Asymptotic results on the product of random probability matrices

1996 J. Phys. A: Math. Gen. 29 3053

(http://iopscience.iop.org/0305-4470/29/12/013)

J. Phys. A: Math. Gen. **29** (1996) 3053–3061. Printed in the UK

Asymptotic results on the product of random probability matrices

X R Wang

Physics Department, The Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong

Received 9 January 1996

Abstract. I study the product of independent identically distributed $D \times D$ random probability matrices. Some exact asymptotic results are obtained. I find that both the left and the right products approach exponentially a probability matrix (asymptotic matrix) in which any two rows are the same. A parameter λ is introduced for the exponential coefficient which can be used to describe the convergent rate of the products. λ depends on the distribution of the individual random matrices. I find $\lambda = \frac{3}{2}$ for $D = 2$ when each element of the individual random probability matrices is uniformly distributed in $[0, 1]$. In this case, each element of the asymptotic matrix follows a parabolic distribution function. The distribution function of the asymptotic matrix elements can be shown numerically to be non-universal. Numerical tests are carried out for a set of random probability matrices with a particular distribution function. I find that λ increases monotonically from $\simeq$1.5 to $\simeq$3 as D increases from 3 to 99, and the distribution of random elements in the asymptotic products can be described by a Gaussian function with a mean of $1/D$.

In recent years, there has been an increasing interest in studying the properties of random matrices [1]. Random matrices can be used to describe disordered systems, chaos and biological problems, and in the statistical description of complex nuclei [1, 2]. There are two different kinds of problem related to random matrices. One of them is to study the statistical properties of a single random matrix. Wigner [3] was the first to use a single large random matrix to explain the statistical behaviour of levels in nuclear physics. The semicircular and the circular theorems [1, 3] had been obtained for the Gaussian unitary, Gaussian orthonormal and Gaussian symplectic ensembles. The second problem is to study the statistical behaviour of a product of random matrices. The product of random matrices has attracted much attention in most recent works because many statistical problems in disordered systems and chaotic dynamical systems can be formulated as the study of such a product. For the product of random matrices, there are beautiful Furstenberg [4] and Oseledec [5] theorems about the existence of Lyapunov characteristic exponents. However, a detailed analysis of the product of random matrices is very difficult due to the non-commutability of random matrices, and there are not many general exact results about structures of the products of random matrices. Therefore, any exact results on the product of a particular type of random matrices should be interesting.

In this work, I present some exact results of the product of independent identically distributed random probability matrices. A probability matrix T of $D \times D$ is defined as $T(ij) \geqslant 0$ and $\sum_{j=1}^{D} T(ij) = 1$, $i = 1, 2, \ldots, D$, where $T(ij)$ are the elements of T.

0305-4470/96/123053+09$19.50

Physically, such matrices can be used to describe the dynamics of a D-state classical system if $T_t(ij)$ is interpreted as the probability for the system in the ith state jumping to the jth state at time t. If the hopping process is stochastic, the evolution of the system is described by the product of the random probability matrices T_t. It is easy to show that the product of two probability matrices is still a probability matrix, i.e. $A = T_1 T_2$ is a probability matrix provided T_1 and T_2 are two probability matrices of order $D \times D$. In this paper, I show that the product of independent identically distributed random probability matrices approaches exponentially, in terms of the number of matrices in the product, a matrix in which any two rows are the same. For 2×2 independent random probability matrices in which any element is uniformly distributed in $[0, 1]$, I find that the asymptotic matrix elements have a parabolic distribution. The main results are described by the following propositions.

Proposition 1. Let $\{T_k\}$ be a set of $D \times D$ independent identically distributed random probability matrices in which all elements are random and have the same distribution function. If $A(n) = \prod_{k=1}^{n} T_k \equiv T_n T_{n-1} \cdots T_2 T_1$ (the left product) and $B(n) = \prod_{k=1}^{n} T_k \equiv T_1 T_2 \cdots T_{n-1} T_n$ (the right product), then

$$\lim_{n\to\infty} A(n) = \begin{pmatrix} a(1) & a(2) & \dots & a(D) \\ a(1) & a(2) & \dots & a(D) \\ \vdots & \vdots & \ddots & \vdots \\ a(1) & a(2) & \dots & a(D) \end{pmatrix} \tag{1}$$

and

$$\lim_{n\to\infty} B(n) = \begin{pmatrix} a(1) & a(2) & \dots & a(D) \\ a(1) & a(2) & \dots & a(D) \\ \vdots & \vdots & \ddots & \vdots \\ a(1) & a(2) & \dots & a(D) \end{pmatrix} \tag{2}$$

where $a(i)$, $(i = 1, 2, \dots D)$, are positive random numbers with $\sum_i a(i) = 1$. However, the values of a in the left product are fixed for a given sequence of random matrices while they keep changing in the right product.

Before we prove this proposition, let us look at a special case of $D = 2$. Let

$$A(n) \equiv T_n \cdots T_2 T_1 = \begin{pmatrix} y_n(1) & 1 - y_n(1) \\ y_n(2) & 1 - y_n(2) \end{pmatrix} \tag{3}$$

then, from $A(n) = T_n A(n-1)$, we obtain the following recursion relations for $y_n(1)$ and $y_n(2)$

$$y_n(1) = x(1) y_{n-1}(1) + (1 - x(1)) y_{n-1}(2) \tag{4a}$$

$$y_n(2) = x(2) y_{n-1}(1) + (1 - x(2)) y_{n-1}(2) \tag{4b}$$

where $x(1)$ and $x(2)$ are the two independent random elements of T_n, i.e.

$$T_n = \begin{pmatrix} x(1) & 1 - x(1) \\ x(2) & 1 - x(2) \end{pmatrix}.$$

Therefore,

$$y_n(1) - y_n(2) = [x(1) - x(2)][y_{n-1}(1) - y_{n-1}(2)]$$

which gives

$$\left| \frac{y_n(1) - y_n(2)}{y_{n-1}(1) - y_{n-1}(2)} \right| = |x(1) - x(2)| \leqslant 1\,. \tag{5}$$

Thus, we expect that $|y_n(1) - y_n(2)|$ approaches zero exponentially. If $x(1)$ and $x(2)$ are uniformly distributed in [0, 1], then, for a large n,

$$y_n(1) - y_n(2) \sim \mathrm{e}^{-\lambda n} \tag{6}$$

with

$$\lambda = -\langle \ln |x(1) - x(2)| \rangle = -\int_0^1 \int_0^1 \ln |x(1) - x(2)| \, \mathrm{d}x(1) \, \mathrm{d}x(2) = \tfrac{3}{2} \, .$$

Therefore, the left product of independent identically distributed 2×2 random probability matrices exponentially approaches a matrix of the form

$$\begin{pmatrix} a_1 & 1 - a_1 \\ a_2 & 1 - a_2 \end{pmatrix} \tag{7}$$

with $\lambda = 1.5$. Similarly, it is easy to show that the same conclusion can be drawn for a right product.

The above approach can be extended to the general cases. Without losing generality, we need only show that the values of elements in the first column of the product of random probability matrices approach each other. Let

$$A_n \equiv \begin{pmatrix} y_n(1) & & \\ y_n(2) & & \\ \cdots & * & \\ \cdots & & \\ y_n(D) & & \end{pmatrix} = T_n A(n-1) \, . \tag{8}$$

We want to show that $\langle |y_n(k) - y_n(1)| \rangle \simeq \exp^{-\lambda n}$, $\lambda > 0$. It is not hard to see that

$$y_n(i) - y_n(D) = \sum_{k=1}^{D-1} [x(i,k) - x(D,k)][y_{n-1}(k) - y_{n-1}(D)] \tag{9}$$

where $x(i, j)$ are matrix elements of T_n which satisfies the conditions of a probability matrix. $\sum_j x(i,k) = 1$ is used in the above derivation. Define

$$z_n(i) = y_n(i) - y_n(D) \qquad i = 1, \ldots, D-1 \, . \tag{10}$$

Equation (9) can be written in the following matrix form:

$$\begin{aligned} \begin{pmatrix} z_n(1) \\ z_n(2) \\ \vdots \\ z_n(D-1) \end{pmatrix} &= C_n \begin{pmatrix} z_{n-1}(1) \\ z_{n-1}(2) \\ \vdots \\ z_{n-1}(D-1) \end{pmatrix} \\ &= \begin{pmatrix} x(1,1) - x(D,1) & \ldots & x(1,D-1) - x(D,D-1) \\ x(2,1) - x(D,1) & \ldots & x(2,D-1) - x(D,D-1) \\ \vdots & \ddots & \vdots \\ x(D-1,1) - x(D,1) & \ldots & x(D-1,D-1) - x(D,D-1) \end{pmatrix} \\ &\quad \times \begin{pmatrix} z_{n-1}(1) \\ z_{n-1}(2) \\ \vdots \\ z_{n-1}(D-1) \end{pmatrix} \end{aligned} \tag{11}$$

where C_n is a $(D-1)\times(D-1)$ matrix related to probability matrix T_n. It is not difficult to show the following relation between T_n and C_n:

$$\|T_n - \mu\mathbf{I}\| = (1-\mu)\|C_n - \mu\mathbf{I}\| \tag{12}$$

that is, the eigenvalues of matrix C_n are $D-1$ of the eigenvalues of matrix T_n (one of the eigenvalues of T_n with value 1 is excluded)†. It is well known that the magnitudes of eigenvalues $\mu_i,\ i = 1,\ldots,D$, of a $D\times D$ probability matrix are not greater than 1, i.e. $|\mu_i| \leqslant 1$. Furthermore, matrices C do not have any common eigenvectors‡. Therefore,

$$\left\langle \begin{pmatrix} z_n(1) \\ z_n(2) \\ \vdots \\ z_n(1) \end{pmatrix} \right\rangle \sim \mathrm{e}^{-\lambda N} \qquad \text{(for a large } N\text{)} \tag{13}$$

with $\lambda > 0$, i.e. the values of elements in the first column of the left product A approach each other exponentially. It is easy to show that the same result is true for any other column of A. Thus, relation (1) holds. Similarly, it can be shown that relation (2) also holds.

This result is not really surprising. It is known that the magnitudes of eigenvalues of a probability matrix are equal to or smaller than 1. A probability matrix always has an eigenvalue 1 with the corresponding eigenvector (mode)

$$\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

(there may exist other eigenvectors with eigenvalue 1). The eigenvalues of the product will either approach 0 or stay at 1 when such matrices are multiplied together. Because each matrix is random and independent, these matrices are not commutable among themselves, and they do not in general have the same eigenvector except

$$\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

which will remain unchanged since its eigenvalue is equal to 1. Therefore, all other modes are mixed together, and decay with the multiplication. The relation (1) is then expected.

Proposition 2. Let $\{T_k\}$ be a set of 2×2 independent identically distributed random probability matrices in which all elements are uniformly distributed in $[0, 1]$. Proposition 1 guarantees that

$$\lim_{n\to\infty} A(n) \equiv \lim_{n\to\infty} T_n T_{n-1}\cdots T_2 T_1 = \begin{pmatrix} a & 1-a \\ a & 1-a \end{pmatrix}$$

† First add $D-1$ columns to the last column of $\|T_n - \mu\mathbf{I}\|$. All elements of the last column are equal to $1-\mu$ because $\sum_j x(i,j) = 1$. Take the common factor $1-\mu$ out, and subtract the last row from the first $D-1$ rows, then the top-left $(D-1)\times(D-1)$ block is exactly $\|C_n - \mu\mathbf{I}\|$.

‡ A probability matrix T_i always has an eigenstate

$$\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

with eigenvalue 1. The eigenvalues and eigenstates are given by matrix C_n which depends on the detailed structure of T_i. In general, C_n do not have common eigenstates.

and

$$\lim_{n\to\infty} B(n) \equiv \lim_{n\to\infty} T_1 T_2 \cdots T_{n-1} T_n = \begin{pmatrix} a & 1-a \\ a & 1-a \end{pmatrix}.$$

Then a is a random variable whose distribution function is

$$f(a) = 6a(1-a)\,. \tag{14}$$

To prove proposition 2, we notice that a is a random variable which obeys the recursion relation

$$a_n = x(1)a_{n-1} + x(2)(1 - a_{n-1}) \tag{15}$$

where $x(1)$ and $x(2)$ are two independent random numbers uniformly distributed in $[0, 1]$. Since $f(a)$ is the asymptotic distribution function of a_n, i.e. $n \to \infty$, a_n and a_{n-1} should have the same distribution function $f(a)$ when $n \to \infty$. Therefore, we can obtain the following equation in the integral form for distribution function $f(a)$

$$f(a) = \int_0^1 \int_0^1 \int_0^1 f(b)\delta(a - bx_1 + bx_2 - x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2\,\mathrm{d}b\,. \tag{16}$$

Substituting $\delta(a - bx_1 + bx_2 - x_2)$ by

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}q(a-bx_1+bx_2-x_2)}\,\mathrm{d}q \tag{17}$$

and integrating over x_1 and x_2, equation (16) becomes

$$f(a) = \frac{1}{2\pi}\int_0^1 \mathrm{d}b\,\frac{f(b)}{b(b-1)}\int_{-\infty}^{\infty}\mathrm{d}q\,\frac{\mathrm{e}^{\mathrm{i}qa} + \mathrm{e}^{\mathrm{i}q(a-1)} - \mathrm{e}^{\mathrm{i}q(a-b)} - \mathrm{e}^{\mathrm{i}q(a+b-1)}}{q^2}\,. \tag{18}$$

Differentiating equation (18) with respect to a twice and noting that $f(a) = f(1-a)$, we can show that $f(a)$ satisfies the differential equation

$$f'' = -\frac{2}{a(1-a)}f \tag{19}$$

with boundary conditions $f(0) = f(1) = 0$. Equation (19) can easily be solved by the power-series expansion method since $a = 0$ (or $a = 1$) is a regular singular point. The solution of this equation (normalized to 1) is that of equation (14).

I have carried out some numerical simulations to further confirm the results in the above propositions. Figure 1 is the distribution function of an element of the asymptotic matrix of the product of 2×2 independent random probability matrices in which elements are uniformly distributed in $[0, 1]$. The full curve is the numerical result, and the broken curve is the analytical expression (14). They agree very well with each other. In order to check whether the distribution function of the product depends on the distribution function of individual random probability matrices, I also study the product of independent random probability matrices in which elements are distributed in $[0, 1]$ according to function

$$f(x) = \int_0^1 \ldots \int_0^1 \delta\left(x - \frac{x_1}{\sum_1^D x_i}\right)\mathrm{d}x_1 \ldots \mathrm{d}x_D\,. \tag{20}$$

Distribution (20) is chosen because it is easy to generate on a computer. Although I cannot find the distribution function for the matrix elements of the asymptotic product analytically in this case, numerical results can easily be obtained. Figure 2 shows the single-variable distribution functions of random matrix elements in the right product of such random probability matrices of the order of 2×2 and 4×4. The broken curves are the numerical results and the full curves are the fits of Gaussian functions. The numerical

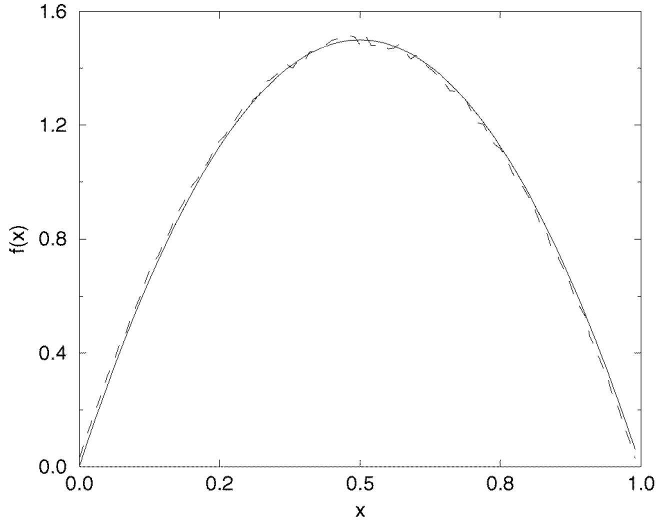

Figure 1. Distribution function of a random matrix element in the right product of independent identically distributed 2×2 random probability matrices in which all elements are uniformly distributed in $[0, 1]$. The broken curve is the numerical result and the full curve is $f(a) = 6a(1-a)$.

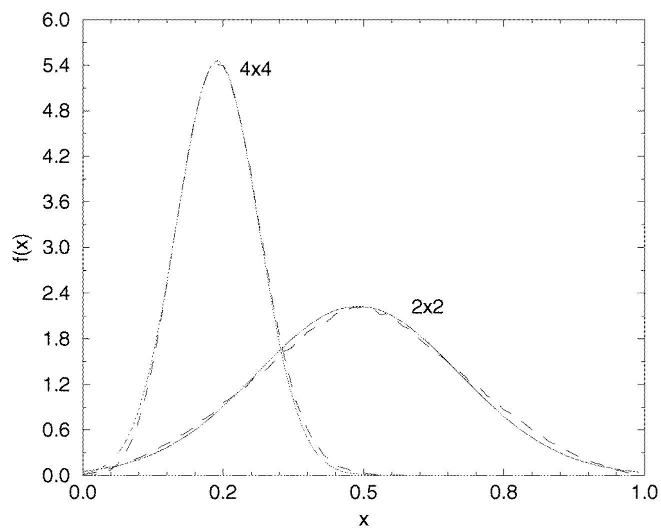

Figure 2. Distribution function of an arbitrary random matrix element in the right product of independent identically distributed 2×2, 4×4 random probability matrices in which all elements are distributed in $[0, 1]$ according to equation (20). The broken curves are the numerical results and the full curves are the fits of Gaussian functions.

results can be well described by a Gaussian function with its mean equal to $1/D$. Compared with that in figure 1, we can see that the distribution function of the product of independent identically distributed random probability matrices depends on the distribution of individual random matrices. In other words, unlike the large number theorem for random numbers, the distribution function of the product is not universal.

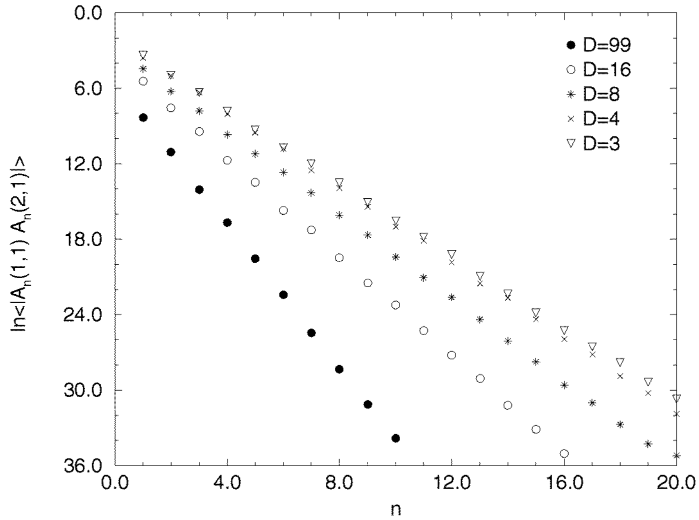

Figure 3. $\ln\langle|A_n(1,1) - A_n(2,1)|\rangle$ versus n of the *left* product of independent identically distributed random matrices of 3×3, 4×4, 8×8, 16×16, 99×99, with 100 ensembles. λ's can be obtained from the slopes. The slopes increase monotonically from $\lambda \simeq 1.5$ to $\lambda \simeq 3$ as D changes from $D=3$ to $D=99$.

The results in proposition 1 are checked numerically by using the random probability matrices whose elements are independently, except for the constraints of a probability matrix, distributed in [0, 1] according to equation (20). I compute the decay of $\langle|A_n(1,1) - A_n(2,1)|\rangle$ with n, where $\langle\cdots\rangle$ denotes ensemble average, and $A_n(i,j)$ are the elements of the product of n random matrices. Figure 3 shows $\ln\langle|A_n(1,1) - A_n(2,1)|\rangle$ versus $\ln n$ for the left product of 3×3, 4×4, 8×8, 16×16 and 99×99 random matrices. 100 ensembles are used in the numerical study. Figure 4 is a similar plot (to figure 1) for a right product. The exponential decay of the quantity is clearly shown in these figures. Numerically, I find that λ increases monotonically from $\simeq$1.5 to $\simeq$3 as D increases from 3 to 99.

In conclusion, I have shown that both left and right products of a sequence of independent identically distributed random probability matrices exponentially approach a probability matrix in which all elements in any column vector are the same. An exponential exponent is used to describe this approach rate. I also find that λ increases monotonically from $\simeq$ 1.5 to $\simeq$ 3 as D increases from 3 to 99 when the distribution function of individual random matrices is described by equation (20). I also find that $\lambda = \frac{3}{2}$ for $D = 2$ when random matrix elements are uniformly distributed in [0, 1]. It is well known that at least one of the eigenvalues of a probability matrix is equal to 1 while the rest of them are distributed in [0, 1] [6]. A large D means that the product has more channel to decay to the stable structure (1) or (2). Thus it is expected that the decay is faster for large D, i.e. λ increases with D. In order to understand the meaning of the results, let us look at a physical model system of D states. Assume the system can move randomly from the ith state to the jth state with probability $T_t(ij)$ at time step t. If the system starts from an initial distribution, one might want to know the probability of the system in the i' state, i.e. the distribution function after a long time. The question may be whether there is a stable distribution (equilibrium state), and/or what it is if there is one. Obviously, the long-time distribution(s) are the non-

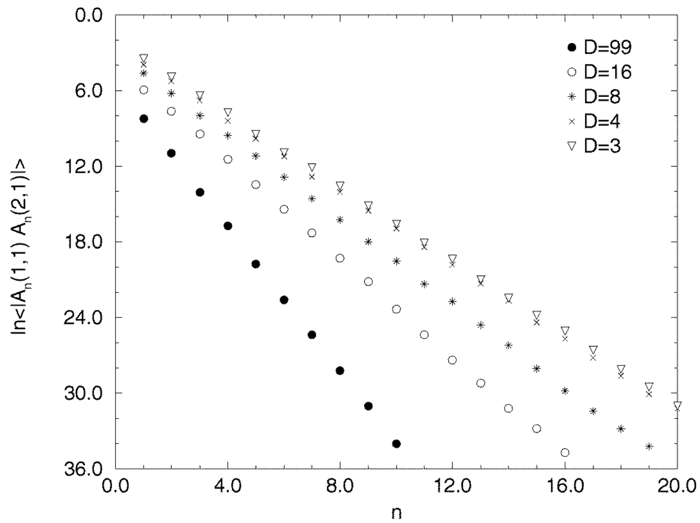

Figure 4. $\ln\langle|A_n(1,1) - A_n(2,1)|\rangle$ versus n of the *right* product of independent identically distributed random matrices of 3×3, 4×4, 8×8, 16×16, 99×99. λ's can be obtained from the slopes. The slopes increase monotonically from $\lambda \simeq 1.5$ to $\lambda \simeq 3$ as D changes from $D = 3$ to $D = 99$.

trivial left eigenstate(s) of the right product of the random matrices T_t. In equation (2), although the structure for a product will not change when n is larger than a certain value, elements $a(i)$ do change as another independent random probability matrix is multiplied by the product. Therefore, the system does not have a stable distribution as expected since the dynamics of the system is a stochastic process, and transition probabilities keep changing with time. Both the right and left products, however, have a non-trivial unique right eigenstate with eigenvalue 1. The eigenstate takes the same value in each of its D components. Unfortunately, I am not able to obtain any meaningful non-trivial results by applying the propositions to this simple model system. It will be interesting to find some interesting physical systems in which the propositions can be used to extract useful information. In contrast to the sum of independent random variables whose distribution is Gaussian no matter what the distribution function of individual random variables is, the distribution function of an element in the product of independent identically distributed random matrices depends on the distribution of individual random matrices. Therefore, the distribution function of the product of independent identically distributed random matrices is not universal.

Acknowledgments

The author would like to point out that proposition 2 results from fruitful discussions with Professor B Derrida. Professor Derrida made the major contribution to the result. The author thanks Professor Derrida for useful suggestions and for reading the manuscript. Stimulating discussions with Professor Y Shapir, Dr W K Ge and Professor E Domany are also acknowledged. This work was supported by UPGC, Hong Kong, through an RGC grant.

References

[1] Mehta M L 1967 *Random Matrices and the Statistical Theory of Energy Levels* (New York: Academic)
Crisanti A, Paladin G and Vulpiani A 1993 *Products of Random Matrices* (New York: Springer)
[2] Cohen J E, Kesten H and Newman C M (eds) *Contemporary Mathematics* vol 50
[3] Wigner E P 1951 *Ann. Math.* **53** 36; 1958 *Ann. Math.* **67** 325
Dyson F J and Mehta M L 1963 *J. Math. Phys.* **4** 701
[4] Furstenberg H and Kesten H 1960 *Ann. Math. Stat.* **31** 457
[5] Oseledec V I 1968 *Trans. Moscow Math. Soc.* **19** 197
[6] See e.g. Feller W 1967 *An Introduction to Probability Theory and its Applications* (New York: Wiley)