}\begin{pmatrix} BDC & BD \\ DC & D \end{pmatrix}\begin{pmatrix} B_2D_2C_2 & B_2D_2 \\ D_2C_2 & D_2 \end{pmatrix} \\ &= \begin{pmatrix} B_1D_1C_1 & B_1D_1 \\ D_1C_1 & D_1 \end{pmatrix}\begin{pmatrix} BD^*C_2 & BD^* \\ D^*C_2 & D^* \end{pmatrix} \\ &= \begin{pmatrix} B_1D^{**}C_2 & B_1D^{**} \\ D^{**}C_2 & D^{**} \end{pmatrix}, \end{aligned} \tag{8}$$

where $D^* = D(CB_2 + I_a)D_2 \in \mathcal{G}$ and $D^{**} = D_1(C_1B + I_a)D^* \in \mathcal{G}$. Now observe that the sets

$$\left\{B \,\middle|\, \begin{pmatrix} BDC & BD \\ DC & D \end{pmatrix} \in \mathcal{A} \text{ for some } D \in \mathcal{G} \text{ and for some } C\right\}$$

and

$$\left\{C \,\middle|\, \begin{pmatrix} BDC & BD \\ DC & D \end{pmatrix} \in \mathcal{A} \text{ for some } D \in \mathcal{G} \text{ and for some } B\right\}$$

are both compact subsets in the usual topology of $(d - a) \times a$ and $a \times (d - a)$ matrices, respectively, because $\mathcal{A}$ is a compact subset of $d \times d$ matrices. [The reason is the following: For every matrix B in the first set above, the entries in the matrix $B{\cdot}D$ ($\equiv B_1$, say), which is the block in the upper right hand corner of the matrix

$$\begin{pmatrix} BDC & BD \\ DC & D \end{pmatrix},$$

which belongs to $\mathcal{A}$, must be bounded because $\mathcal{A}$ is a compact set. Because the matrix D belongs to the compact group $\mathcal{G}$ of matrices with full rank, it is clear that the entries in all such matrices $B = B_1D^{-1}$ must be bounded. A similar reason applies for the second set above.] This observation, along with the form of the product in (8), implies that the set $\mathcal{A} \cdot m(S_1) \cdot \mathcal{A}$ is a compact subset of S_1. This establishes our claim and consequently it follows from (4) and (5) that

$$\begin{aligned} \lim_{n\to\infty} \lambda^{n+2N}(\mathcal{A} \cdot m(S_1) \cdot \mathcal{A}) &\geq \lim_{n\to\infty} \lambda^N(\mathcal{A}) \cdot \lambda^n(m(S_1)) \cdot \lambda^N(\mathcal{A}) \\ &> 1 - \varepsilon. \end{aligned}$$

It follows that the sequence (λ^n) and, therefore, the sequence (μ^n) are tight.

Let us now prove the "only if" part in (iii). Let us then assume that the sequence (μ^n) is tight. Then it is well known (and as mentioned earlier, see also [11], page 32) that the sequence $(1/n)\sum_{k=1}^{n} \mu^k$ converges weakly to some probability measure ν such that $S(\nu)$, the support of ν, is the (completely simple) minimal ideal of S and, consequently, $S(\nu) = m(S)$. Now we must exploit the algebraic structure of $m(S)$, a completely simple subsemigroup of S. (For details of properties of such semigroups, see [11], pages 6–9.) Let $e = e^2$ be an idempotent element of $m(S)$. [$m(S)$, being a completely simple subsemigroup, has at least one idempotent.] Then e has rank a and there is an invertible $d \times d$ matrix y such that

$$y^{-1}ey = \begin{pmatrix} 0 & 0 \\ 0 & I_a \end{pmatrix}, \tag{9}$$

where I_a is the $a \times a$ identity matrix.

Let $x \in S$. Let us write $y^{-1}xy$ as

$$y^{-1}xy = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \tag{10}$$

where D is $a \times a$, B is $(d-a) \times a$, A is $(d-a) \times (d-a)$ and C is $a \times (d-a)$. Now because $m(S)$ is completely simple and the support of an idempotent probability measure ν, the set $eSe = em(S)e$ is a compact group; see Theorem 3.16 and the proof of Theorem 3.15 in [11]. Observe that

$$y^{-1}(exe)y = \begin{pmatrix} 0 & 0 \\ 0 & I_a \end{pmatrix}\begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & I_a \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & D \end{pmatrix}.$$

Let $\mathcal{G} = \{D \mid \text{there exist } A, B, C \text{ as in (10) such that } (\begin{smallmatrix} A & C \\ B & D \end{smallmatrix}) \in y^{-1}Sy\}$. Then it is clear that $\mathcal{G}$ is a compact group of $a \times a$ matrices. Now we recall from [11] (Theorem 2.14, page 9) that

$$m(S) = E\big(m(S)e\big) \cdot \big[em(S)e\big] \cdot E\big(em(S)\big), \tag{11}$$

where $E(W)$ denotes the set of all idempotent elements in the set W. A typical element in $y^{-1}[m(S)e]y$ is of the form

$$y^{-1}(xe)y = \big(y^{-1}xy\big)\big(y^{-1}ey\big) = \begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & I_a \end{pmatrix} = \begin{pmatrix} 0 & B \\ 0 & D \end{pmatrix}$$

and if this element is idempotent, then

$$\begin{pmatrix} 0 & B \\ 0 & D \end{pmatrix} = \begin{pmatrix} 0 & B \\ 0 & D \end{pmatrix}\begin{pmatrix} 0 & B \\ 0 & D \end{pmatrix} = \begin{pmatrix} 0 & BD \\ 0 & D^2 \end{pmatrix}$$

so that D must be I_a. Thus, elements in $y^{-1} \cdot E(m(S)e)y$ are of the form $(\begin{smallmatrix} 0 & B \\ 0 & I_a \end{smallmatrix})$. Similarly, the elements in $y^{-1}E(em(S))y$ are of the form $(\begin{smallmatrix} 0 & 0 \\ C & I_a \end{smallmatrix})$. Then it follows from (11) that a typical element in $y^{-1}m(S)y$ is of the form

$$\begin{pmatrix} 0 & B \\ 0 & I_a \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & D \end{pmatrix}\begin{pmatrix} 0 & 0 \\ C & I_a \end{pmatrix} = \begin{pmatrix} BDC & BD \\ DC & D \end{pmatrix}, \tag{12}$$

where D belongs to the compact group $\mathcal{G}$. Now let us consider an arbitrary element in $y^{-1}Sy$. Let it be $y^{-1}zy$, which looks like

$$\begin{pmatrix} A_1 & B_1 \\ C_1 & D_1 \end{pmatrix}, \qquad D_1 \in \mathcal{G}, \tag{13}$$

expressed in the form described in (10). Then notice that $y^{-1}(ze)y \in y^{-1}m(S)y$ and, therefore,

$$\begin{pmatrix} A_1 & B_1 \\ C_1 & D_1 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & I_a \end{pmatrix} = \begin{pmatrix} 0 & B_1 \\ 0 & D_1 \end{pmatrix}$$

must have the same form as in (12) so that $B_1 = BD_1$ for some $(d-a)\times a$ matrix B. Notice that this B is unique when B_1 and D_1 are given because $D_1 \in \mathcal{G}$ and $\mathcal{G}$ is a group. Similarly, $y^{-1}(ez)y \in y^{-1}m(S)y$ and, therefore,

$$\begin{pmatrix} 0 & 0 \\ 0 & I_a \end{pmatrix}\begin{pmatrix} A_1 & B_1 \\ C_1 & D_1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ C_1 & D_1 \end{pmatrix}$$

must have the form as in (12) and as such, $C_1 = D_1 C$ for some unique $a \times (d-a)$ matrix. Thus, it follows from (13) that every matrix in $y^{-1}Sy$ is of the form

$$\begin{pmatrix} A_1 & BD_1 \\ D_1C & D_1 \end{pmatrix}, \tag{14}$$

where D_1 is an element of the compact group $\mathcal{G}$ of $a \times a$ matrices and (14) has the same form described in (10). The proof of (iii) will be complete once we show that whenever the matrix

$$\begin{pmatrix} A & BD \\ DC & D \end{pmatrix},$$

with the same form as in (14), is in $y^{-1}Sy$, then the matrix

$$\begin{pmatrix} BDC & BD \\ DC & D \end{pmatrix}$$

must also be an element of $y^{-1}Sy$ and, therefore, in $y^{-1}m(S)y$.

To prove this part, let us take x in S such that

$$y^{-1}xy = \begin{pmatrix} A & BD \\ DC & D \end{pmatrix}$$

with the same form as in (14). Notice that $eSe = em(S)e$ is a group and as such, $(exe)^{-1}$, the inverse of exe in this group, belong to S. If $z \equiv y^{-1}(x(exe)^{-1}x)y$, then

$$z = \begin{pmatrix} A & BD \\ DC & D \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & D^{-1} \end{pmatrix}\begin{pmatrix} A & BD \\ DC & D \end{pmatrix} = \begin{pmatrix} BDC & BD \\ DC & D \end{pmatrix}.$$

Thus, we have proven that the set $m(S)$ consisting of all matrices in S with the minimal rank coincides with the set $y\mathcal{M}y^{-1}$. Because $(1/n)\sum_{k=1}^{n}\mu^k$ converges weakly to the probability measure ν and the support of ν is $m(S)$, the condition (b) now follows immediately from Theorem 4.3 in [7]. [The idea here is that all the weak*-limit points of the sequence (μ^n), which are also weak limit points in this case, must have their supports contained inside $S(\nu)$.] This proves our previous assertion and the proof of (iii) is complete. The proof of the last part (that is, the weak convergence part) is immediate from Theorem 2.1 in [6]. □

Next we consider the case of 2×2 real matrices. In this case, as we show in the following text, we can be more specific.

THEOREM 2. *Let μ be a (Borel) probability measure on 2×2 real matrices and let S be the closed (w.r.t. the usual topology) multiplicative semigroup generated by $S(\mu)$, the support of μ. Suppose that the sequence (μ^n) is tight and that the rank of the matrices in $m(S)$ is 1. (When this rank is 0 or 2, exactly what happens is clear from Theorem 1.)*

Suppose also that $m(S)$ does not contain a group of the form $\{1, -1\}$. Then either there is a common right nonzero eigenvector for every matrix in S with common eigenvalue 1 or there is a common left nonzero eigenvector for every matrix in S with common eigenvalue 1. In particular, there is an invertible 2×2 matrix y such that in case of the first possibility, $y^{-1}Sy \subset \{(\begin{smallmatrix} a & 0 \\ c & 1 \end{smallmatrix})$: a, c scalars$\}$ and in case of the second possibility, $y^{-1}Sy \subset \{(\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix})$: a, b scalars$\}$.

Suppose now that $m(S)$ contains a group of the form $\{1, -1\}$. Then either $m(S)$ consists of exactly eight elements, or there is a common right eigenvector for every matrix in S with eigenvalue 1 or -1, or there is a common left eigenvector for every matrix in S with eigenvalue 1 or -1. In case of the last two possibilities, there is an invertible 2×2 matrix y such that in case of the second possibility,

$$y^{-1}Sy \subset \left\{ \begin{pmatrix} a & 0 \\ c & b \end{pmatrix} : a, c \text{ scalars and } b = \pm 1 \right\},$$

and in case of the third possibility,

$$y^{-1}Sy \subset \left\{ \begin{pmatrix} a & c \\ 0 & b \end{pmatrix} : a, c \text{ scalars and } b = \pm 1 \right\}.$$

(It is relevant to point out that a similar result in the context of nonnegative matrices appeared in [5].)

PROOF. Suppose that (μ^n) is tight and that the rank of the matrices in $m(S)$ is one. Then as we have seen in the proof of Theorem 1, $m(S)$ is a completely simple subsemigroup of S with a compact group factor. Let $e = e^2$ be a fixed idempotent element of $m(S)$. Because rank $(e) = 1$, there is an invertible 2×2 matrix y such that

$$y^{-1}ey = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Let us consider the set $y^{-1}(m(S)e)y$. Because the rank of the matrices in the set $y^{-1}m(S)y$ is 1, a typical element in the set $y^{-1}(m(S)e)y$ is of the form

$$\begin{pmatrix} a_1b_1 & a_1b_2 \\ a_2b_1 & a_2b_2 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & a_1b_2 \\ 0 & a_2b_2 \end{pmatrix}$$

and if this element is idempotent, then $a_2b_2 = 1$ so that the set $E(y^{-1}m(S)ey) = y^{-1}(E(m(S)e))y$, where $E(W)$ is the set of idempotent elements in W, consists of elements of the form $\left(\begin{smallmatrix}0 & a\\ 0 & 1\end{smallmatrix}\right)$. Similarly, the set $E(y^{-1}em(S)y) = y^{-1}E(em(S))y$ consists of elements of the form $\left(\begin{smallmatrix}0 & 0\\ b & 1\end{smallmatrix}\right)$. Now suppose that

$$E\big(m(S)e\big) \neq e \quad \text{and} \quad E\big(em(S)\big) \neq e. \tag{15}$$

Then there exist elements x_1, x_2 in $m(S)$ such that $x_1e = (x_1e)^2$, $ex_2 = (ex_2)^2$ and

$$y^{-1}(x_1e)y = \begin{pmatrix}0 & a\\ 0 & 1\end{pmatrix}, \qquad a \neq 0,$$

and

$$y^{-1}(ex_2)y = \begin{pmatrix}0 & 0\\ b & 1\end{pmatrix}, \qquad b \neq 0.$$

Notice that then

$$y^{-1}(ex_2x_1e)y = \begin{pmatrix}0 & 0\\ b & 1\end{pmatrix}\begin{pmatrix}0 & a\\ 0 & 1\end{pmatrix} = \begin{pmatrix}0 & 0\\ 0 & ab+1\end{pmatrix}, \tag{16}$$

where $ab + 1 \neq 1$ because $ab \neq 0$.

However, because $y^{-1}(em(S)e)y$ is a compact group of matrices of the form $\left(\begin{smallmatrix}0 & 0\\ 0 & c\end{smallmatrix}\right)$, c being a nonzero scalar, it is clear that in (16), $ab + 1 \in \{1, -1\}$. Because $ab \neq 0$, $ab + 1 = -1$ or $ab = -2$.

It follows that if (15) does occur, then we must have

$$E\Big[y^{-1}\big(em(S)\big)y\Big] = \left\{\begin{pmatrix}0 & 0\\ -\frac{2}{a} & 1\end{pmatrix}, \begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}\right\} \tag{17}$$

and

$$E\Big[y^{-1}\big(m(S)e\big)y\Big] = \left\{\begin{pmatrix}0 & a\\ 0 & 1\end{pmatrix}, \begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}\right\}$$

for some $a \neq 0$. It is also clear that when $m(S)$ does not contain a group of the form $\{1, -1\}$, the element $ab + 1$ must be 1 so that $ab = 0$, which will mean that (15) cannot occur in this case.

Let us now observe that the set $y^{-1}(em(S)e)y$, being a compact group, must be either $\{\left(\begin{smallmatrix}0 & 0\\ 0 & 1\end{smallmatrix}\right)\}$ or $\{\left(\begin{smallmatrix}0 & 0\\ 0 & 1\end{smallmatrix}\right), \left(\begin{smallmatrix}0 & 0\\ 0 & -1\end{smallmatrix}\right)\}$. As a consequence, because $m(S)$ is completely simple, it follows easily that every element in $m(S)$ is idempotent when $y^{-1}(em(S)e)y$ is a singleton and that when this set is not a singleton, then $m(S) = -m(S)$ so that every element in $m(S)$ is either an idempotent or the negative of an idempotent.

Because we have the set equality $m(S) = E(m(S)e) \cdot em(S)e \cdot E(em(S))$, it follows that when (17) occurs, $y^{-1}m(S)y$ consists of exactly the following eight elements, that is, there is $a \neq 0$ such that

$$\begin{aligned} y^{-1}m(S)y = \Bigg\{&\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}, \begin{pmatrix}0 & 0\\ 0 & -1\end{pmatrix}, \begin{pmatrix}0 & a\\ 0 & 1\end{pmatrix}, \begin{pmatrix}0 & -a\\ 0 & -1\end{pmatrix},\\ &\begin{pmatrix}0 & 0\\ -\frac{2}{a} & 1\end{pmatrix}, \begin{pmatrix}0 & 0\\ \frac{2}{a} & -1\end{pmatrix}, \begin{pmatrix}2 & -a\\ \frac{2}{a} & -1\end{pmatrix}, \begin{pmatrix}-2 & a\\ -\frac{2}{a} & 1\end{pmatrix}\Bigg\}. \end{aligned} \tag{18}$$

Now we look into the situation when $m(S)$ does not contain a group of the form $\{1, -1\}$. In this case, (15) does not occur and then we have either

$$Se = m(S)e = E\big(m(S)e\big) = \{e\}, \tag{19}$$

or

$$eS = em(S) = E\big(em(S)\big) = \{e\}. \tag{20}$$

Suppose that (19) holds. Then, for any $s \in S$, let $y^{-1}sy = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. Then $y^{-1}(se)y = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, in which case $b = 0$ and $d = 1$. This means that when (19) occurs,

$$y^{-1}Sy \subset \left\{ \begin{pmatrix} a & 0 \\ c & 1 \end{pmatrix} : a, c \text{ scalars} \right\}.$$

Similarly, when (20) occurs,

$$y^{-1}Sy \subset \left\{ \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} : a, b \text{ scalars} \right\}.$$

Finally, it can so happen that $m(S)$ contains a group of the form $\{1, -1\}$ and also (15) does not occur. In this case, as is clear from above, we must have either

$$Se = m(S) \cdot e = \{e, -e\}$$

or

$$eS = e \cdot m(S) = \{e, -e\}.$$

The rest of the details are clear and left to the reader.

The proof is complete. □

Let us remark that when $\left(\begin{smallmatrix} \xi_i & \eta_i \\ 0 & \varsigma_i \end{smallmatrix}\right)$, $1 \le i < \infty$, is a sequence of i.i.d. random matrices, where $\xi_i, \eta_i, \varsigma_i$ are real random variables with $\varsigma_i = \pm 1$, then it is easy to prove using methods in [2] that if $P(\xi_i = 0) = 0$, $E \log|\xi_i| < 0$ and $E \log \max\{|\eta_i|, 1\} < \infty$, then the sequence (μ^n) is tight, where μ is the distribution of $\left(\begin{smallmatrix} \xi_i & \eta_i \\ 0 & \varsigma_i \end{smallmatrix}\right)$. In the same manner one can treat an i.i.d. sequence $\left(\begin{smallmatrix} \xi_i & 0 \\ \eta_i & \varsigma_i \end{smallmatrix}\right)$ with ξ_i, η_i and ς_i as before, the backward product in this case corresponding to the forward product in the former case.

Let us now illustrate Theorem 1 with a simple example. We consider a probability measure μ with a three-point support in 3×3 real matrices so that

$$S(\mu) = \{x_1, x_2, x_3\},$$

where

$$x_1 = \begin{pmatrix} 1 & \cos\theta_1 - \sin\theta_1 & -\sin\theta_1 - \cos\theta_1 \\ 0 & \cos\theta_1 & -\sin\theta_1 \\ 0 & \sin\theta_1 & \cos\theta_1 \end{pmatrix}, \qquad x_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and

$$x_3 = \begin{pmatrix} 2 & \cos\theta_2 - \sin\theta_2 & -\sin\theta_2 - \cos\theta_2 \\ \frac{1}{\sqrt{2}}(\cos\theta_2 + \sin\theta_2) & \cos\theta_2 & -\sin\theta_2 \\ \frac{1}{\sqrt{2}}(\sin\theta_2 - \cos\theta_2) & \sin\theta_2 & \cos\theta_2 \end{pmatrix},$$

where $\theta_1, \theta_2 \geq 0$. Notice that to keep everything simple, the support of μ has been taken here as in the proof of Theorem 1 as $y^{-1}S(\mu)y$ (for some appropriate invertible y). Also observe that

$$x_1 = \begin{pmatrix} A_1 & B_1D_1 \\ D_1C_1 & D_1 \end{pmatrix}, \qquad x_2 = \begin{pmatrix} 0 & 0 \\ 0 & I_2 \end{pmatrix}, \qquad x_3 = \begin{pmatrix} A_2 & B_2D_2 \\ D_2C_2 & D_2 \end{pmatrix},$$

where

$$D_1 = \begin{pmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{pmatrix}, \qquad D_2 = \begin{pmatrix} \cos\theta_2 & -\sin\theta_2 \\ \sin\theta_2 & \cos\theta_2 \end{pmatrix},$$

$$C_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad B_1 = (1, -1) = B_2, \qquad C_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}.$$

Here, $C_1B_1 + I_2, C_1B_2 + I_2, C_2B_1 + I_2$ and $C_2B_2 + I_2$ are all of the form $\left(\begin{smallmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{smallmatrix}\right)$. Thus, the closed semigroup generated by $S(\mu)$ will consist of elements of the form $\left(\begin{smallmatrix} A & BD \\ DC & D \end{smallmatrix}\right)$, as described in Theorem 1, where D is an element of the compact group

$$\left\{ \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} : \theta \geq 0 \right\}.$$

It is clear that the rank of the matrices in $m(S)$ is 2 and that $x_2 \in m(S)$ so that $\mu(m(S)) > 0$, which implies that $\lim_{n\to\infty} \mu^n(m(S)) = 1$. It follows from Theorem 1 that the sequence (μ^n) is tight. Because $S(\mu)$ contains an idempotent element in $m(S)$, it also follows from Theorem 1 that the sequence (μ^n) converges weakly. The same will be the conclusion if we considered $S(\mu)$ to be a set $\{y^{-1}x_1y, y^{-1}x_2y, y^{-1}x_3y\}$ for any invertible 3×3 matrix y. Notice that without the use of Theorem 1 (and this the reader can verify for herself/himself), it may not be a simple problem to establish the weak convergence of the sequence (μ^n) even in this very simple situation.

Acknowledgment. We are grateful to an anonymous referee for his very constructive comments, which helped us revise, clarify and correct a number of points in an earlier version.

REFERENCES

[1] BOUGEROL, P. (1987). Tightness of products of random matrices and stability of linear stochastic systems. *Ann. Probab.* **15** 40–74 .

[2] GRINCEVICIUS, A. K. (1974). On the continuity of the distribution of a sum of dependent variables connected with independent walks on lines. *Theory. Probab. Appl.* **19** 163–168.

[3] HÖGNAS, G. (1987). A note on products of random matrices. *Statist. Probab. Lett.* **5** 367–370.

[4] HÖGNAS, G. and MUKHERJEA, A. (1980). Recurrent random walks and invariant measures on semigroups of $n \times n$ matrices. *Math. Z.* **173** 69–94.

[5] KESTEN, H. and SPITZER, F. (1984). Convergence in distribution for products of random matrices. Z. Wahrsch. *Verw. Gebiete* **67** 363–386.

[6] LO, C. C. and MUKHERJEA, A. (1991). Convergence in distribution for products of $d \times d$ random matrices. *J. Math. Anal. Appl.* **162**(1): 71–91.

[7] MUKHERJEA, A. (1979). Limit theorems: Stochastic matrices, ergodic Markov chains and measures on semigroups. In *Probability Analysis and Related Topics* **2** 143–203. Academic Press, New York.

[8] MUKHERJEA, A. (1987). Convergence in distribution of products of random matrices: A semigroup approach. *Trans. Amer. Math. Soc.* **303** 395–411.

[9] MUKHERJEA, A. (1991). Tightness of products of i.i.d. random matrices. *Probab. Theory Related Fields* **87** 389–401.

[10] Mukherjea, A. (1992). Recurrent random walks in nonnegative matrices: Attractors of certain iterated function systems. *Probab. Theory Related Fields* **91** 297–306.

[11] MUKHERJEA, A. and TSERPES, N. A. (1976). *Measures on Topological semigroups: Convolution Products and Random Walks. Lecture Notes in Math.* **547**. Springer, Berlin.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF SOUTH FLORIDA
TAMPA, FLORIDA 33620-5700
ARUN@ MATH.USF.EDU