*The Annals of Probability*
1997, Vol. 25, No. 4, 1545–1587

# LIMIT THEOREMS FOR PRODUCTS OF POSITIVE RANDOM MATRICES

BY H. HENNION

*Université de Rennes I*

Let $S$ be the set of $q \times q$ matrices with positive entries, such that each column and each row contains a strictly positive element, and denote by $S^\circ$ the subset of these matrices, all entries of which are strictly positive. Consider a random ergodic sequence $(X_n)_{n\geq 1}$ in $S$. The aim of this paper is to describe the asymptotic behavior of the random products $X^{(n)} = X_n \cdots X_1$, $n \geq 1$, under the main hypothesis $P(\bigcup_{n\geq 1}[X^{(n)} \in S^\circ]) > 0$. We first study the behavior "in direction" of row and column vectors of $X^{(n)}$. Then, adding a moment condition, we prove a law of large numbers for the entries and lengths of these vectors and also for the spectral radius of $X^{(n)}$. Under the mixing hypotheses that are usual in the case of sums of real random variables, we get a central limit theorem for the previous quantities. The variance of the Gaussian limit law is strictly positive except when $(X^{(n)})_{n\geq 1}$ is tight. This tightness property is fully studied when the $X_n$, $n \geq 1$, are independent.

## 1. Statement of results.

1.1. *Framework.* Let $S$ be the multiplicative semigroup of $q \times q$ matrices with real positive (greater than or equal to zero) entries which are allowable [30]; that is, every row and every column contains a strictly positive element. The product of $g$ and $g' \in S$ is denoted by $gg'$. The subset of $S$ composed of matrices with strictly positive entries is a subsemigroup denoted by $S^\circ$.

Both $S$ and $S^\circ$ may be described in terms of endomorphisms. Consider the linear space $\mathbb{R}^q$ endowed with its canonical basis $(e_i)_{i=1,\ldots,q}$, with the scalar product $\langle \cdot, \cdot \rangle$ for which this basis is orthonormal and with the norm defined by

$$x \in \mathbb{R}^q, \qquad ||x|| = \sum_{i=1}^{q} |\langle x, e_i \rangle|.$$

Moreover, introduce the cones $C$ and $\overline{C}$:

$$C = \{x: x \in \mathbb{R}^q,\ \forall i = 1, \ldots, q,\ \langle x, e_i \rangle > 0\},$$
$$\overline{C} = \{x: x \in \mathbb{R}^q,\ \forall i = 1, \ldots, q,\ \langle x, e_i \rangle \geq 0\}.$$

In this context, a $q \times q$ matrix with positive entries is identified with an endomorphism of $\mathbb{R}^q$ that preserves $\overline{C}$. Let $g$ be such a matrix. The $g$ image

Received April 1996; revised December 1996.
*AMS* 1991 *subject classifications*. 60F99, 60F05.
*Key words and phrases*. Positive random matrices, mixing, limit theorems.

of $x \in \mathbb{R}^q$ is denoted by $gx$. We associate with $g$ the real numbers

$$||g|| = \sup\{||gx||: x \in \mathbb{R}^q, ||x|| = 1\} = \sup\{||gx||: x \in \overline{C}, ||x|| = 1\},$$

$$v(g) = \inf\{||gx||: x \in \overline{C}, ||x|| = 1\}.$$

Note that $g$ belongs to $S$ if and only if $v(g) > 0$ and $v(g^*) > 0$, where $g^*$ is the adjoint matrix of $g$.

Now, let $(\Omega, \mathscr{F}, P, \theta)$ be an ergodic dynamical system and let $X_0$ be a random variable on $\Omega$ with values in $S$.

For $n \geq 0$, we set

$$X_n = X_0 \circ \theta^n.$$

Then we define the random products $X^{(n)}$, $n \geq 1$, by

$$X^{(1)}(\omega) = X_1(\omega), \qquad X^{(n+1)}(\omega) = X_{n+1}(\omega) X^{(n)}(\omega).$$

We denote by $\Lambda_n(\omega)$ the spectral radius of $X^{(n)}(\omega)$.

Our aim is the study of the asymptotic behavior of the sequence $(X^{(n)})_{n\geq 1}$. For the main part, our results concern the sequences of real random variables

$$\big(\langle y, X^{(n)} x\rangle\big)_{n\geq 1}, \qquad x, y \in \overline{C}\setminus\{0\}.$$

In these are included the sequences of matrix entries $(\langle e_i, X^{(n)} e_j\rangle)_{n\geq 1}$, $i, j = 1, \dots, q$, as well as the sequences of norms of the transforms of a positive vector $(||X^{(n)} x||)_{n\geq 1}$, $x \in \overline{C}\setminus\{0\}$, as is seen by choosing for $y$ the vector all coordinates of which are equal to 1.

The hypothesis common to all the following statements is

$$(\mathscr{C}) \qquad P\bigg(\bigcup_{n\geq 1}\big[X^{(n)} \in S^\circ\big]\bigg) > 0.$$

Consider the particular case of independent $X_n$, $n \geq 1$, to which we will henceforth refer as the "independent case." Then it is clear that condition $(\mathscr{C})$ only depends on the support of the random matrix $X_1$.

Returning to the general case, we notice that $S^\circ$ is an ideal of $S$; that is, if $g \in S^\circ$ and $g' \in S$, then $g'g \in S^\circ$, so that, for $x, y \in \overline{C}\setminus\{0\}$, $\langle y, g'gx\rangle > 0$. Consequently, the stopping time $T$ defined by

$$T(\omega) = \inf\{n: n \geq 1,\ X^{(n)}(\omega) \in S^\circ\}$$

is of great importance in what follows. It will be established in Lemma 3.1 that, under $(\mathscr{C})$, $P[T < +\infty] = 1$.

Before we proceed, let us agree on two notations. First, if $X$ is a random variable in $S$, we set $X^*(\omega) = (X(\omega))^*$. Second, if $1 \leq p$, we denote

$$m_p = E\big[\big|\ln ||X_0^*||\big|^p\big]^{1/p} + E\big[\big|\ln v(X_0^*)\big|^p\big]^{1/p}.$$

Notice that, according to the norm equivalence on the space of $q \times q$ matrices, we may write equally well $E[|\ln ||X_0^*||\,|^p] < +\infty$ or $E[|\ln ||X_0||\,|^p] < +\infty$.

In the sequel, unless otherwise stated, $e_i$, $e_j$ are arbitrary basis vectors.

1.2. *Asymptotic behavior in direction.* If $a, b \in \mathbb{R}^q$, we denote by $a \otimes b$ the $q \times q$ matrix defined by

$$\langle e_i, (a \otimes b)e_j \rangle = \langle e_i, a \rangle \langle e_j, b \rangle, \qquad i, j = 1, \dots, q.$$

When $\langle a, b \rangle = 1$, $a \otimes b$ is the matrix of the projector on span$(a)$ associated with the decomposition $\mathbb{R}^q = \text{span}(a) \oplus b^\perp$.

According to the Perron–Frobenius theorem ([22] and [30]), if $g \in S$ has spectral radius $\lambda$, then $\lambda$ is a strictly positive eigenvalue of $g$ and $g^*$ and it is possible to choose $r, l \in \overline{C} \setminus \{0\}$ such that

$$gr = \lambda r, \qquad g^* l = \lambda l, \qquad ||l|| = 1, \qquad \langle l, r \rangle = 1.$$

If $g \in S^\circ$, if the eigenvalue $\lambda$ is simple and all other eigenvalues have a strictly smaller modulus, and if $r$ and $l$ are uniquely given by the preceding relations and belong to $C$, then we have $\lim_n \lambda^{-n} g^n = r \otimes l$.

THEOREM 1. *Assume ($\mathscr{C}$). Let $R_n$, $L_n$ be random vectors in $\overline{C} \setminus \{0\}$ such that*

$$X^{(n)} R_n = \Lambda_n R_n, \qquad X^{(n)*} L_n = \Lambda_n L_n, \qquad ||L_n|| = 1, \qquad \langle L_n, R_n \rangle = 1.$$

*Then:*

(i)

$$\lim_n \frac{\langle e_i, X^{(n)} e_j \rangle}{\Lambda_n \langle R_n, e_i \rangle \langle L_n, e_j \rangle} 1_{[T \le n]} = 1 \text{ a.s.},$$

$$\lim \left( \Lambda_n^{-1} X^{(n)} - R_n \otimes L_n \right) = 0 \text{ a.s.};$$

(ii) *there exist random unit vectors $Z_1$ and $Z_1'$ in $C$ such that*
  (a) $(L_n)_{n \ge 1}$ *converges almost surely to $Z_1$,*
  (b) $(R_n / ||R_n||)_{n \ge 1}$ *converges weakly to $Z_1'$;*

(iii) *moreover, if $\theta$ is mixing:*
  (a) *the sequence $(R_n/||R_n||, L_n)_{n \ge 1}$ converges weakly to $\nu' \times \nu$, where $\nu$ and $\nu'$ are the laws of $Z_1$ and $Z_1'$;*
  (b) *the sequence $(\Lambda_n^{-1} X^{(n)})_{n \ge 1}$ converges weakly to the probability measure on $S^\circ$ which is the image of $\nu' \times \nu$ under the function $h$ defined on $C \times C$ by $h(z', z) = z' \otimes z / \langle z', z \rangle$.*

Point (i) compares the random matrices $\Lambda_n^{-1} X^{(n)}$ and the random projectors $R_n \otimes L_n$. With respect to almost sure convergence, the first assertion says that at infinity the entries of these matrices are equivalent; the second mimics what was recalled previously for the powers of an element $g \in S^\circ$. Joined to (ii), this first statement shows that, in direction, the sequence of row vectors of $X^{(n)}$ converges almost surely, while the sequence of column vectors converges weakly.

Notice that, if the matrices $X_n$ are stochastic, then the sequence $(X^{(n)})_{n \ge 1}$ converges a.s.

The asymptotic link between $\Lambda_n$ and $||X^{(n)}||^{1/n}$ will appear in Theorem 2 under $m_1 < +\infty$. However, the preceding theorem may be adapted to normalized matrix products. Recall [22] that a norm $|||\cdot|||$ on a vector space ordered by a cone $\overline{C}$ is said to be monotone if, for all $x, y \in \overline{C}$ such that $y - x \in \overline{C}$, $|||x||| \le |||y|||$.

COROLLARY 1. *Let $|||\cdot|||$ denote both a monotone norm on $\mathbb{R}^q$ and the induced norm on the space of $q \times q$ matrices. Under the hypothesis of Theorem 1, we have:*

(i)

$$\lim_n \Lambda_n \frac{|||R_n \otimes L_n|||}{|||X^{(n)}|||} = 1 \text{ a.s.},$$

$$\lim_n \left( \frac{X^{(n)}}{|||X^{(n)}|||} - \frac{R_n \otimes L_n}{|||R_n \otimes L_n|||} \right) = 0 \text{ a.s.};$$

(ii) *if $\theta$ is mixing, the sequence $(X^{(n)}/|||X^{(n)}|||)_{n\ge1}$ converges weakly to the probability measure on $S^\circ$ which is the image of $\nu' \times \nu$ by the function $h_1 = |||h|||^{-1}h$.*

1.3. *The strong law of large numbers.* Recall [23] that, under the hypothesis $E[\ln^+ ||X_1||] < +\infty$ which is satisfied if $m_1 < +\infty$, the greatest characteristic exponent of the sequence $(X_n)_{n\ge1}$ is the element $\gamma_1$ of $\mathbb{R} \cup \{-\infty\}$ defined by

$$\gamma_1 = \lim_n \frac{1}{n} E[\ln ||X^{(n)}||],$$

and that, by means of the subadditive ergodic theorem, we have almost surely

$$\lim_n \frac{1}{n} \ln ||X^{(n)}|| = \gamma_1.$$

It appears that, in the present context, the exponent $\gamma_1$ also governs the almost sure asymptotic behavior of the entries and of the spectral radius of the matrices $X^{(n)}$.

THEOREM 2. *Suppose $(\mathscr{C})$ and $m_1 < +\infty$. Then $\gamma_1 > -\infty$ and we have almost surely*

$$\lim_n \sup\left\{ \left| \frac{1}{n} 1_{[T\ge n]} \ln\langle y, X^{(n)}x\rangle - \gamma_1 \right| : x, y \in \overline{C}, ||x|| = ||y|| = 1 \right\} = 0,$$

$$\lim_n \frac{1}{n} \ln \Lambda_n = \gamma_1.$$

In Section 3, $(\mathscr{C})$ will be understood as a contraction property of the sequence $(X^{(n)})_{n\ge1}$ with which we can associate a coefficient $\kappa \in [0, 1[$. As in the case of independent matrices ([3], Proposition 3-6-4), $\kappa$ allows us to compare the two leading characteristic exponents.

COROLLARY 2. *Let $\gamma_2 \in \mathbb{R} \cup \{-\infty\}$ be the second characteristic exponent of the sequence $(X_n)_{n\ge1}$. Under the hypothesis of Theorem 2, we have*

$$\gamma_2 \le \gamma_1 - \ln \frac{1}{\kappa}.$$

1.4. *The central limit theorem.* For $n \geq 1$, we denote

$\mathscr{F}_n$ the $\sigma$-field generated by the random variables $X_k$, $0 \leq k \leq n$,
$\mathscr{F}^n$ the $\sigma$-field generated by the random variables $X_k$, $n \leq k$.

We set

$$\alpha_n = \sup_{k \geq 0} \sup \Big\{ |P(A \cap B) - P(A)P(B)|: A \in \mathscr{F}_k,\ B \in \mathscr{F}^{n+k} \Big\},$$

$$\rho_n = \sup_{k \geq 0} \sup \Big\{ |\operatorname{Cor}(Y, Z)|: Y \in L^2(\mathscr{F}_k),\ Z \in L^2(\mathscr{F}^{n+k}),\ Y \neq 0,\ Z \neq 0 \Big\},$$

where

$$\operatorname{Cor}(Y, Z) = \frac{E\big[(Y - E[Y])(Z - E[Z])\big]}{\sigma(Y)\sigma(Z)}.$$

We will consider the two following situations:

(A) *there exists* $\delta > 0$ *such that* $m_{2+\delta} < +\infty,\ \sum_{n \geq 1} \alpha_n^{\delta/(2+\delta)} < +\infty$;
(B) $m_2 < +\infty,\ \sum_{n \geq 1} \rho_n < +\infty$.

The independent case is a particular case of (B). Conditions (A) and (B) are verified when $(X_n)_{n \geq 1}$ is a finite stationary ergodic Markov chain on $S$.

THEOREM 3. *Assume that* $(\mathscr{C})$ *and one of the conditions* (A) *or* (B) *is satisfied. Then, for all sequences* $(x_n)_{n \geq 1}$ *and* $(y_n)_{n \geq 1}$ *of unit vectors of* $\overline{C}$, *the random sequences*

$$\left( \frac{1}{\sqrt{n}} 1_{[T \leq n]} \Big( \ln \langle y_n, X^{(n)} x_n \rangle - n\gamma_1 \Big) \right)_{n \geq 1},$$
$$\left( \frac{1}{\sqrt{n}} \Big( \ln \Lambda_n - n\gamma_1 \Big) \right)_{n \geq 1}$$

*converge weakly to a centered normal law whose variance is denoted by* $\sigma^2$.

Notice that, in each case, moments and mixing conditions are the same as in the standard theorems for sums of real random variables. For the preceding result to be a meaningful central limit theorem, it is necessary that $\sigma^2 > 0$. So we have now to study the case $\sigma^2 = 0$.

COROLLARY 3. *Under the hypothesis of Theorem 3, if* $\sigma^2 = 0$, *then the sequence* $(e^{-n\gamma_1} \|X^{(n)}\|)_{n \geq 1}$ *is tight in* $]0, +\infty[$.

Since we may always substitute the sequence $(e^{-\gamma_1} X_n)_{n \geq 1}$ for the initial sequence $(X_n)_{n \geq 1}$, we are led to study the tightness properties of nonnormalized random products.

1.5. *Tightness.* Although $(X^{(n)})_{n\geq 1}$ is a sequence of random matrices in $S$, it is natural to consider a tightness property on the closed semigroup $\overline{S}$ of $q\times q$ matrices with positive entries.

Tightness of the sequence $(X^{(n)})_{n\geq 1}$ may be deduced from the tightness of certain associated sequences of real random variables. For the statement of the precise result, we consider functions $\jmath$ on $\overline{S}$ which are continuous, strictly positive on $S^\circ$, homogeneous and increasing; that is, for $g, g' \in \overline{S}$, $\lambda > 0$, $\jmath(g) \geq 0$, $\jmath(\lambda g) = \lambda \jmath(g)$, if $g' - g \in \overline{S}$, $\jmath(g') \geq \jmath(g)$ and, if $g \in S^\circ$, $\jmath(g) > 0$. The norm and the function $v$ on $\overline{S}$ induced by an increasing norm on $\mathbb{R}^q$ are examples of the function $\jmath$, as is $v_m$, given by

$$v_m(g) = \min\{\langle e_i, g e_j\rangle: i, j = 1, \ldots, q\}.$$

THEOREM 4. (i) *Assume that $\theta$ is mixing, that condition $(\mathscr{C})$ is satisfied and that the sequence $(\jmath(X^{(n)}))_{n\geq 1}$ is tight in* $[0, +\infty[$. *Then $(X^{(n)})_{n\geq 1}$ is tight in $\overline{S}$ and every weak limit value $\pi$ of this sequence satisfies $\pi(\{0\}\cup S^\circ) = 1$. Moreover, if $\pi\{0\} = 0$, $\pi$ is carried by the subset $Q = \{a\otimes b: a, b \in C\}$ of rank one elements in $S^\circ$.*

(ii) *Assume that $\theta$ is strongly mixing* ($\lim_n \alpha_n = 0$), *that condition $(\mathscr{C})$ is satisfied and that the sequence $(X^{(n)})_{n\geq 1}$ is tight in $\overline{S}$. Then $(X^{(n)})_{n\geq 1}$ converges weakly to the unit mass at* 0 *or each weak limit value $\pi$ of this sequence satisfies $\pi(Q) = 1$.*

If $(X^{(n)})_{n\geq 1}$ is tight in $\overline{S}$, it is tight in $S$ if and only if any of its weak limit value $\pi$ satisfies $\pi(S) = 1$. Under the hypothesis of (i), this condition reduces to $\pi\{0\} = 0$ and follows from

$$\lim_{\ell} \liminf_{n} P\big[\ell^{-1} \leq \|X^{(n)}\|\big] = 1.$$

In the independent case, the tightness in $S$ may be characterized in a geometric way and we get a convergence result. Some additional notation is necessary.

We denote by $\mathscr{K}^+$ the collection of the affine subspaces $A$ of $\mathbb{R}^q$ such that $A\cap C \neq \varnothing$ and $(A-A)\cap\overline{C} = \{0\}$ (or, equivalently, $A\cap C$ is a nonempty bounded subset). The semigroup of elements in $S$ preserving $A \in \mathscr{K}^+$ is denoted by $S_A$.

Let $T$ denote the closed subsemigroup of $S$ spanned by the identity matrix and the support of $X_1$ in $S$.

THEOREM 5. *Assume that the random matrices $X_n$, $n \geq 1$, are independent and that condition $(\mathscr{C})$ is verified.*

*Then $(X^{(n)})_{n\geq 1}$ is tight in $S$ if and only if there exists $A \in \mathscr{K}^+$ such that $T \subset S_A$.*

*The set $A$ may be described as the affine subspace spanned by $Tr_m$, where $r_m$ is a vector of $C$ such that $E[X_1]r_m = r_m$, $E[X_1]$ being the mean matrix $\langle e_i, E[X_1]e_j\rangle = E[\langle e_i, X_1 e_j\rangle]$.*

To proceed, we need to describe the structure of elements in $S_A$, $A \in \mathscr{K}^+$. There exist measurable functions on $S_A$: $r(\cdot)$, $l(\cdot)$ and $w(\cdot)$ such that

$$r(g) \in A \cap \overline{C}, \quad gr(g) = r(g), \quad l(g) \in \overline{C}, \quad g^* l(g) = l(g), \quad \langle r(g), l(g) \rangle = 1,$$

$$w(g)A \subset A - A, \qquad w(g) \text{ has spectral radius } \lambda(w(g)) \leq 1,$$

$$g = r(g) \otimes l(g) + w(g).$$

Notice that now the normalization of left and right eigenvectors is not the same as in Section 1.2.

COROLLARY 4. *Under the hypothesis of Theorem 5,* $(X^{(n)})_{n\geq 1}$ *converges weakly to the probability* $h(\tilde{\nu}' \times \tilde{\nu})$*, where* $\tilde{\nu}'$ *and* $\tilde{\nu}$ *are the probability distributions of the almost surely converging series*

$$\sum_{n\geq 1} w(X_1^*) \cdots w(X_{n-1}^*) l(X_n), \qquad \sum_{n\geq 1} w(X_{-1}) \cdots w(X_{-n+1}) r(X_{-n}).$$

The preceding results clearly apply to products of random stochastic matrices.

1.6. *Connection with previous results.* As far as I know, stationarity and $(\mathscr{C})$ are the weakest conditions that have been considered when dealing with random products of positive matrices in order to establish limit theorems. The method of this paper differs from those of the previous ones by the systematic use of the contracting action of matrices on the projective space, a method which is usual when studying products of random, independent, invertible matrices ([3], [5], [14] and [17]).

*Behavior in direction.* Behavior in direction is related to the notion of weak ergodicity for products of random, positive matrices as developed by Cohn, Nerman and Peligrad [6], and others [30]. The work of Kesten and Spitzer [20] gives some answers for this problem in the independent case. Theorem 1 takes place in a more general context; however, [20] has inspired the algebraic lemma (Lemma 4.1). The stationary sequence $(Z_k)_{k\geq 1}$ in Lemma 3.3 is the projective image of the generalized eigenvector introduced by Evstigneev [12]. The direction $Z_1$ is of main importance to state a kind of multiplicative ergodic theorem for random, strictly positive matrices [1]; it may be generalized to the case of transfer operators ([1] and [13]).

*Law of large numbers.* Generalizing a result of Furstenberg and Kesten [15], Kingman [21] has established that

$$\lim_n \frac{1}{n} \ln\langle e_i, X^{(n)} e_j \rangle = \gamma_1 \quad \text{a.s.,}$$

under the hypothesis $\max_{i,\, j=1,\ldots,q} E[|\ln\langle e_i, X_1 e_j\rangle|] < +\infty$, which implies $m_1 < +\infty$ and strengthens $(\mathscr{C})$. The proof is based on the subadditive ergodic theorem stated by this author. Avoiding this powerful tool, Cohn, Nerman and Peligrad [6] have established the same result in an elementary way. The relation between the two leading characteristic exponents and the contraction rate stated in Corollary 2 has been obtained by Peres [28] for the case of the Hilbert metric, using an algebraic formula due to Hopf.

*Central limit theorem.* Set

$$\beta_n = \sup_{k\geq 0} \sup\{|P(B|A) - P(B)|: A \in \mathscr{F}_k,\ P(A) > 0,\ B \in \mathscr{F}^{n+k}\}.$$

Note that $\beta_n \geq \alpha_n$. Assuming:

(i) almost surely $\min_{i,\, j=1,\dots,q}\langle e_i, X_0 e_j\rangle > 0$ and $\max_{i,\, j=1,\dots,q}\langle e_i, X_0 e_j\rangle / \min_{i,\, j=1,\dots,q}\langle e_i, X_0 e_j\rangle$ is bounded;
(ii) there exists $\delta > 0$ such that $\max_{i,\, j=1,\dots,q} E[|\ln\langle e_i, X_0 e_j\rangle|^{2+\delta}] < +\infty$;
(iii) $(\beta_n)_{n\geq 1}$ tends to 0 exponentially fast.

Furstenberg and Kesten [15] have obtained weak convergence to a normal law.

This result is strengthened by Cohn, Nerman and Peligrad [6], since their first set of hypotheses only supposes the existence of $\delta > 0$ such that:

(i′) $\max_{i,\, j=1,\dots,q} E[|\ln\langle e_i, X_1 e_j\rangle|^{2+\delta}] < +\infty$;
(ii′) $\sum_{n\geq 1} \alpha_n^{\delta/(2+\delta)} < +\infty$.

The second set of hypotheses considered by these authors to get a central limit theorem is:

(i″) $\max_{i,\, j=1,\dots,q} E[|\ln\langle e_i, X_1 e_j\rangle|^{2}] < +\infty$;
(ii″) $\sum_{n\geq 1} \rho_n < +\infty$.

It is clear that the conditions of our Theorem 3 are weaker than the previous ones. Notice that the proofs of [6] are based on Theorem 5.4 of [18], whose proof was corrected by Esseen and Janson [11]; see also Volny [31].

*Tightness.* The results of Theorem 5 are more general and precise than those of Kesten and Spitzer [20]; moreover, the proofs are simpler. The hypotheses of [20] are $(\mathscr{C})$ and the fact that $(v_m(X^{(n)}))_n$ is tight in $[0, +\infty[$; as already noted, $v_m$ is a particular function $\jmath$. In [20] it is proved that the sequence $(X^{(n)})_n$ converges weakly to a probability on $S^\circ$ and in a particular case it is shown that the limit distribution is of the form described here. The central idea of the proof of Theorem 5 is an argument of Raugi [29]. A similar technique has been used by Bougerol [2] to describe the structure of tight products of independent not necessarily positive matrices. The favorable case of positive matrices allows more precise statements and shorter proofs. Notice also that, using semigroup methods, Mukherjea [26] gave alternative proofs of some of Kesten and Spitzer's results.

The limit theorems stated previously are similar to those obtained for products of independent or Markov-dependent, invertible, random matrices, by

Furstenberg [14], Le Page [24] and Guivarc'h and Raugi [17]. For a more precise comparison, look at the results obtained here in the independent case. Positivity allows us to avoid the invertibility of $X_1$ and the strong irreducibility of its support and, in the central limit theorem, permits the use of a method which eliminates the exponential moment condition. Although the case of positive matrices is much easier to handle than the general one, it is worth noting that, in both cases, the crux is a contraction property.

2. Notation and general points. First, it will be convenient to modify the dynamical system setting. Without restricting the generality of the previously stated results, we may replace the original dynamical system by a natural extension, still denoted $(\Omega, \mathscr{F}, P, \theta)$, for which $\theta$ is invertible [7]. The sequence $(X_n)_{n\in\mathbb{Z}}$ is then defined by $X_n = X_0 \circ \theta^n,\ n \in \mathbb{Z}$.

The definitions of the $\sigma$-fields $\mathscr{F}_n$, $\mathscr{F}^n$ are suitably extended.

The dual action of the random matrices $X_n$, $n \in \mathbb{Z}$, will have a leading role in the sequel, so we set

$$Y_n = X_n^*, \qquad n \in \mathbb{Z}, \qquad Y^{(n)} = X^{(n)*} = Y_1 \cdots Y_n, \qquad n \geq 1.$$

Recall that the norm used on $\mathbb{R}^q$ is defined by

$$x \in \mathbb{R}^q, \qquad ||x|| = \sum_{i=1}^{q} |\langle x, e_i \rangle|,$$

which is not the norm associated with the scalar product $\langle \cdot, \cdot \rangle$; however, the inequality $|\langle x, y\rangle| \leq ||x||\,||y||$ is satisfied. We have also set

$$C = \{x\colon x \in \mathbb{R}^q,\ \forall i = 1, \dots, q,\ \langle x, e_i \rangle > 0\},$$

$$\overline{C} = \{x\colon x \in \mathbb{R}^q,\ \forall i = 1, \dots, q,\ \langle x, e_i \rangle \geq 0\}.$$

As usual, when studying random matrix products, it is convenient to introduce the projective action of these matrices [3].

Projectively, that is, when only the directions are considered, the elements of $C$ and $\overline{C}$ are represented by points of the open and closed "polygons"

$$B = C \cap \{x\colon x \in \mathbb{R}^q,\ ||x|| = 1\}, \qquad \overline{B} = \overline{C} \cap \{x\colon x \in \mathbb{R}^q,\ ||x|| = 1\}.$$

We pick out in $B$ the point $\chi = (1/q, \dots, 1/q)$. On $\overline{C}$, the norm and scalar product are then connected by the relation $q\langle x, \chi \rangle = ||x||$.

We have that $S$ acts on $\overline{B}$. Precisely, every $g \in S$ induces a transformation of $\overline{B}$; the image of $x \in \overline{B}$ is denoted $g \cdot x$ and defined by

$$g \cdot x = \frac{gx}{||gx||}.$$

If $e$ stands for the identity matrix and $g, g' \in S$, $x \in \overline{B}$,

$$e \cdot x = x, \qquad (gg') \cdot x = g \cdot (g' \cdot x).$$

Notice the importance of $\cdot$ to distinguish between the action of elements of $S$ on $\overline{C}$ and on $\overline{B}$.

From now on, condition $(\mathscr{C})$ is assumed. It implies the contraction properties stated in Section 3. Having these at hand, we establish Theorem 1 in Section 4 and tackle the proofs of Theorems 2 and 3 in Section 5.

## 3. Contraction properties.

3.1. *Contraction properties of positive matrices.* We construct a bounded distance $d$ on $\overline{B}$, such that, with respect to $d$, any element $g \in S$ acts on $\overline{B}$ as a contraction.

Let $x = (x_1, \dots, x_q)$, $y = (y_1, \dots, y_q)$ be two points in $\overline{B}$. We write

$$m(x, y) = \sup\{\lambda: \lambda \in \mathbb{R}_+,\ \forall i = 1, \dots, q,\ \lambda y_i \le x_i\}$$
$$= \min\left\{\frac{x_i}{y_i}: i = 1, \dots, q,\ y_i > 0\right\}.$$

As $\sum_{i=1}^q x_i = \sum_{i=1}^q y_i = 1$, we have $0 \le m(x, y) \le 1$. We write

$$d(x, y) = \varphi(m(x, y)m(y, x)),$$

where $\varphi$ is the one-to-one function on $[0, 1]$ defined by

$$\varphi(s) = \frac{1 - s}{1 + s}.$$

PROPOSITION 3.1. *$d$ is a distance on $\overline{B}$ having the following properties:*

(i) $\sup\{d(x, y): x, y \in \overline{B}\} = 1$;
(ii) *if* $x, y \in \overline{B}$, $\|x - y\| \le 2d(x, y)$.

*For $g \in S$, there exists $c(g) \le 1$ such that:*

(iii) *if* $x, y \in \overline{B}$, $d(g \cdot x, g \cdot y) \le c(g)d(x, y) \le c(g)$;
(iv) $c(g) < 1$ *if and only if* $g \in S^\circ$;
(v) *if* $g' \in S$, $c(gg') \le c(g)c(g')$;
(vi) $c(g^*) = c(g)$.

This result will be proved in Section 10, where we shall also establish a formula that expresses $c(g)$ as a function of the entries of the matrix $g$.

REMARK. The Hilbert distance on $B$ is defined by

$$d_H(x, y) = -\ln\big(m(x, y)m(y, x)\big).$$

It is connected to the preceding distance by

$$d(x, y) = \tanh\big(\tfrac{1}{2}d_H(x, y)\big).$$

On one hand, one knows ([3], page 59, [4] and [22]) that properties (iii)–(vi) are satisfied by $d_H$. On the other hand,

$$\sup\{d_H(x, y): x, y \in B\} = +\infty,$$

so that (i) makes $d$ more convenient than $d_H$.

The distance $d$ has been introduced in [4], for the two-dimensional case, and, more recently, in [27], the relevant contraction coefficient $c(g)$ has been associated with the norm distance on $\overline{B}$. As quoted in [4], if $c_H(g)$ is the contraction coefficient of $g \in S$ with respect to $d_H$, we have $c_H(g) \le c(g)$. But it is of no interest for the sequel to use the best contraction coefficient.

3.2. *Stochastic contraction properties.* Recall that $T(\omega) = \inf\{n\colon n \ge 1, X^{(n)}(\omega) \in S^\circ\}$.

LEMMA 3.1. *We have $P[T < +\infty] = 1$, $E[T] < +\infty$ and, for $n \ge T(\omega)$, $X^{(n)}(\omega) \in S^\circ$.*

PROOF. From $(\mathscr{C})$, there exists $b \in \mathbb{N}^*$ such that $P[X^{(b)} \in S^\circ] > 0$. Set $T'(\omega) = \inf\{n\colon n \ge 1,\ X^{(b)}(\theta^n(\omega)) \in S^\circ\}$. From Kac's lemma [7], we know that $P[T' < +\infty] = 1$ and that $E[T'] < +\infty$. But, if $g \in S$ and $g' \in S^\circ$, $g'g$ and $gg' \in S^\circ$. From this we deduce that $T(\omega) \le T'(\omega) + b$ and that, for $n \ge T(\omega)$, $X^{(n)}(\omega) \in S^\circ$. □

We now explain the strictly contracting action of the matrices $Y^{(n)}$ and $X^{(n)}$, $n \ge 1$, on $\overline{B}$.

LEMMA 3.2. *The contraction coefficient of the sequence $(Y^{(n)})_{n\ge 1}$ is the real number $\kappa \in [0, 1[$ defined by*

$$\ln \kappa = \lim_n \frac{1}{n} E[\ln c(Y^{(n)})] = \inf_n \frac{1}{n} E[\ln c(Y^{(n)})].$$

*We have $\lim_n (c(Y^{(n)}))^{1/n} = \kappa$ a.s. In particular, $\lim_n c(Y^{(n)}) = \lim_n c(X^{(n)}) = 0$ a.s.*

PROOF. We use the properties of $c$ given previously. For $m$ and $n \ge 1$,

$$\ln c(Y^{(m+n)}) \le \ln c(Y^{(m)}) + \ln c(Y_{m+1} \cdots Y_{m+n}) = \ln c(Y^{(m)}) + \ln c(Y^{(n)} \circ \theta^m).$$

As $\ln c(Y^{(n)}) \le 0$, the subadditive ergodic theorem does the main part of the proof. It is clear that $\kappa < 1$, because by $(\mathscr{C})$ there exists $b \in \mathbb{N}^*$ such that $-\infty \le (1/b)E[\ln c(Y^{(b)})] < 0$. At last, since $c(g) = c(g^*)$, $\kappa$ is also the contraction coefficient of the sequence $(X_n)_{n\ge 1}$. □

LEMMA 3.3. *There exists a stationary, ergodic sequence $(Z_k)_{k\in\mathbb{Z}}$ of random elements of $B$ such that:*

(i) *$Z_k$ is $\mathscr{F}^k$-measurable;*
(ii) *for $k \in \mathbb{Z}$, $Y_k \cdot Z_{k+1} = Z_k$;*
(iii) *for $y \in \overline{B}$ and $k, n \in \mathbb{Z}$, $k \le n$, $d((Y_k \cdots Y_n) \cdot y, Z_k) \le c(Y_k \cdots Y_n)$.*

PROOF. Set $\Omega_1 = \{\omega\colon \lim_n c(Y^{(n)}(\omega)) = 0\}$ and note that $P(\Omega_1) = 1$.

Let $\omega \in \Omega_1$. The "polygons" $K_n(\omega) = Y^{(n)}(\omega) \cdot (\overline{B})$ form a decreasing sequence of compact subsets of $\overline{B}$ endowed with the canonical topology on $\mathbb{R}^q$, so that $K(\omega) = \bigcap_{n\geq 1} K_n(\omega) \neq \varnothing$. For the distance $d$, the diameter $\Delta(\omega)$ of $K(\omega)$ is equal to 0. In fact, for $n \geq 1$,

$$\Delta(\omega) \leq \Delta_n(\omega) = \sup\{d(Y^{(n)}(\omega)\cdot x, Y^{(n)}(\omega)\cdot y)\} \leq c(Y^{(n)}(\omega)).$$

We define $Z_1(\omega)$ by $K(\omega) = \{Z_1(\omega)\}$.

$Z_1(\omega) \in B$, since for $n$ large enough $Y^{(n)}(\omega) \in S^{\circ}$.

Let $y \in \overline{B}$. As $Y^{(n)}(\omega)\cdot y \in K_n(\omega)$, we have, from (ii) in Proposition 3.1, Lemma 3.2 and the preceding inequality,

$$\lim_n \|Y^{(n)}(\omega)\cdot y - Z_1(\omega)\| = \lim_n d(Y^{(n)}(\omega)\cdot y, Z_1(\omega)) = 0,$$

and, more precisely, considering distance $d$,

$$d(Y^{(n)}(\omega)\cdot y, Z_1(\omega)) \leq \Delta_n(\omega) \leq c(Y^{(n)}(\omega)).$$

For $k \in \mathbb{Z}$, we set $Z_k = Z_1 \circ \theta^{(k-1)}$.

We have a.s. $Z_k = \lim_n (Y_k \cdots Y_n)\cdot y = Y_k \cdot Z_{k+1}$. The other properties of $(Z_k)_{k\in\mathbb{Z}}$ are clear. □

## 4. Proof of Theorem 1 and Corollary 1.

4.1. *Proof of* (i). Let $\Omega_1 \in \mathscr{F}$ with $P(\Omega_1) = 1$ such that, for $\omega \in \Omega_1$, $\lim_n c(Y^{(n)})(\omega) = 0$ and $T(\omega) < +\infty$.

We fix $\omega \in \Omega_1$. Since $Z_1(\omega) \in B$, $\inf\{\langle Z_1(\omega), x\rangle\colon x \in \overline{B}\} = a > 0$. For $x, y \in \overline{B}$, omitting $\omega$ for a while, we write

$$\left|\frac{\langle Y^{(n)}\cdot y, x\rangle}{\langle Z_1, x\rangle} - 1\right| = \left|\frac{\langle Y^{(n)}\cdot y - Z_1, x\rangle}{\langle Z_1, x\rangle}\right| \leq \frac{\|Y^{(n)}\cdot y - Z_1\|\|x\|}{\langle Z_1, x\rangle}$$
$$\leq 2\frac{d(Y^{(n)}\cdot y, Z_1)}{\langle Z_1, x\rangle} \leq \frac{2}{a} c(Y^{(n)}).$$

Because

$$\langle Y^{(n)}\cdot y, x\rangle = \frac{\langle Y^{(n)} y, x\rangle}{\|Y^{(n)} y\|} = \frac{\langle y, X^{(n)} x\rangle}{q\langle X^{(n)}\chi, y\rangle},$$

we get

$$\langle y, X^{(n)}(\omega)x\rangle = q\langle X^{(n)}(\omega)\chi, y\rangle\langle Z_1(\omega), x\rangle(1 + \varepsilon_n(x, y)(\omega)),$$

where $(\varepsilon_n(x, y)(\omega))_n$ converges to 0, uniformly for $x, y \in \overline{B}$. Using homogeneity, we may still write this relation for $x, y \in \overline{C}$, with $\lim_n \sup\{|\varepsilon_n(x, y)(\omega)|\colon x, y \in \overline{C}\} = 0$.

Points (i) of Theorem 1 and Corollary 1 are direct consequences of the following lemma inspired by [20]. Notice, for later use, that, if $\jmath$ is as in Theorem 4, the last statement of this lemma implies

$$\lim_n \frac{\jmath(R_n \otimes L_n)}{\langle e_i, (R_n \otimes L_n) e_j\rangle} \frac{\langle e_i, X^{(n)} e_j\rangle}{\jmath(X^{(n)})} 1_{[T\le n]} = 1 \quad \text{a.s.}$$

LEMMA 4.1. *Let $(g_n)_n$ be a sequence in $S^\circ$, $x_0$ an element of $B$ and $(y_n)_n$ a sequence in $C$ such that, for $x, y \in \overline{C}$,*

$$\langle y, g_n x\rangle = \langle y_n, y\rangle \langle x_0, x\rangle \big(1 + \varepsilon_n(x, y)\big),$$

*where $\lim_n \sup\{|\varepsilon_n(x, y)|: x, y \in \overline{C}\} = 0$. Then, if $\lambda_n$ is the spectral radius of $g_n$ and if $r_n$, $l_n$ are vectors in $C$ defined by*

$$g_n^* l_n = \lambda_n l_n, \qquad g_n r_n = \lambda_n r_n, \qquad \|l_n\| = 1, \qquad \langle l_n, r_n\rangle = 1,$$

*we have*

$$\lim_n l_n = x_0, \qquad \lim_n \frac{\langle e_i, g_n e_j\rangle}{\lambda_n \langle r_n, e_i\rangle \langle l_n, e_j\rangle} = 1, \qquad \lim_n \big(\lambda_n^{-1} g_n - r_n \otimes l_n\big) = 0.$$

*Moreover, if $|||\cdot|||$ is as in Corollary 1 and $\jmath$ is as in Theorem 4,*

$$\lambda_n \sim \frac{|||g_n|||}{|||r_n \otimes l_n|||}, \qquad \lim_n \left( \frac{g_n}{|||g_n|||} - \frac{r_n \otimes l_n}{|||r_n \otimes l_n|||} \right) = 0,$$

$$\lim_n \frac{\jmath(r_n \otimes l_n)}{\langle e_i, (r_n \otimes l_n) e_j\rangle} \frac{\langle e_i, g_n e_j\rangle}{\jmath(g_n)} = 1.$$

PROOF. Let $\varepsilon$, $0 < \varepsilon < 1$, and let $n$ be such that, for $x, y \in \overline{C} \backslash \{0\}$,

$$1 - \varepsilon \le \frac{\langle y, g_n x\rangle}{\langle y_n, y\rangle \langle x_0, x\rangle} \le 1 + \varepsilon.$$

Consider the two double inequalities obtained from the preceding one by substituting $r_n$ for $x$ and $l_n$ for $y$. Term-by-term multiplication of these inequalities leads to

$$(1 - \varepsilon)^2 \le \frac{\lambda_n^2 \langle y, r_n\rangle \langle l_n, x\rangle}{\langle y_n, y\rangle \langle x_0, r_n\rangle \langle y_n, l_n\rangle \langle x_0, x\rangle} \le (1 + \varepsilon)^2.$$

Setting $y = \chi$, we see that there exists a real positive sequence $(c_n)_n$ such that $(c_n(\langle l_n, x\rangle / \langle x_0, x\rangle))_n$ converges. As $\|l_n\| = \|x_0\| = 1$ it follows that $\lim_n l_n = x_0$.

Dividing the middle terms of the two double inequalities already written, we now get $d_n$ such that, for all $x, y \in \overline{C} \backslash \{0\}$,

$$\frac{1 - \varepsilon}{(1 + \varepsilon)^2} \le d_n \frac{\langle y, g_n x\rangle}{\langle y, r_n\rangle \langle l_n, x\rangle} \le \frac{1 + \varepsilon}{(1 - \varepsilon)^2}.$$

Assigning simultaneously to $y$ and $x$ the values $l_n$ and $r_n$, we obtain

$$\frac{1-\varepsilon}{(1+\varepsilon)^2} \le d_n \lambda_n \le \frac{1+\varepsilon}{(1-\varepsilon)^2}.$$

Hence

$$\left(\frac{1-\varepsilon}{1+\varepsilon}\right)^3 \le \frac{\langle y, g_n x\rangle}{\lambda_n \langle y, r_n\rangle \langle l_n, x\rangle} \le \left(\frac{1+\varepsilon}{1-\varepsilon}\right)^3,$$

which gives the second assertion.

Since $\lim_n l_n = x_0 \in B$, there exists $a > 0$ and $n_0$, such that, for $n \ge n_0$ and all $i = 1, \dots, q$, $\langle l_n, e_i\rangle \ge a$, the relation $\langle l_n, r_n\rangle = 1$ then implies $\langle r_n, e_i\rangle \le 1/a$. We have $\langle r_n, e_i\rangle\langle l_n, e_j\rangle \le \|l_n\|/a \le 1/a$, so that $\lim_n(\lambda_n^{-1} g_n - r_n \otimes l_n) = 0$. Let $\varepsilon$, $0 < \varepsilon < 1$, and let $n$ be such that, for $i, j = 1, \dots, q$,

$$0 \le \lambda_n \langle r_n, e_i\rangle\langle l_n, e_j\rangle(1-\varepsilon) \le \langle e_i, g_n e_j\rangle \le \lambda_n \langle r_n, e_i\rangle\langle l_n, e_j\rangle(1+\varepsilon).$$

From the properties of $\jmath$, we get

$$\lambda_n \jmath(r_n \otimes l_n)(1-\varepsilon) \le \jmath(g_n) \le \lambda_n \jmath(r_n \otimes l_n)(1+\varepsilon).$$

It follows that

$$\lim_n \frac{\jmath(g_n)}{\lambda_n \jmath(r_n \otimes l_n)} = 1.$$

Replacing $\lambda_n$ by the equivalent so supplied, we deduce that, for $i, j = 1, \dots, q$,

$$\lim_n \frac{\jmath(r_n \otimes l_n)}{\langle e_i, (r_n \otimes l_n) e_j\rangle} \frac{\langle e_i, g_n e_j\rangle}{\jmath(g_n)} = 1.$$

As $r_n \otimes l_n$ is a projector $|||r_n \otimes l_n||| \ge 1$ and

$$\limsup_n \frac{\langle e_i, (r_n \otimes l_n) e_j\rangle}{|||r_n \otimes l_n|||} \le 1/a,$$

the preceding convergence with $\jmath = |||\cdot|||$ leads to

$$\lim_n \left(\frac{\langle e_i, g_n e_j\rangle}{|||g_n|||} - \frac{\langle e_i, (r_n \otimes l_n) e_j\rangle}{|||r_n \otimes l_n|||}\right) = 0. \qquad \square$$

4.2. *Proof of Theorem* 1(ii). We have $d(L_n, Z_1) = d(Y^{(n)} \cdot L_n, Y^{(n)} \cdot Z_{n+1}) \le c(Y^{(n)})$. Hence

$$\lim_n d(L_n, Z_1) = \lim_n \|L_n - Z_1\| = 0 \quad \text{a.s.}$$

The sequence whose general term is $X'_n = X^*_{-n}$, $n \in \mathbb{Z}$, satisfies property $(\mathscr{C})$, so that what was done for $(X_n)_{n\ge 1}$ may be done for $(X'_n)_{n\in\mathbb{Z}}$. Assigning a $'$ to the corresponding elements, we have

$$\lim_n d(L'_n, Z'_1) = \lim_n \|L'_n - Z'_1\| = 0 \quad \text{a.s.,}$$

with the relations

$$L'_n = (Y'_1 \cdots Y'_n) \cdot L'_n = (X_{-1} \cdots X_{-n}) \cdot L'_n$$

and

$$Z'_1 = \lim_n (Y'_1 \cdots Y'_n) \cdot \chi = \lim_n (X_{-1} \cdots X_{-n}) \cdot \chi \quad \text{a.s.}$$

Set $L''_n = L'_n \circ \theta^{n+1}$. The sequence $(L''_n)_{n\geq 1}$ converges weakly to $Z'_1$ and satisfies $X^{(n)} \cdot L''_n = L''_n$. For $n \geq T(\omega)$, $X^{(n)}(\omega) \in S^\circ$, so that $L''_n(\omega) = \|R_n(\omega)\|^{-1} R_n(\omega)$. As $P[T < +\infty] = 1$, we conclude that $(\|R_n\|^{-1} R_n)_{n\geq 1}$ converges weakly to $Z'_1$. □

4.3. *Proof of Theorem* 1(iii). Set $R^1_n = \|R_n\|^{-1} R_n$. Let $\phi$, $\psi$ be two continuous functions on the compact set $\overline{B}$ and $k$, $n$ integers such that $1 \leq k \leq n$. Let us write

$$u_n = E[\phi(L_n)\psi(R^1_n)] - E[\phi(Z_1)]E[\psi(Z'_1)] = a(n,k) + b(n,k) + c(n,k) + d(k),$$

where

$$\begin{aligned} a(n,k) &= E\big[\phi(Y^{(n)} \cdot L_n)\psi(X^{(n)} \cdot R^1_n)\big] - E\big[\phi(Y^{(k)} \cdot \chi)\psi(X^{(n)} \cdot R^1_n)\big], \\ b(k,n) &= E\big[\phi(Y^{(k)} \cdot \chi)\psi(X^{(n)} \cdot R^1_n)\big] - E\big[\phi(Y^{(k)} \cdot \chi)\psi((X_n \cdots X_{n-k}) \cdot \chi)\big], \\ c(n,k) &= E\big[\phi(Y^{(k)} \cdot \chi)\psi((X^{(k+1)} \circ \theta^{n-k-1} \cdot \chi)\big] \\ &\quad - E\big[\phi(Y^{(k)} \cdot \chi)\big]E\big[\psi(X^{(k+1)} \cdot \chi)\big], \\ d(k) &= E\big[\phi(Y^{(k)} \cdot \chi)\big]E\big[\psi((X_{-1} \cdots X_{-k-1}) \cdot \chi)\big] - E[\phi(Z_1)]E[\psi(Z'_1)]. \end{aligned}$$

Fix $k$. To bound the sequences $(a(n,k))_n$ and $(b(n,k))_n$, we introduce the continuity modulus $\eta_1$, $\eta_2$ of the functions $\phi$, $\psi$ and their uniform bounds $r_1$, $r_2$. Using Proposition 3.1, we get

$$\begin{aligned} |a(n,k)| &\leq r_2 E[\eta_1(\|Y^{(n)} \cdot L_n - Y^{(k)} \cdot \chi)\|] \leq r_2 E[\eta_1(2c(Y^{(k)}))], \\ |b(n,k)| &\leq r_1 E[\eta_2(\|X^{(n)} \cdot R^1_n - (X_n \cdots X_{n-k}) \cdot \chi\|)] \\ &\leq r_1 E[\eta_2(2c(X_n \cdots X_{n-k}))] \leq r_1 E[\eta_2(2c(X^{(k+1)}))]. \end{aligned}$$

On the other hand, by the mixing property $\lim_n c(n,k) = 0$, so that, for any $k$,

$$\limsup_n |u_n| \leq r_2 E[\eta_1(2c(Y^{(k)}))] + r_1 E[\eta_2(2c(X^{(k+1)}))] + d(k).$$

Letting $k \to \infty$, we conclude that $\lim_n u_n = 0$. In fact, we have seen in the preceding section that $\lim_k d(k) = 0$, while, by Lemma 3.2, $\lim_k c(Y^{(k)}) = \lim_k c(X^{(k)}) = 0$ a.s. It is therefore established that the sequence $((L_n, R^1_n))_{n\geq 1}$ converges weakly to $\nu \times \nu'$.

Since the function $h$ defined in the statement of Theorem 1 is continuous, the sequence of positive matrices $(h(R^1_n, L_n))_{n\geq 1}$ converges weakly to $h(\nu' \times \nu)$. As $\lim_n (\Lambda_n^{-1} X^{(n)} - h(R^1_n, L_n)) = 0$ a.s., it follows that the sequence

$(\Lambda_n^{-1} X^{(n)})_{n\geq 1}$ also converges weakly to $h(\nu' \times \nu)$. Finally, $\nu$ and $\nu'$ being carried by $B$, $h(\nu' \times \nu)$ is carried by $S^\circ$.

Part (ii) of Corollary 1 is obtained along the same lines. □

## 5. Preliminaries to the proofs of Theorems 2 and 3.

5.1. *Main lines of these proofs and general facts.* In each case, the idea is to show that we may reduce the proof to that of a limit theorem for a certain stationary, ergodic sequence of real random variables. This is done in two steps: the first is Lemma 5.1 which is common to the two proofs; the second consists of Lemma 7.1 or Lemma 8.1, depending whether an almost sure or a weak convergence is considered. When this is done, Theorem 2 is nearly proved. However, a lot of work is still necessary to establish Theorem 3 using Gordin's method.

To be precise and to carry out the program, we need some more objects.

We have already defined in Section 2 the space $\overline{B}$ and the action of $g \in S$ on this space. As usual in this context [3], we consider the real positive function $\rho$ defined on $S \times \overline{B}$ by

$$\rho(g, x) = ||gx||.$$

This function is connected to the previously quoted action by the cocycle property: if $g, g' \in S$ and $x \in \overline{B}$,

$$\rho(gg', x) = \rho(g, g' \cdot x)\rho(g', x).$$

On the other hand, for $y \in \overline{B}$, $k, n \in \mathbb{Z}$, $k \leq n$, set

$$Z^y_{n+1,\, n} = y, \qquad Z^y_{k,\, n} = (Y_k \cdots Y_n) \cdot y.$$

These are random vectors in $\overline{B}$.

Provided with these new tools, we can write

$$\begin{aligned} \ln ||Y^{(n)} y|| &= \ln \rho(Y_1 \cdots Y_n, y) = \ln \rho(Y^{(n-1)}, Y_n \cdot y) + \ln \rho(Y_n, y) \\ &= \sum_{k=1}^{n} \ln \rho(Y_k, Z^y_{k+1,\, n}). \end{aligned}$$

We have established in Lemma 3.3(iii) that, when $n \to +\infty$, the sequence $(Z^y_{k+1,\, n})_n$ converges almost surely to the random element $Z_{k+1}$ and that the sequence $(Z_k)_{k\in\mathbb{Z}}$ is stationary and ergodic. Consequently, we are led to replace the preceding sum by the Birkhoff sum

$$\sum_{k=1}^{n} \ln \rho(Y_k, Z_{k+1}).$$

As explained at the beginning, our task will be to show that such a substitution does not affect the asymptotic behavior.

The law of large numbers and connected results are established in Section 7. The central limit theorem is proved in Section 8. The properties of random submultiplicative sequences listed in Section 6 play a central part in these proofs, but only Lemma 6.1 is useful for the law of large numbers.

5.2. *From scalar product to norm.*

LEMMA 5.1. *If $\omega$ is such that $T(\omega) < +\infty$, the sequence whose general term is*

$$D_n(\omega) = \sup\Big\{ \big| 1_{[T\le n]}(\omega) \ln\langle y, X^{(n)}(\omega)x\rangle - \ln \|Y^{(n)}(\omega)y\|\big| : x, y \in \overline{B} \Big\}$$

*is bounded.*

*Consequently, we have* $\lim_n (1/n) D_n = 0$, $\lim_n (1/\sqrt{n}) D_n = 0$ *a.s.*

PROOF. Choose $n \ge T(\omega)$ and denote by $a, b$ two strictly positive real numbers such that, for $i, j = 1, \dots, q$,

$$a \le \langle e_i, X^{(T)}(\omega) e_j\rangle \le b.$$

For $x \in \overline{B}$ and $i = 1, \dots, q$, $a = a\|x\| \le \langle e_i, X^{(T)}(\omega)x\rangle \le b\|x\| = b$. Picking $y \in \overline{B}$, we write

$$\big\langle y, X^{(n)}(\omega)x\big\rangle = \big\langle \big(Y_{T+1}\cdots Y_n\big)(\omega)y, X^{(T)}(\omega)x\big\rangle.$$

Hence, successively,

$$\big\langle \big(Y_{T+1}\cdots Y_n\big)(\omega)y, a(q\chi)\big\rangle \le \big\langle y, X^{(n)}(\omega)x\big\rangle \le \big\langle \big(Y_{T+1}\cdots Y_n\big)(\omega)y, b(q\chi)\big\rangle,$$

$$\big|\ln\langle y, X^{(n)}(\omega)x\rangle - \ln\big\|\big(Y_{T+1}\cdots Y_n\big)(\omega)y\big\|\big| \le \max\{|\ln a|, |\ln b|\}.$$

Substituting $\chi$ for $x$ in this relation, we get

$$\left|\ln\frac{1}{q}\|Y^{(n)}(\omega)y\| - \ln\big\|\big(Y_{T+1}\cdots Y_n\big)(\omega)y\big\|\right| \le \max\{|\ln a|, |\ln b|\}.$$

These two inequalities imply that $(D_n(\omega))_{n\ge1}$ is bounded.

The stated convergences follow now from the fact that $T$ is almost surely finite. □

5.3. *Some technical ingredients.*

LEMMA 5.2. *For $g \in S$ and $x \in \overline{B}$,*

$$v(g) \le \rho(g, x) \le \|g\| \quad \textit{and} \quad \big|\ln\rho(g,x)\big| \le \big|\ln v(g)\big| + \big|\ln\|g\|\big|.$$

PROOF. The proof is obvious. □

It will be of great importance to control the increments of the function $\ln\rho(g,\cdot)$.

LEMMA 5.3. *For $g \in S$, set $r(g) = ||g||/v(g)$. We have:*

(i) *for $x, y \in \overline{B}$, $|\ln \rho(g, x) - \ln \rho(g, y)| \leq 2r(g)d(x, y)$;*
(ii) *for $x, y \in B$, $|\ln \rho(g, x) - \ln \rho(g, y)| \leq 2\ln(1/(1 - d(x, y)))$;*
(iii) *if $g' \in S^\circ$ and if $c(g') \leq \eta < 1$, there exists $\eta_1$ such that, for $x, y \in \overline{B}$,*

$$\left|\ln \rho(g, g' \cdot x) - \ln \rho(g, g' \cdot y)\right| \leq \eta_1 d(g' \cdot x, g' \cdot y) \leq \eta_1 c(g').$$

PROOF. From the mean value inequality,

$$\left|\ln \rho(g, x) - \ln \rho(g, y)\right| \leq \frac{|\rho(g, x) - \rho(g, y)|}{\min\{\rho(g, x), \rho(g, y)\}} \leq \frac{||g||}{v(g)}||x - y||,$$

so that (i) follows from point (ii) of Proposition 3.1.

If $x = (x_1, \dots, x_q)$, $y = (y_1, \dots, y_q)$ and $g = [g_{ij}]_{i,\, j=1,\dots,q}$,

$$\rho(g, x) = \sum_{i=1}^{q}\sum_{j=1}^{q} g_{ij}x_j \geq m(x, y)\sum_{i=1}^{q}\sum_{j=1}^{q} g_{ij}y_j = m(x, y)\rho(g, y).$$

Hence, using the symmetry in $x$ and $y$,

$$m(x, y) \leq \frac{\rho(g, x)}{\rho(g, y)} \leq \frac{1}{m(y, x)}.$$

As $x, y \in B$, $m(x, y), m(y, x) \in ]0, 1]$, so that we get

$$\begin{aligned}\left|\ln \rho(g, x) - \ln \rho(g, y)\right| &\leq \max\{-\ln m(y, x), -\ln m(x, y)\}\\ &\leq -\ln m(y, x) - \ln m(x, y)\\ &= -\ln \varphi^{-1}\big(d(x, y)\big) = \ln \frac{1 + d(x, y)}{1 - d(x, y)}.\end{aligned}$$

For $t \in [0, 1[$,

$$2\ln \frac{1}{1 - t} - \ln \frac{1 + t}{1 - t} = \ln \frac{1}{1 - t^2} \geq 0,$$

and (ii) follows.

Since $g' \in S^\circ$ the inequality (ii) applies to $g' \cdot x$ and $g' \cdot y$, so

$$\left|\ln \rho(g, g'x) - \ln \rho(g, g'y)\right| \leq 2\ln \frac{1}{1 - d(g' \cdot x, g' \cdot y)}.$$

Moreover, $d(g' \cdot x, g' \cdot y) \leq c(g')d(x, y) \leq \eta$. The function $f(t) = 2\ln 1/(1 - t)$ being convex and such that $f(0) = 0$, for all $t \in [0, \eta]$, we have $f(t) \leq (f(\eta)/\eta)t$, hence (iii) holds with $\eta_1 = f(\eta)/\eta$. □

To get rid of the integrability questions, let us end this section with the following obvious statements.

LEMMA 5.4. *For* $p \geq 1$, *denote by* $\|\cdot\|_p$ *the norm on* $L^p(\Omega)$ *and recall that* $m_p = \|\ln\|X_0^*\|\|_p + \|\ln v(X_0^*)\|_p$. *Then:*

(i) *for all* $x \in \overline{B}$, $\|\ln \rho(Y_1, x)\|_p \leq m_p$;
(ii) $\|\ln r(Y_1)\|_p \leq m_p$.

6. Submultiplicative real random sequences. In this section, $(M_n)_{n\geq 1}$ is a sequence of random variables in $[0,1]$ defined on $(\Omega, \mathscr{F}, P, \theta)$, with the following submultiplicative property: for all $m, n \in \mathbb{N}^*$,

$$M_{m+n} \leq M_m \; M_n \circ \theta^m.$$

We define $\tau \in [0,1]$ by

$$\ln \tau = \inf\left\{\frac{1}{n}E[\ln M_n]: n \geq 1\right\} = \lim_n \frac{1}{n}E[\ln M_n],$$

and we set, for $0 \leq m < n$,

$$M_{m,n} = M_{n-m} \circ \theta^m.$$

6.1. *Almost sure convergence.*

LEMMA 6.1. *Choose* $\alpha$, $0 < \alpha \leq 1$, *and let* $n_\alpha$ *be the integral part of* $(1-\alpha)n$. *Then:*

(i) $\limsup_n (1/n)\ln M_{n_\alpha, n} \leq \alpha \ln \tau$ *a.s. in* $\mathbb{R} \cup \{-\infty\}$;
(ii) *if* $\tau < 1$, $(M_n)_{n\geq 1}$ *converges to* 0 *almost surely and in the mean.*

PROOF. Fix $b \in \mathbb{N}^*$ and define the integers $k_n$, $\ell_n$ by

$$k_n b < n_\alpha \leq (k_n+1)b, \qquad (\ell_n+1)b \leq n < (\ell_n+2)b.$$

One easily verifies that

$$\lim_n \frac{k_n}{n} = \frac{1-\alpha}{b}, \qquad \lim_n \frac{\ell_n}{n} = \frac{1}{b}$$

and that, for $\varepsilon > 0$,

$$M_{n_\alpha, n} \leq \prod_{i=k_n+1}^{\ell_n} M_{ib,(i+1)b} \leq \prod_{i=k_n+1}^{\ell_n} M^{\varepsilon}_{ib,(i+1)b},$$

where $M^{\varepsilon}_{ib,(i+1)b} = \sup\{M_{ib,(i+1)b}, \varepsilon\}$. With a left-hand member in $\mathbb{R}_- \cup \{-\infty\}$, we write

$$\frac{1}{n}\ln M_{n_\alpha, n} \leq \frac{\ell_n+1}{n}\frac{1}{\ell_n+1}\sum_{i=0}^{\ell_n} \ln M^{\varepsilon}_{ib,(i+1)b} - \frac{k_n+1}{n}\frac{1}{k_n+1}\sum_{i=0}^{k_n}\ln M^{\varepsilon}_{ib,(i+1)b}.$$

Applying the ergodic theorem, we get

$$\limsup_n \frac{1}{n}\ln M_{n_\alpha, n} \leq \left(\frac{1}{b} - \frac{1-\alpha}{b}\right)E[\ln M^{\varepsilon}_{0,b}] = \frac{\alpha}{b}E[\ln M^{\varepsilon}_{0,b}] \quad \text{a.s.}$$

As this inequality is valid for arbitrary $\varepsilon$ and $b$, (i) follows.

Choose $\alpha = 1$. If $\tau < 1$, we have $\lim_n M_n = 0$ a.s. Since $0 \le M_n \le 1$, this implies $\lim_n E[M_n] = 0$. $\square$

In the context of our study, we immediately deduce the following variation on Lemma 3.2.

COROLLARY 6.1. *Under hypothesis* $(\mathscr{C})$, *if* $0 < \alpha \le 1$ *and if* $n_\alpha$ *is the integral part of* $(1-\alpha)n$, *we have*

$$\limsup_n \frac{1}{n} \ln c(Y_{n_\alpha+1} \cdots Y_n) \le \alpha \ln \kappa \quad \text{a.s.}$$

6.2. *Mean convergence and mixing.* From now on we suppose that $(M_n)_{n\ge1}$ has the following properties:

$\tau < 1$, for all $m, n$, $0 \le m < n$, $M_{m,n}$ is both $\mathscr{F}_n$ and $\mathscr{F}^m$ measurable.

Recall that these $\sigma$-fields have been defined in Section 1.4 and that their definitions have been extended in Section 2.

We just noted that, if $\tau < 1$, the sequence $(E[M_n])_{n\ge1}$ converges to 0. Using the mixing properties, we are now going to estimate the speed of convergence of this sequence. At the end of this section, Corollary 6.2 gathers the properties that will be used later in the proofs of Theorems 3 and 4. Notice that the proofs of Lemmas 6.2 and 6.3 implement arguments that appear in Lemmas 3 and 4 of [6].

Let us recall the integral properties of the mixing coefficients $\alpha_n$ and $\rho_n$:

$(I_\alpha)$ for $p, q, r \in [1, +\infty]$ such that $1/p + 1/q + 1/r = 1$ and for $X \in L^p(\mathscr{F}_k)$, $Y \in L^q(\mathscr{F}^{k+n})$,

$$|E[XY] - E[X]E[Y]| \le 8\alpha_n^{1/r} ||X||_p ||Y||_q.$$

$(I_\rho)$ for $X \in L^2(\mathscr{F}_k)$, $Y \in L^2(\mathscr{F}^{k+n})$,

$$|E[XY] - E[X]E[Y]| \le \rho_n ||X||_2 ||Y||_2.$$

See, for example, [8] or [10] for $(I_\alpha)$.

LEMMA 6.2. *Assume that, for some* $\lambda > 0$ *and* $c$, $\alpha_n \le c/8n^\lambda$. *Then, for every real sequence* $(\ell(n))_n$ *satisfying*

$$\lim_n \frac{\ln n}{\ell(n)} = \lim_n \frac{\ell(n)}{n} = 0,$$

*there exists* $c'$ *such that, for all* $n \ge 1$,

$$E[M_n] \le c' \left( \frac{\ell(n)}{n} \right)^\lambda.$$

PROOF. It follows from submultiplicativity and $0 \le M_n \le 1$ that the sequence $(E[M_n])_{n\ge 1}$ is decreasing. Using, moreover, the stationarity and the inequality $(I_\alpha)$ with $p = q = +\infty$ and $r = 1$, we get, for $u, v, s \ge 1$,

$$\begin{aligned} E[M_{u+s+v}] &\le E[M_u M_{u,\,u+s} M_{u+s,\,u+s+v}] \\ &\le E[M_u M_{u+s,\,u+s+v}] \\ &\le E[M_u]E[M_v] + 8\alpha_s. \end{aligned}$$

Indeed, $M_u$ is $\mathscr{F}_u$-measurable while $M_{u+s,\,u+s+v}$ is $\mathscr{F}^{u+s}$-measurable.

So the sequence with general term $m_n = E[M_n]$ is decreasing and satisfies

$$m_{u+s+v} \le m_u m_v + \frac{c}{s^\lambda}.$$

Let $k \ge 1$, $u = 2ks$ and $v = s$. Then

$$m_{2(k+1)s} \le m_{2ks} m_s + \frac{c}{s^\lambda}.$$

Iterating this leads to

$$m_{2ks} \le (m_s)^{k-1} m_{2s} + \frac{c}{s^\lambda} \sum_{i=0}^{k-2} (m_s)^i.$$

For $n \ge 1$, we denote by $s(n)$ the integer defined by

$$2s(n)\ell(n) \le n < 2(s(n)+1)\ell(n).$$

Since $\lim_n s(n) = +\infty$, we may choose $n_0$ so large that, for $n \ge n_0$, $m_{s(n)} \le e^{-1}$. For $n \ge n_0$, the preceding upper bound for $m_{2ks}$ gives

$$m_n \le m_{2s(n)\ell(n)} \le \exp(-\ell(n)) + \frac{ec}{e-1} \exp\left(-\lambda \ln\left(\frac{n}{2\ell(n)} - 1\right)\right),$$

$$\left(\frac{n}{2\ell(n)}\right)^\lambda m_n \le \exp\left(-\ell(n) + \lambda \ln \frac{n}{2\ell(n)}\right) + ec \exp\left(-\lambda \ln\left(1 - \frac{2\ell(n)}{n}\right)\right).$$

The hypothesis on $(\ell(n))_n$ shows that the right-hand side of the inequality is bounded. □

LEMMA 6.3. *Assume that* $\lim_n \rho_n = 0$. *Then, for every* $k \in \mathbb{N}$, *there exists* $c \in \mathbb{R}$ *such that, for all* $n \ge 1$,

$$E[M_n] \le \frac{c}{n^k}.$$

PROOF. For the same reasons as in the preceding lemma, if $u, s, v \ge 1$, we have

$$\begin{aligned} E[M_{u+s+v}] \le E[M_u M_{u+s,\,u+s+v}] &\le E[M_u]E[M_v] + \rho_s E[M_u^2]^{1/2} E[M_v^2]^{1/2}, \\ &\le E[M_u]E[M_v] + \rho_s E[M_u]^{1/2} E[M_v]^{1/2}, \end{aligned}$$

so that the decreasing sequence $(m_n)_n$, $m_n = E[M_n]$, satisfies

$$m_{2u+s} \le m_u^2 + \rho_s m_u = (m_u + \rho_s) m_u.$$

Choose $n_0$ and $s$ such that $s \le n_0$ and $(m_{n_0} + \rho_s) \le 1/4^k$. For $i \ge 1$ define $n_i$ by $n_i = 2n_{i-1} + s$. We get

$$m_{n_i} \le \frac{1}{4^{ik}} m_{n_0}.$$

Now if $n$ is a large integer, we may write, with a suitable $j$, $n_j \le n < n_{j+1}$. As $n_i \le 4n_{i-1}$, this $j$ satisfies $n \le 4^{j+1} n_0$. Hence

$$m_n \le m_{n_j} \le \exp(-kj \ln 4) m_{n_0} \le 4^k \exp\left(-k \ln \frac{n}{n_0}\right) m_{n_0} = \frac{c}{n^k}. \qquad \square$$

COROLLARY 6.2. *Assume* $(\mathscr{C})$. *Choose* $\eta$, $0 < \eta < 1$, *and set*

$$T_1(\omega) = \inf\{n: c(Y^{(n)}(\omega)) \le \eta\}.$$

*Then we have* $P[T_1 < +\infty] = 1$. *Moreover:*

(i) *if, for some* $\lambda > 0$, $\sum_{n \ge 1} \alpha_n^{1/\lambda} < +\infty$, *then there exists a constant* $K$ *such that, for all* $n \ge 1$,

$$\max\{P[T_1 > n],\ E[c(Y^{(n)})]\} \le K \left(\frac{\ln^2 n}{n}\right)^{\lambda};$$

(ii) *if* $\lim_n \rho_n = 0$, *then there exists a constant* $K$ *such that, for all* $n \ge 1$,

$$\max\{P[T_1 > n], E[c(Y^{(n)})]\} \le K \frac{1}{n^8}.$$

PROOF. Since $(\alpha_n)_{n \ge 1}$ is decreasing and the series associated with $(\alpha_n^{1/\lambda})_{n \ge 1}$ converges, we have $\lim_n n\alpha_n^{1/\lambda} = 0$. It follows that there exists a constant $c$ such that, for all $n \ge 1$, $\alpha_n \le cn^{-\lambda}$. In each case, the hypotheses of Lemma 6.2 or 6.3 are satisfied.

The assertions on $E[c(Y^{(n)})]$ follow from the application of these lemmas to the submultiplicative random sequence $(c(Y^{(n)}))_{n \ge 1}$.

Since $\lim_n c(Y^{(n)}) = 0$ a.s., $T_1 < +\infty$ a.s.

To estimate the tail of $T_1$ law, we only have to notice that, as $(c(Y^{(n)}))_n$ is decreasing,

$$P[T_1 > n] = P[c(Y^{(n)}) > \eta] \le \frac{1}{\eta} E[c(Y^{(n)})]. \qquad \square$$

7. Law of large numbers. The hypotheses of this section are those of Theorem 2.

7.1. *Proof of Theorem* 2. As indicated in Section 5.1, we first reduce the study to the case of an ergodic mean, using the notation of Section 5.1.

LEMMA 7.1. *Assume* $(\mathscr{E})$ *and* $m_1 < +\infty$. *Then the sequences with general terms*

$$\sup\left\{\left|\frac{1}{n}1_{[T\le n]}\ln\langle y, X^{(n)}x\rangle - \frac{1}{n}\sum_{k=1}^{n}\ln\rho(Y_k, Z_{k+1})\right|: x, y\in\overline{B}\right\},$$

$$\left|\frac{1}{n}\ln\Lambda_n - \frac{1}{n}\sum_{k=1}^{n}\ln\rho(Y_k, Z_{k+1})\right|$$

*converge to* 0 *almost surely.*

PROOF. From Lemma 5.1, to obtain the first assertion, it suffices to study the left-hand member of the following inequality, written for $y\in\overline{B}$ with the notation of Section 5.1:

$$\begin{aligned}\left|\ln\|Y^{(n)}y\| - \sum_{k=1}^{n}\ln\rho(Y_k, Z_{k+1})\right| &\le \Delta_n(y)\\ &= \sum_{k=1}^{n}\left|\ln\rho(Y_k, Z^y_{k+1,n}) - \ln\rho(Y_k, Z_{k+1})\right|.\end{aligned}$$

Setting $\Delta_n = \sup\{\Delta_n(y): y\in\overline{B}\}$, we will prove that $\lim_n(1/n)\Delta_n = 0$ a.s.

Pick out $\alpha$, $0<\alpha<1$, and denote by $n_\alpha$ the integral part of $(1-\alpha)n$. Write

$$\Delta_n(y) = \left(\sum_{k=1}^{n_\alpha} + \sum_{k=n_\alpha+1}^{n}\right)\left|\ln\rho(Y_k, Z^y_{k+1,n}) - \ln\rho(Y_k, Z_{k+1})\right| = \Delta_n^1(y) + \Delta_n^2(y)$$

and set, for $i = 1, 2$, $\Delta_n^i = \sup\{\Delta_n^i(y): y\in\overline{B}\}$.

From point (i) of Lemma 5.3, Lemma 3.3 and the submultiplicativity of $c$,

$$\begin{aligned}\Delta_n^1(y) &\le 2\sum_{k=1}^{n_\alpha} r(Y_k)d\big(Z^y_{k+1,n}, Z_{k+1}\big) \le 2\sum_{k=1}^{n_\alpha} r(Y_k)c(Y_{k+1}\cdots Y_n)\\ &\le 2\sum_{k=1}^{n_\alpha} r(Y_k)c(Y_{n_\alpha+1}\cdots Y_n).\end{aligned}$$

Choose $\varepsilon$, $\varepsilon > 0$, such that $(1-\alpha)\varepsilon + \alpha\ln\kappa < 0$. Since $E[\ln r(Y_1)] < +\infty$, the series $\sum_{k\ge1} P[\ln r(Y_k)\ge \varepsilon k]$ converges. By the Borel–Cantelli lemma, $P(\limsup_k[r(Y_k)\ge e^{\varepsilon k}]) = 0$; that is to say, the real random variable $A = \sup_k r(Y_k)e^{-\varepsilon k}$ is a.s. finite. This leads to the estimate

$$\Delta_n^1 \le 2A\left(\sum_{k=1}^{n_\alpha} e^{\varepsilon k}\right)c(Y_{n_\alpha+1}\cdots Y_n) \le \frac{2Ae^{\varepsilon}}{e^{\varepsilon}-1}\, e^{\varepsilon n_\alpha}c(Y_{n_\alpha+1}\cdots Y_n).$$

Using Corollary 6.1, we get

$$\limsup_n \frac{1}{n}\ln\Delta_n^1 \le \varepsilon\lim_n\frac{n_\alpha}{n} + \limsup_n\frac{1}{n}\ln c(Y_{n_\alpha+1}\cdots Y_n) \le (1-\alpha)\varepsilon + \alpha\ln\kappa < 0.$$

Hence

$$\lim_n \Delta_n^1 = 0 \quad \text{a.s.}$$

Moreover, through Lemma 5.2,

$$\Delta_n^2 \le 2 \sum_{n_\alpha+1}^{n} \left|\ln \|Y_k\|\right| + \left|\ln v(Y_k)\right|,$$

and, by means of the ergodic theorem, we get

$$\limsup_n \frac{1}{n}\Delta_n^2 \le 2\alpha m_1 \quad \text{a.s.}$$

As $\alpha$ is arbitrary, we obtain a.s. convergence of the first sequence considered in the statement.

According to the uniformity in $x, y \in \overline{B}$, we still have almost sure convergence to 0, if, in the sequence we just studied, we omit the supremum and replace $y$ by the random vector $L_n$ and $x$ by $\chi$. As $\langle L_n, X^{(n)}\chi\rangle = \langle Y^{(n)}L_n, \chi\rangle = \Lambda_n/q$ and $P[T < +\infty] = 1$, we get the convergence of the second sequence. □

The real random variable $\ln \rho(Y_1, Z_2)$ being integrable, the ergodic theorem shows that

$$\lim_n \frac{1}{n}\sum_{k=1}^{n} \ln \rho(Y_k, Z_{k+1}) = E\big[\ln \rho(Y_1, Z_2)\big] = \gamma \quad \text{a.s.}$$

So we have established Theorem 2 with $\gamma$ instead of $\gamma_1$. Specifying $x$ and $y$ in this preliminary form, we get almost surely, for $i = 1, \dots, q$, $\lim_n (1/n) \ln \|X^{(n)}e_i\| = \gamma$. Hence $\lim_n (1/n) \ln \|X^{(n)}\| = \gamma$ and $\gamma = \gamma_1$.

7.2. *Remarks.*

REMARK 1. Since, for $g \in S$, $v(g) = \inf\{\|ge_i\|: i = 1, \dots, q\}$, we have

$$\lim_n \frac{1}{n} \ln v(X^{(n)}) = \gamma_1 \quad \text{a.s.}$$

REMARK 2. Another proof of the convergence of the sequence $((1/n) \ln \|X^{(n)}x\|)_{n\ge1}$, $x \in \overline{B}$, can also be achieved using some version of Oseledets' theorem adapted to the case of noninvertible matrices. One only has to adapt the arguments contained in the proof of Theorem 1 in [19]. The preceding proof is more elementary.

REMARK 3. It is possible to modify the reasoning leading to $\gamma = \gamma_1$ in such a way that the almost sure convergence

$$\lim_n \frac{1}{n} \ln \|X^{(n)}\| = \gamma_1$$

will no more be an argument but a corollary.

In fact, from $\lim_n (1/n) \ln\langle y, X^{(n)}x\rangle = \gamma$ a.s., we deduce that

$$\lim_n (1/n) \ln ||Y^{(n)}x|| = \gamma \quad \text{a.s.}$$

But, for $g, g' \in S$, $v(g)v(g') \le v(gg') \le ||gg'|| \le ||g||\,||g'||$, so that

$$\frac{1}{n}\sum_{k=1}^{n} \ln v(Y_k) \le \frac{1}{n} \ln ||Y^{(n)}|| \le \frac{1}{n}\sum_{k=1}^{n} \ln ||Y_k||,$$

this double inequality shows that the random variables $(1/n) \ln ||Y^{(n)}||$, $n \ge 1$, are uniformly integrable. We conclude that $\gamma = \lim_n (1/n) E[\ln ||Y_1 \cdots Y_n||] = \gamma_1$.

REMARK 4. As a by-product we get

$$\gamma_1 = E\big[\ln \rho\big(Y_1, Z_2\big)\big] = E\big[\ln ||Y_1 Z_2||\big].$$

7.3. *Proof of Corollary* 2. The second characteristic exponent [23] of the sequence $(X_n)_{n\ge1}$ is the element $\gamma_2 \in \mathbb{R} \cup \{-\infty\}$ which satisfies

$$\gamma_1 + \gamma_2 = \lim_n \frac{1}{n} E\big[\ln || \wedge^2 X^{(n)}||\big].$$

From the subadditive ergodic theorem,

$$\gamma_1 + \gamma_2 = \lim_n \frac{1}{n} \ln || \wedge^2 X^{(n)}|| \quad \text{a.s.},$$

and, since the vectors $e_i \wedge e_j$, $i, j = 1, \dots, q$, $i < j$, form a basis of $\wedge^2 R^q$, we also have

$$\gamma_1 + \gamma_2 = \sup_{i, j=1,\dots,q, i<j} \limsup_n \frac{1}{n} \ln\big|\big|\big(X^{(n)}e_i\big) \wedge \big(X^{(n)}e_j\big)\big|\big| \quad \text{a.s.}$$

For $x \in \mathbb{R}^q$, set $||x||' = \langle x, x\rangle^{1/2}$. The scalar product and the Euclidean norm extend to $\wedge^2 \mathbb{R}^q$. The angular distance of $x, y \in \overline{B}$ is defined by

$$d_a(x, y) = \frac{||x \wedge y||'}{||x||'||y||'}.$$

Let us compare $d_a$ to the distance $d$ which has been used to define the contraction coefficient of the sequence $(X_n)_{n\ge1}$.

LEMMA 7.3. *For all $x, y \in \overline{B}$, $d_a(x, y) \le 2\sqrt{q}\,d(x, y)$.*

PROOF. We have

$$||x \wedge y||' = ||x \wedge (y - x)||' \le ||x||'||y - x||' \le ||x||'||y - x||'(\sqrt{q})||y||'),$$

because, by convexity, for $y \in \overline{B}$, $\sqrt{q}||y||' \ge 1$. So we obtain

$$d_a(x, y) \le \sqrt{q}||y - x||' \le \sqrt{q}||y - x|| \le 2\sqrt{q}\,d(x, y). \qquad \square$$

For all $x, y \in \overline{B}$, we therefore have

$$d_a(X^{(n)} \cdot x, X^{(n)} \cdot y) \le 2\sqrt{q}d(X^{(n)} \cdot x, X^{(n)} \cdot y) \le 2\sqrt{q}c(X^{(n)}).$$

Hence

$$\ln ||(X^{(n)}x) \wedge (X^{(n)}y)||' - \ln ||X^{(n)}x||' - \ln ||X^{(n)}y||' \le \ln 2\sqrt{q} + \ln c(X^{(n)}).$$

From the norm equivalence, it follows that

$$\limsup_n \frac{1}{n} \ln ||(X^{(n)}x) \wedge (X^{(n)}y)|| \le 2\gamma_1 + \ln \kappa \quad \text{a.s.}$$

Finally, using basis vectors,

$$\gamma_1 + \gamma_2 \le 2\gamma_1 + \ln \kappa \quad \text{a.s.} \qquad \square$$

## 8. Central limit theorem.

8.1. *Proof of Theorem* 3. We first show how this proof may be reduced to the case of a stationary sequence. The following statement is similar to Lemma 7.1, it differs by the normalization and the type of convergence involved.

LEMMA 8.1. *Assume* $(\mathscr{C})$ *and* $m_1 < +\infty$. *Then the sequences with general terms*

$$\sup\left\{\left|\frac{1}{\sqrt{n}}1_{[T\le n]} \ln\langle y, X^{(n)}x\rangle - \frac{1}{\sqrt{n}}\sum_{k=1}^{n} \ln \rho(Y_k, Z_{k+1})\right| : x, y \in \overline{B}\right\},$$

$$\left|\frac{1}{\sqrt{n}} \ln \Lambda_n - \frac{1}{\sqrt{n}}\sum_{k=1}^{n} \ln \rho(Y_k, Z_{k+1})\right|$$

*converge to* 0 *in probability.*

PROOF. To establish the first assertion, we use the notation of the proof of Lemma 7.1. From Lemma 5.1, it suffices to establish that the sequence $((1/\sqrt{n})\Delta_n)_{n\ge 1}$ converges to 0 in probability.

As seen in the course of the proof of Lemma 7.1, we may write

$$\frac{1}{\sqrt{n}}\Delta_n \le \frac{2}{\sqrt{n}}\sum_{k=1}^{n} r(Y_k)c(Y_{k+1}\cdots Y_n) = \frac{2}{\sqrt{n}}A_n,$$

so that the proof will be over if we show that $(A_n)_{n\ge 1}$ converges weakly. The random variable $A_n$ has the same law as

$$A'_n = A_n \circ \theta^{-n-1} = \sum_{\ell=1}^{n} r(Y_{-\ell})c(Y_{-\ell+1}\cdots Y_{-1}).$$

This is the partial sum of a series whose general term satisfies

$$\limsup_{\ell} \frac{1}{\ell} \ln\bigl[r(Y_{-\ell})c(Y_{-\ell+1}\cdots Y_{-1})\bigr] \le \limsup_{\ell} \frac{1}{\ell} \ln r(Y_{-\ell}) + \limsup_{\ell} \frac{1}{\ell} \ln c(Y_{-\ell+1}\cdots Y_{-1}).$$

Since the random variables $\ln r(Y_{-\ell})$, $\ell \ge 1$, are identically distributed and integrable, the first term on the right-hand side is equal to 0. Owing to the properties of $c$ and the subadditive ergodic theorem, the second term may be written as

$$\lim_{\ell} \frac{1}{\ell} \ln c(Y_{-\ell+1}\cdots Y_{-1}) = \inf_{\ell} \frac{1}{\ell} E[\ln c(Y_{-\ell+1}\cdots Y_{-1})] = \inf_{\ell} \frac{1}{\ell} E[\ln c(Y^{(\ell)})] = \ln \kappa < 0,$$

so that the considered series converges almost surely. We conclude that $(A'_n)_{n\ge 1}$ and hence $(A_n)_{n\ge 1}$ converges weakly. As in Lemma 7.1, the second assertion follows from the first one. □

Set

$$U_n = \ln \rho(Y_n, Z_{n+1}) - \gamma_1, \qquad n \ge 0.$$

The preceding lemma shows that Theorem 3 will be established when the weak convergence to a normal law of the sequence $((1/\sqrt{n})\sum_{k=1}^n U_k)_{n\ge 1}$ is proved. For this purpose, we employ Gordin's method [16]. It is convenient to use this method through the following statement.

LEMMA 8.2. *Let $p \ge 2$ be such that $m_p < +\infty$ and let $q$ be defined by $1/p + 1/q = 1$.*
*Assume*

$$\sum_{n\ge 1} ||E[U_0|\mathscr{F}^n]||_q < +\infty.$$

*Then the sequence $((1/\sqrt{n})\sum_{k=1}^n U_k)_n$ converges weakly to a centered normal law with variance $\sigma^2$. Moreover, if $\sigma^2 = 0$ there exists a stationary sequence $(W_n)_n$, such that, for each $n$,*

$$W_n \in L^q(\mathscr{F}^n), \qquad U_n = W_{n+1} - W_n.$$

This lemma is proved in [9]. It is also easily deduced from Liverani [25], Theorem 1.1. The case considered here allows a short proof based on the same argument.

PROOF OF LEMMA 8.2.

$$\sum_{j\ge 0} \big|\big| \, |E[U_{-j}|\mathscr{F}^0]| \, \big|\big|_1 \le \sum_{j\ge 0} ||E[U_{-j}|\mathscr{F}^0]||_q = \sum_{j\ge 0} ||E[U_0|\mathscr{F}^j]||_q < +\infty,$$

so that $\sum_{j\geq 0} E[U_{-j}|\mathscr{F}^0]$ converges in $L^q$ and also absolutely on a set of probability one. In the sense of these convergences, we set

$$V_0 = \sum_{j\geq 0}\big(E[U_{-j}|\mathscr{F}^0] - E[U_{-j}|\mathscr{F}^1]\big), \qquad W_0 = \sum_{j\geq 1} E[U_{-j}|\mathscr{F}^0],$$

and, for $n \in \mathbb{Z}$, $V_n = V_0 \circ \theta^n$ and $W_n = W_0 \circ \theta^n$. It is clear that

$$V_n = U_n + W_n - W_{n+1}, \qquad E[V_n|\mathscr{F}^{n+1}] = 0.$$

Now, to get the result from the Ibragimov–Billingsley theorem (cf. [10]), it is sufficient to prove that $V_0$ or, equivalently, $W_1 - W_0$ is in $L^2$. For this purpose, we introduce the parameter $\lambda$, $0 < \lambda < 1$, and we set, in the sense of $L^2$ convergence,

$$W_0^\lambda = \sum_{j\geq 1} \lambda^{j-1} E[U_{-j}|\mathscr{F}^0].$$

A standard result in the theory of power series shows that we have $\lim_{\lambda\to 1-} W_0^\lambda = W_0$ a.s. Therefore, if it is proved that

$$\sup\big\{||W_1^\lambda - W_0^\lambda||_2 \colon 0 < \lambda < 1\big\} < +\infty,$$

Fatou's lemma leads to the required result. The following computation based on the stationarity gives the desired boundedness. From the equality $\lambda W_0^\lambda = (\lambda W_0^\lambda - W_1^\lambda) + W_1^\lambda$, we get

$$\begin{aligned}||\lambda W_0^\lambda - W_1^\lambda||_2^2 &= 2\langle W_1^\lambda - \lambda W_0^\lambda, W_1^\lambda\rangle - (1-\lambda^2)||W_0^\lambda||_2^2 \\ &\leq 2\langle W_1^\lambda - \lambda E[W_0^\lambda|\mathscr{F}^1], W_1^\lambda\rangle.\end{aligned}$$

But

$$W_1^\lambda - \lambda E[W_0^\lambda|\mathscr{F}^1] = \sum_{j\geq 1}\lambda^{j-1}E[U_{-j+1}|\mathscr{F}^1] - \lambda\sum_{j\geq 1}\lambda^{j-1}E[U_{-j}|\mathscr{F}^1] = E[U_0|\mathscr{F}^1],$$

so

$$\begin{aligned}||\lambda W_0^\lambda - W_1^\lambda||_2^2 &\leq 2\sum_{j\geq 1}\lambda^{j-1}\langle E[U_0|\mathscr{F}^1], E[U_{-j+1}|\mathscr{F}^1]\rangle \\ &\leq 2||U_0||_p \sum_{j\geq 1}||E[U_{-j}|\mathscr{F}^0]||_q. \end{aligned}$$ □

LEMMA 8.3. *For $\eta$, $0 < \eta < 1$, set $T_1 = \inf\{n \colon n \geq 1,\ c(Y^{(n)}) \leq \eta\}$. Then there exists $\eta_1$ such that, if, denoting by $n_2$ the integral part of $n/2$, we set*

$$\begin{aligned}a_n(2+\delta) &= 8(4m_{2+\delta}+\eta_1)\alpha_{n-n_2}^{\delta/(2+\delta)} + 4m_{2+\delta}P[T_1 > n_2]^{(1+\delta)/(2+\delta)} \\ &\quad + 2\eta_1 E[c(Y^{(n_2)})], \qquad \delta > 0,\end{aligned}$$

$$a_n(2) = (4m_2+\eta_1)\rho_{n-n_2} + 4m_2 P[T_1 > n_2]^{1/2} + 2\eta_1 E[c(Y^{(n_2)})],$$

*we have, according to the conditions $m_{2+\delta} < +\infty$ or $m_2 < +\infty$,*

$$||E[U_{-n}|\mathscr{F}^0]||_{(2+\delta)/(1+\delta)} \leq a_n(2+\delta), \qquad ||E[U_{-n}|\mathscr{F}^0]||_2 \leq a_n(2).$$

PROOF. Let $(p,q)=(2+\delta,(2+\delta)/(1+\delta))$ or $(p,q)=(2,2)$.

The method is to get an upper bound for the quantity $|E[U_0G_n]|$, where $G_n$ is an arbitrary $\mathscr{F}^n$-measurable random variable in $L^p$. This is done using terms related to the contraction properties and the mixing coefficients.

Recalling that $Y^{(n_2)}\cdot Z_{n_2+1}=Z_1$ and that $\gamma_1=E[\ln\rho(Y_0,Z_1)]$, we may write

$$\begin{aligned}U_0&=\ln\rho(Y_0,Z_1)-\gamma_1\\&=\ln\rho(Y_0,Y^{(n_2)}\cdot Z_{n_2+1})-\ln\rho(Y_0,Y^{(n_2)}\cdot\chi)\\&\quad+\ln\rho(Y_0,Y^{(n_2)}\cdot\chi)-E[\ln\rho(Y_0,Y^{(n_2)}\cdot\chi)]\\&\quad+E[\ln\rho(Y_0,Y^{(n_2)}\cdot\chi)]-E[\ln\rho(Y_0,Y^{(n_2)}\cdot Z_{n_2+1})],\end{aligned}$$

or, equivalently,

$$\begin{aligned}U_0&=A_n+B_n-E[A_n],\\A_n&=\ln\rho(Y_0,Y^{(n_2)}\cdot Z_{n_2+1})-\ln\rho(Y_0,Y^{(n_2)}\cdot\chi),\\B_n&=\ln\rho(Y_0,Y^{(n_2)}\cdot\chi)-E[\ln\rho(Y_0,Y^{(n_2)}\cdot\chi)].\end{aligned}$$

To bound $A_n$, we use the stopping time $T_1$. On the event $[T_1\le n_2]$, we have $c(Y^{(n_2)})\le\eta$. Therefore, by point (iii) of Lemma 5.3, there exists $\eta_1$ such that $|A_n|\le\eta_1c(Y^{(n_2)})$, and hence, by Lemma 5.2,

$$|A_n|\le A_n'=2\big(|\ln\|Y_0\||+|\ln v(Y_0)|\big)1_{[T_1>n_2]}+\eta_1c(Y^{(n_2)}).$$

For the sequel notice the straightforward inequalities

$$E[A_n']\le 2m_pP[T_1>n_2]^{(p-1)/p}+\eta_1E[c(Y^{(n_2)})],$$

$$\|A_n'\|_p\le 2m_p+\eta_1,\qquad \|B_n\|_p\le 2m_p.$$

As $E[B_n]=0$, we have

$$\begin{aligned}|E[U_0G_n]|&\le|E[A_nG_n]|+|E[B_nG_n]|+|E[A_n]E[G_n]|\\&\le E[A_n'|G_n|]+|E[B_nG_n]|+E[A_n']E[|G_n|]\\&\le\big|E[A_n'|G_n|]-E[A_n']E[|G_n|]\big|+\big|E[B_nG_n]-E[B_n]E[G_n]\big|\\&\quad+2E[A_n']E[|G_n|].\end{aligned}$$

Noting that $G_n$ is $\mathscr{F}^n$-measurable, while $A_n'$ and $B_n$ are $\mathscr{F}_{n_2}$-measurable, we can bound above the first two terms of the last member by means of the mixing inequalities.

Assume $p=2+\delta$. Using inequality $(I_\alpha)$ with $p=q=2+\delta$ and $r=(2+\delta)/\delta$, we get

$$\begin{aligned}|E[U_0G_n]|&\le 8\alpha_{n-n_2}^{\delta/(2+\delta)}\big(\|A_n'\|_{2+\delta}+\|B_n\|_{2+\delta}\big)\|G_n\|_{2+\delta}+2E[A_n']\|G_n\|_{2+\delta}\\&\le a_n(2+\delta)\|G_n\|_{2+\delta}.\end{aligned}$$

Assume $p = 2$. Using inequality $(I_\rho)$, we get

$$|E[U_0 G_n]| \le \rho_{n-n_2}\big(||A'_n||_2 + ||B_n||_2\big)||G_n||_2 + 2E[A'_n]||G_n||_2$$
$$\le a_n(2)||G_n||_2. \qquad \square$$

To conclude, let us assume that the hypotheses of Theorem 3 are satisfied. From Corollary 6.2, if $p = 2+\delta$ and $\sum_{n\ge 1} \alpha_n^{\delta/(2+\delta)} < +\infty$, then

$$E[c(Y^{(n)})] \le K'\left(\frac{\ln^2 n}{n}\right)^{(2+\delta)/\delta}, \qquad P[T_1 > n]^{(1+\delta)/(2+\delta)} \le K''\left(\frac{\ln^2 n}{n}\right)^{(1+\delta)/\delta},$$

while, if $p = 2$ and $\sum_{n\ge 1} \rho_n < +\infty$, then

$$E[c(Y^{(n)})] \le \frac{K'}{n^8}, \qquad P[T_1 > n]^{1/2} \le \frac{K''}{n^4}.$$

In each case, the series $\sum_{n\ge 1} a_n(p)$ converges. $\square$

8.2. *Proof of Corollary* 3.

LEMMA 8.5. *Under hypothesis* $(\mathscr{C})$, *if there exists a stationary sequence* $(W_n)_{n\in\mathbb{Z}}$ *such that, for* $n \in \mathbb{Z}$,

$$U_n = \ln \rho(Y_n, Z_{n+1}) - \gamma_1 = W_{n+1} - W_n,$$

*then the sequence* $(e^{-n\gamma_1}||X^{(n)}||)_{n\ge 1}$ *is tight in* $]0, +\infty[$.

PROOF. Let $\varepsilon > 0$. Choose $c$ and $c' > 0$ such that

$$P[|W_0| > c] \le \varepsilon/4, \qquad P\Big[\inf_{i=1,\dots,q} \langle Z_0, e_i\rangle < c'\Big] \le \varepsilon/2.$$

The last choice is possible since $P[Z_0 \in B] = 1$.

Set $D_n = [|W_n - W_0| \le 2c] \cap [\inf_{i=1,\dots,q}\langle Z_{n+1}, e_i\rangle \ge c']$.

From $[|W_n - W_0| > 2c] \subset [|W_n| > c] \cup [|W_0| > c]$ and the stationarity of the sequences $(W_n)_{n\in\mathbb{Z}}$, $(Z_n)_{n\in\mathbb{Z}}$, it follows that $P(D_n) \ge 1-\varepsilon$.

On $D_n$,

$$\exp(-2c) \le \exp(W_n - W_0) = \exp(-n\gamma_1)||Y^{(n)}Z_{n+1}|| \le \exp(2c)$$

and

$$c'q||Y^{(n)}\chi|| \le ||Y^{(n)}Z_{n+1}|| \le q||Y^{(n)}\chi||.$$

Hence

$$\frac{e^{-2c}}{q}e^{n\gamma_1} \le ||Y^{(n)}\chi|| = \frac{1}{q}\sum_{i,\,j=1}^{q}\langle Y^{(n)}e_i, e_j\rangle \le \frac{e^{2c}}{c'q}e^{n\gamma_1}.$$

On the finite-dimensional vector space of $q\times q$ matrices, the norms defined by

$$||h|| = \sup\big\{||hx||\colon ||x|| = 1\big\}, \qquad ||h||' = \sum_{i,\,j=1}^{q} |\langle he_i, e_j\rangle|$$

are equivalent. Hence there exist constants $a', b'$, $0 < a' \le b'$, so that, on $D_n$,

$$a' e^{n\gamma_1} \le ||X^{(n)}|| \le b' e^{n\gamma_1}.$$

Finally, for all $n$, $P[a' e^{n\gamma_1} \le ||X^{(n)}|| \le b' e^{n\gamma_1}] \ge 1 - \varepsilon$. □

## 9. Tightness.

### 9.1. *Proof of Theorem* 4.

PROOF OF (i). Recall first that, as noted before Lemma 4.1,

$$\lim_n \frac{\jmath(R_n \otimes L_n)}{\langle e_i, (R_n \otimes L_n) e_j \rangle} \frac{\langle e_i, X^{(n)} e_j \rangle}{\jmath(X^{(n)})} 1_{[T \le n]} = 1 \quad \text{a.s.},$$

and second that, from Theorem 1(iii), continuity, homogeneity of function $\jmath$ and $\nu \times \nu' \{g\colon g \in S, \jmath(g) > 0\} = 1$,

$$\left( \frac{R_n \otimes L_n}{\jmath(R_n \otimes L_n)} 1_{[\jmath(R_n \otimes L_n) > 0]} \right)_n$$

converges weakly to the probability $h_2(\nu' \times \nu)$ on $S^\circ$, where $h_2 = \jmath(h)^{-1} h$.

Joining these to the fact that $\big(\jmath(X^{(n)})\big)_{n \ge 1}$ is tight, we deduce that, for each $\varepsilon > 0$, there exist $a > 1$ and $n_0$ such that, for $n \ge n_0$, the events

$$A_1^n = \left( \bigcap_{i,j} \left[ a^{-1} \frac{\langle e_i, (R_n \otimes L_n) e_j \rangle}{\jmath(R_n \otimes L_n)} \le \frac{\langle e_i, X^{(n)} e_j \rangle}{\jmath(X^{(n)})} \le a \frac{\langle e_i, (R_n \otimes L_n) e_j \rangle}{\jmath(R_n \otimes L_n)} \right] \right)$$
$$\cap [T \le n],$$
$$A_2^n = \left( \bigcap_{i,j} \left[ a^{-1} \le \frac{\langle e_i, (R_n \otimes L_n) e_j \rangle}{\jmath(R_n \otimes L_n)} \le a \right] \right) \cap [\jmath(R_n \otimes L_n) > 0],$$
$$A_3^n = \big[ \jmath(X^{(n)}) \le a \big],$$

satisfy $P(A_i^n) \ge 1 - \varepsilon/3$, so that $P(A_1^n \cap A_2^n \cap A_3^n) \ge 1 - \varepsilon$. But $A_1^n \cap A_2^n \cap A_3^n \subset [X^{(n)} \in K]$, where $K$ is the compact subset of $\overline{S}$ defined by

$$K = \big\{ g\colon g \in \overline{S},\ \forall i, j,\ \langle e_i, g e_j \rangle \le a^3 \big\}.$$

It follows that the sequence $(\mu_n)_n$, $\mu_n$ being the law of $X^{(n)}$, is tight in $\overline{S}$. By Prohorov's theorem, it is weakly conditionally compact.

Let $\tilde{\mu}$ and $(n_k)_k$ be such that $\lim_k \mu_{n_k} = \tilde{\mu}$. We now prove that $\tilde{\mu}(S^\circ \cup \{0\}) = 1$. Fix $c > 0$. For $b > 0$ and $i, j = 1, \ldots, q$, set

$$D_{ij}^b = \big\{ g\colon g \in \overline{S},\ ||g|| > c,\ \langle e_i, g e_j \rangle < b \big\},$$
$$D_{ij} = \big\{ g\colon g \in \overline{S},\ ||g|| > c,\ \langle e_i, g e_j \rangle = 0 \big\}.$$

On $(A_1^n \cap A_2^n) \cap [X^{(n)} \in D_{ij}^b]$, we have

$$a^{-2} \le \min_{(i',j')} \frac{\langle e_{i'}, X^{(n)} e_{j'} \rangle}{\jmath(X^{(n)})} \le \max_{(i',j')} \frac{\langle e_{i'}, X^{(n)} e_{j'} \rangle}{\jmath(X^{(n)})} \le a^2, \qquad \langle e_i, X^{(n)} e_j \rangle < b,$$

and we get successively

$$\jmath(X^{(n)}) \le a^2 b, \qquad \max_{(i',j')} \langle e_{i'}, X^{(n)} e_{j'} \rangle \le a^4 b, \qquad ||X^{(n)}|| \le q a^4 b.$$

If $b$ is chosen such that $qa^4 b < c$, $(A_1^n \cap A_2^n) \cap [X^{(n)} \in D_{ij}^b] = \varnothing$, so that, for $n \ge n_0$,

$$P\big[X^{(n)} \in D_{ij}^b\big] = P\big([X^{(n)} \in D_{ij}^b] \cap \big(A_1^n \cap A_2^n\big)^c\big) \le P\big(A_1^n \cap A_2^n\big)^c \le \frac{2\varepsilon}{3}.$$

It follows that

$$\tilde{\mu}(D_{ij}) \le \tilde{\mu}(D_{ij}^b) \le \liminf_n P\big[X^{(n)} \in D_{ij}^b\big] \le \frac{2\varepsilon}{3}.$$

Since $\varepsilon$ is arbitrary, we get $\tilde{\mu}(D_{ij}) = 0$. At last, letting $c$ and $(i, j)$ vary, we get

$$\tilde{\mu}\big(\overline{S} \backslash S^{\circ} \cup \{0\}\big) = 0.$$

If $\tilde{\mu}(\{0\}) = 0$, the function $h_2$ is defined and continuous except on a set of $\tilde{\mu}$ probability 0, so that $\lim_k h_2(\mu_{n_k}) = h_2(\tilde{\mu})$. But the sequence $(h_2(\mu_n))_n$ has the limit $h_2(\nu' \times \nu)$ and hence $1 = h_2(\tilde{\mu})(Q) = \tilde{\mu}(Q)$. □

PROOF OF (ii). In this section we suppose that $\lim_n \alpha_n = 0$, where $\alpha_n$ is the mixing coefficient defined in Section 1.4.

LEMMA 9.1. *For each $\varepsilon > 0$, there exist $s_0 \in \mathbb{N}$ and $\eta_0 > 0$ such that, for all $s \ge s_0$, $\eta \le \eta_0$ and $n \ge 1$,*

$$P\big[||X^{(2n+s)}|| \le \eta\big] \ge 2P\big[||X^{(n)}|| \le \eta^3\big] - \big(P[||X^{(n)}|| \le \eta^3]\big)^2 - \varepsilon.$$

PROOF. Let $\eta$, $0 < \eta < 1$. Setting

$$\begin{aligned} A &= \big[||X^{(n)}|| \le \eta^3\big] \cap \big[||X^{(n)}|| \circ \theta^{n+s} \le \eta^{-1}\big], \\ B &= \big[||X^{(n)}|| \le \eta^{-1}\big] \cap \big[||X^{(n)}|| \circ \theta^{n+s} \le \eta^3\big], \\ C &= \big[||X^{(s)}|| \circ \theta^n \le \eta^{-1}\big], \end{aligned}$$

we have

$$\big[||X^{(2n+s)}|| \le \eta\big] \supset (A \cap C) \cup (B \cap C).$$

As

$$P((A \cup B) \cap C) \ge P(A \cup B) - P(C^c) = P(A) + P(B) - P(A \cap B) - P(C^c),$$

we get

$$\begin{aligned}
P\big[\|X^{(n+s+n)}\| \le \eta\big] \ge\ & P\big(\big[\|X^{(n)}\| \le \eta^3\big] \cap \big[\|X^{(n)}\| \circ \theta^{n+s} \le \eta^{-1}\big]\big) \\
&+ P\big(\big[\|X^{(n)}\| \le \eta^{-1}\big] \cap \big[\|X^{(n)}\| \circ \theta^{n+s} \le \eta^3\big]\big) \\
&- P\big[\|X^{(n)}\| \le \eta^3\big] \cap \big[\|X^{(n)}\| \circ \theta^{n+s} \le \eta^3\big] \\
&- P\big(\big[\|X^{(s)}\| > \eta^{-1}\big]\big) \\
\ge\ & 2P\big[\|X^{(n)}\| \le \eta^3\big]P\big[\|X^{(n)}\| \le \eta^{-1}\big] - \big(P[\|X^{(n)}\| \le \eta^3]\big)^2 \\
&- 3\alpha_s - P\big[\|X^{(s)}\| > \eta^{-1}\big].
\end{aligned}$$

As $(\|X^{(k)}\|)_{k\ge 1}$ is tight in $\mathbb{R}_+$, for all $\varepsilon > 0$, there exists $\eta_0$ such that, for $\eta \le \eta_0$,

$$\sup_k P\big[\|X^{(k)}\| > \eta^{-1}\big] \le \varepsilon/5.$$

To conclude, we just choose $s_0$ such that, for $s \ge s_0$, $3\alpha_s \le \varepsilon/5$. □

Suppose $(\mu_n)_n$ is tight in $\overline{S}$ and there exist a probability $\tilde{\mu}$ on $\overline{S}$ and a subsequence $(n_k)_k$ such that $0 < \tilde{\mu}\{0\} < 1$ and $\lim_k \mu_{n_k} = \tilde{\mu}$. We apply Lemma 9.1 with $\varepsilon \le (1/2)\tilde{\mu}\{0\}(1 - \tilde{\mu}\{0\})$. From the sequence $(2n_k + s_0)_k$, we may extract a subsequence $(n'_k)_k$ such that $\lim_k \mu_{n'_k} = \tilde{\mu}'$. If $\eta$ is chosen such that $\tilde{\mu}\{g\colon g \in \overline{S},\ \|g\| = \eta^3\} = 0$, we have

$$\tilde{\mu}'\{g\colon \|g\| \le \eta\} \ge 2\,\tilde{\mu}\{g\colon \|g\| \le \eta^3\} - (\tilde{\mu}\{g\colon \|g\| \le \eta^3\})^2 - \tfrac{1}{2}\tilde{\mu}\{0\}(1 - \tilde{\mu}\{0\}).$$

Hence

$$\tilde{\mu}'\{0\} \ge \tfrac{1}{2}\tilde{\mu}\{0\}(3 - \tilde{\mu}\{0\}).$$

The sequence $(u_n)_n$ defined by $u_0 \in [0, 1]$ and $u_{n+1} = \frac{1}{2}u_n(3 - u_n)$ converges to 1. It follows that, for all $\varepsilon > 0$, it is possible to construct a probability $\tilde{\mu}$ on $\overline{S}$ and a sequence $(n_k)_k$ such that $\tilde{\mu}\{0\} > 1 - \varepsilon$ and $\lim_k \mu_{n_k} = \tilde{\mu}$. Under the conditions of Lemma 9.1, if $\eta$ is chosen outside a suitable countable subset, we get

$$\begin{aligned}
\liminf_k \inf_{s\ge s_0} P[\|X^{(2n_k+s)}\| \le \eta] &\ge 2\,\tilde{\mu}\{g\colon \|g\| \le \eta^3\} - (\tilde{\mu}\{g\colon \|g\| \le \eta^3\})^2 - \varepsilon \\
&\ge \tilde{\mu}\{0\}(2 - \tilde{\mu}\{0\}) - \varepsilon \ge 1 - 2\varepsilon.
\end{aligned}$$

This means that $(\|X^{(n)}\|)_n$ converges in probability to 0, or else that $(\mu_n)_n$ converges weakly to the probability carried by $\{0\}$.

Finally, if $(\mu_n)_n$ is tight in $\overline{S}$ and does not converge to the unit mass at 0, all its weak limit values $\tilde{\mu}$ satisfy $\tilde{\mu}\{0\} = 0$. This joined to (i) implies $\tilde{\mu}(Q) = 1$.

9.2. *Case of independent random matrices.* In this section, the random matrices $X_n$, $n \geq 1$, are independent, identically distributed according to the probability measure $\mu$ on $S$. We denote by $T$ the closed subsemigroup of $S$ spanned by the support of $\mu$ and the identity matrix.

If $\pi, \pi'$ are probabilities on the multiplicative semigroup $S$, $\pi * \pi'$ denote their convolution product while $\pi^{*n}$ is the $n$th convolution power of $\pi$.

REMARKS. In this context the contraction condition may be written as

$(\mathscr{C})$ there exists $n \geq 1$ such that $\int c(g)\, d\mu^{*n}(g) < 1$.

Otherwise, let $\mathscr{P}_q$ be the collection of all nonempty subsets of $\{1, \dots, q\}$ and, for $I \in \mathscr{P}_q$, set $\xi_I = \sum_{i \in I} e_i$. For $I$ and $J$ in $\mathscr{P}_q$, we write $I \overset{\mu}{\to} J$ if there exist $n$, $g$ in the support of $\mu^{*n}$ and $\lambda > 0$ such that $g\xi_I - \lambda \xi_J \in \overline{C}$. According to positivity and independence, the relation $\overset{\mu}{\to}$ is transitive, so that it is determined by the connections established through the elements of the support of $\mu$. It is clear that

$$(\mathscr{C}) \text{ is equivalent to the irreducibility of the graph } \big(\mathscr{P}_q, \overset{\mu}{\to}\big).$$

Notice that $(\mathscr{C})$ cannot be deduced from the condition $E[X_1] \in S^\circ$ as may be seen when $q = 2$ and $\mu = \frac{1}{2}(\delta_a + \delta_b)$, $a = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, $b = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$.

*Conjugate affine subspaces and structure of elements of $S_A$.* Let $A \in \mathscr{K}^+$. To describe $S_A$, it is convenient to introduce the conjugate subspace $A'$ of $A$:

$$A' = \{a' \colon a' \in \mathbb{R}^q,\ \forall a \in A,\ \langle a, a' \rangle = 1\}.$$

Denote by $a_0 \in C$ the orthogonal projection of 0 on $A$. As each $a' \in A'$ satisfies $\langle a_0, a' \rangle = 1$, $A' \cap C$ is bounded. It is nonempty since it contains $a_0' = a_0 / \langle a_0, a_0 \rangle$ and hence $A' \in \mathscr{K}^+$. Notice that the relations $g \in S_A$ and $g^* \in S_{A'}$ are equivalent.

From these remarks we may deduce a dual form of Theorem 5 in which the tightness condition is the existence of $A^* \in \mathscr{K}^+$ such that $T^* \subset S_{A^*}$, $A^*$ being now described as the affine subset spanned by $T^* \ell_m$, where $\ell_m \in C$ satisfies $E[X_1]^* \ell_m = \ell_m$.

Using norm equivalence, it is seen that the condition $\langle a_0', g a_0 \rangle = 1$ implies that there exists a constant $c$, such that, for each $g \in S_A$, $\|g\| \leq c$.

Let $g \in S_A$. Consider the sequence $(r_n)_n$ of the compact subset $A \cap \overline{C}$ defined by $r_n = (1/n) \sum_{k=0}^{n-1} g^k a_0$. If $r$ is a limit value of this sequence, then $r \in A \cap \overline{C} \not\ni 0$ and $gr = r$. From this and the fact that $(\|g^n\|)_n$ is bounded, we conclude that $g$ has spectral radius $\lambda(g) = 1$.

We define $r(g)$ as the unique element of $K_g = \{r \colon r \in A \cap \overline{C},\ gr = r\}$ such that $\langle r(g), r(g) \rangle = \min\{\langle r, r \rangle \colon r \in K_g\}$. The function $r(\cdot)$ is then measurable. Note that $l(\cdot)$ is constructed similarly, replacing $A$ by $A'$ and $g$ by $g^*$. At last we set $w(g) = g - r(g) \otimes l(g)$.

As $\langle l(g), r(g) \rangle = 1$, we may write $\mathbb{R}^q = \operatorname{span}(r(g)) \oplus l(g)^\perp$; these two subsets are preserved by $g$. It is clear that $w(g) r(g) = 0$, while if $x \in l(g)^\perp$,

$w(g)x = gx$. This implies that $\lambda(w(g)) \le \lambda(g) = 1$ and $\lambda(w(g)) < 1$ if $g \in S^\circ$. For $a \in A$, $w(g)a = ga - \langle l(g), a\rangle r(g) = ga - r(g) \in A - A$. $\square$

PROOF OF THEOREM 5. The "if" part is readily obtained from the preceding discussion.

The converse will be proved through the two following lemmas.

LEMMA 9.2. *Set $\bar\mu_n = (1/n)\sum_{k=1}^n \mu^{*k}$, assume that $(\bar\mu_n)_n$ is tight in $S$ and denote by $\pi$ a weak limit value of this sequence. Then, if $T$ is the closed subsemigroup of $S$ spanned by the support of $\mu$ and the identity matrix,*

$$\pi(T) = 1, \qquad \mu * \pi = \pi * \mu = \pi$$

*and*

$$\text{for all } g \in T,\ \pi * \delta_g * \pi = \pi.$$

PROOF. As $T$ is closed in $S$ and $\mu^{*n}(T) = 1$, $\pi(T) = 1$.

By the relation $\mu * \bar\mu_n = \bar\mu_n - (1/n)\mu + (1/n)\mu^{*(n+1)}$, we have $\mu * \pi = \pi$. Operating in the same way but on the left-hand side, we get $\pi * \mu = \pi$. One step further gives $\pi * \pi = \pi$, which is the second relation for $g = e$.

To cope with the general case, we use a technique due to Raugi (cf. [3], [17] and [29]) based on a martingale argument introduced in [14]. Set $M_n = X_1 \cdots X_n$. To any real continuous bounded function $f$ on $S$, we associate $F$ on $S$ defined by $F(g) = \int f(gk)\, d\pi(k)$. It is easily verified, using the independence and the invariance $\mu * \pi = \pi$, that $(F(M_n))_n$ is a martingale with respect to the filtration $(\mathscr{F}_n)_n$. Fix $j \ge 1$ and consider the sequence

$$u_n = \int E\big[\big(F(M_n g) - F(M_n)\big)^2\big]\, d\mu^{*j}(g).$$

By the martingale property we have

$$u_n = E\Big[\big(F(M_{n+j}) - F(M_n)\big)^2\Big] = E\big[F(M_{n+j})^2\big] - E\big[F(M_n)^2\big].$$

Using independence, we deduce

$$0 \le \sum_{n\ge1} u_n = \int E\bigg[\sum_{n\ge1}\big(F(M_n g) - F(M_n)\big)^2\bigg]\, d\mu^{*j}(g) \le 2\sup_h |F(h)| < +\infty.$$

It follows that $\lim_n(F(M_n g) - F(M_n)) = 0$ $P \times \mu^{*j}$ a.s. This implies that there exists a subset $D_j$ of $S$ with $\mu^{*j}(D_j) = 1$ such that, for $g \in D_j$,

$$\lim_n\big(E[F(M_n g)] - E[F(M_n)]\big) = \lim_n\bigg(\int F(hg)\, d\mu^{*n}(h) - \int F(h)\, d\mu^{*n}(h)\bigg) = 0,$$

so that, for the limit value $\pi$, $\int F(hg)\, d\pi(h) = \int F(h)\, d\pi(h)$. Since the set of elements $g \in S$ satisfying this relation is closed and contains $D_j$ for arbitrary $j$, it contains $T$. Returning to $f$, we thus have, for all $g \in T$,

$$\int f(hgk)\, d\pi(h)\, d\pi(k) = \int f(hk)\, d\pi(h)\, d\pi(k) = \int f(h)\, d\pi(h). \qquad \square$$

LEMMA 9.3. *Let $T$ be a subsemigroup of $S$ containing the identity matrix and $\pi$ a probability on $T$ such that, for each $g \in T$, $\pi * \delta_g * \pi = \pi$. Then there exists $c \in \mathbb{R}$ such that, for each $g \in T$, $1 \leq ||g|| \leq c$.*

*Assume, moreover, that $\pi(S^\circ) > 0$. Then the mean matrix $h_0 = \int h\, d\pi(h)$ may be written as $h_0 = r_0 \otimes l_0$, with $r_0, l_0 \in C$ and $\langle r_0, l_0 \rangle = 1$. If $A$ and $A^*$ denote the affine subsets spanned by $Tr_0$ and $T^* l_0$, $A$ and $A^*$ are conjugate elements of $\mathscr{K}^+$ and $T \subset S_A \cap S^*_{A^*}$.*

Recall that, with respect to some probability distribution $\pi$ on $S$, the mean matrix entries are the mean values of the entries under $\pi$.

PROOF. Let $f$ be a real continuous function with compact support in $S$, and let $(g_n)_n$ be a sequence in $T$. For each $n$,

$$\int f(hg_n k)\, d\pi(h)\, d\pi(k) = \int f(h)\, d\pi(h).$$

Suppose there exists a $g \in T$ with $||g|| < 1$. Then $(g^n)_n \to 0$, so that, setting $g_n = g^n$, the limit of the left-hand side of the preceding relation is $f(0)$, which leads to a contradiction.

Assume the existence of a sequence $(g_n)_n$ such that $\lim_n ||g_n|| = +\infty$. By considering a subsequence, we may suppose that there exist $i, j$ such that $\lim_n \langle e_i, g_n e_j \rangle = +\infty$. Let $h_1, h_2 \in S$. We may choose $i'$ and $j'$ such that $\langle e_{i'}, h_1 e_i \rangle > 0$ and $\langle e_j, h_2 e_{j'} \rangle > 0$. It follows that $\lim_n ||h_1 g^n h_2|| = +\infty$. Hence we have $\pi \times \pi\{(h_1, h_2)\colon \lim_n ||h_1 g^n h_2|| = +\infty\} = 1$. The limit of the left-hand side is now 0, contradicting the fact that $\pi$ is a probability on $S$. The first assertion is established.

Since $\pi$ has a bounded support, we may consider the mean matrix $h_0$. From $\pi(S^\circ) > 0$, we get $h_0 \in S^\circ$. The convolution relation in the hypothesis gives $h_0 g h_0 = h_0$, for all $g \in T$. Replacing $g$ by the identity matrix, we see that $h_0$ is a projector. As it is in $S^\circ$ it has rank 1 (its main eigenvalue is simple) and so may be written as $h_0 = r_0 \otimes l_0$, where $r_0, l_0 \in C$ and $\langle r_0, l_0 \rangle = 1$. Then, returning to an arbitrary $g \in T$, we have $r_0 \otimes l_0 = (r_0 \otimes l_0) g (r_0 \otimes l_0) = \langle l_0, g r_0 \rangle (r_0 \otimes l_0)$. We conclude that $\langle l_0, g r_0 \rangle = 1$.

Obviously, for $g \in T$, $g(Tr_0) \subset Tr_0$ while $g^*(T^* l_0) \subset T^* l_0$. The same is true of the spanned affine subspaces. Notice that $A \cap C \ni r_0$ and $A - A \subset l_0^\perp$. As $l_0 \in C$, it follows that $(A - A) \cap \overline{C} = \{0\}$, so $A$ and hence $A^*$ are in $\mathscr{K}^+$. The last relation of the preceding paragraph shows that, for $a \in A$ and $a' \in A^*$, $\langle a, a' \rangle = 1$, so $A$ and $A^*$ are conjugated. □

END OF THE PROOF OF THEOREM 5. If $(X^{(n)})_{n \geq 1}$ is tight, so is $(\bar{\mu}_n)_n$. To conclude from Lemma 9.3, it just remains to verify that we may choose $r_0 = r_m$. From the relation $\mu * \pi = \pi * \mu = \pi$, we get $E[X_1] h_0 = h_0 E[X_1] = h_0$, so $E[X_1] r_0 = r_0$, as there exists $n$ such that $E[X_1]^n = E[X^{(n)}] \in S^\circ$; $r_0$ and $r_m$ have the same direction. □

PROOF OF COROLLARY 4. Lemma 9.3 shows that every element of $T$ has spectral radius 1. So $\Lambda_n = 1$ a.s. and, from Theorem 1(iii), $(X^{(n)})_n$ converges weakly to the probability $h(\nu' \times \nu)$.

If $g \in S_A$, $g^*$ preserves $A'$ and acts on this affine subspace according to the formula $g^* a' = l(g) + w(g)^* a'$. For a sequence $(g_k)_{k=1}^n$ in $S_A$, we get

$$g_1^* \cdots g_n^* a' = \sum_{k=1}^{n} w(g_1)^* \cdots w(g_{k-1})^* l(g_k) + w(g_1)^* \cdots w(g_n)^* a',$$

so that

$$Y^{(n)} l(X_{n+1}) = X_1^* \cdots X_n^* l(X_{n+1}) = \sum_{k=1}^{n+1} w(X_1)^* \cdots w(X_{k-1})^* l(X_k).$$

From Lemma 3.3, we know that

$$\lim_n \frac{Y^{(n)} l(X_{n+1})}{\langle \chi, Y^{(n)} l(X_{n+1}) \rangle} = Z_1 \quad \text{a.s.}$$

This implies that, if $a_0 \in A \cap C$, we have

$$\lim_n \frac{Y^{(n)} l(X_{n+1})}{\langle a_0, Y^{(n)} l(X_{n+1}) \rangle} = \frac{Z_1}{\langle a_0, Z_1 \rangle} \quad \text{as } \langle a_0, Y^{(n)} l(X_{n+1}) \rangle = 1.$$

We get

$$\frac{Z_1}{\langle a_0, Z_1 \rangle} = \sum_{k \geq 1} w(X_1)^* \cdots w(X_{k-1})^* l(X_k).$$

If $g \in S_A$, $g^* \in S_{A'}$ and we may write $g^* = r'(g^*) \otimes l'(g^*) + w'(g^*)$, with $r'(g^*) = l(g)$, $l'(g^*) = r(g)$ and $w'(g^*) = w(g)^*$. Using the construction of $Z_1'$ given in Section 4.2 and proceeding as before, we get, with $a_0' \in A' \cap C$,

$$\begin{aligned} \frac{Z_1'}{\langle a_0', Z_1' \rangle} &= \sum_{k \geq 1} w'(X_{-1}^*)^* \cdots w'(X_{-k+1}^*)^* l'(X_{-k}^*) \\ &= \sum_{k \geq 1} w(X_{-1}) \cdots w(X_{-k+1}) r(X_{-k}). \end{aligned}$$

To conclude, just notice that, since

$$h\left( \frac{z'}{\langle a_0', z' \rangle}, \frac{z}{\langle a_0, z \rangle} \right) = h(z', z),$$

we have $h(\tilde{\nu}' \times \tilde{\nu}) = h(\nu' \times \nu)$. $\square$

10. Proofs of facts about distance $d$. The aim of this section is to establish Proposition 3.1 and some connected formulas.

Recall that, for $x = (x_1, \dots, x_q),\ y = (y_1, \dots, y_q) \in \overline{B}$,

$$m(x, y) = \sup\{\lambda\colon \lambda \in \mathbb{R}_+,\ \forall i = 1, \dots, q,\ \lambda y_i \le x_i\}$$
$$= \min\left\{\frac{x_i}{y_i}\colon i = 1, \dots, q, y_i > 0\right\},$$

and that

$$d(x, y) = \varphi(m(x, y)m(y, x)), \qquad \varphi(s) = \frac{1-s}{1+s}, \qquad s \in [0, 1].$$

The reader will easily verify the following properties of function $m$, keeping in mind that, for $x \in \overline{B}$, $\sum_{i=1}^{q} x_i = 1$.

LEMMA 10.1. *Let $x, y, z \in \overline{B}$. Then*

(i) $m(x, y) \in [0, 1]$;
(ii) $m(x, z)m(z, y) \le m(x, y)$;
(iii) $m(x, y)m(y, x) = 1$ *if and only if* $x = y$;
(iv) $m(x, y) = 0$ *if and only if there exists* $i_0$ *such that* $x_{i_0} = 0$ *and* $y_{i_0} \ne 0$.

LEMMA 10.2. (i) *$d$ is a distance on $\overline{B}$:*
(ii) $\sup\{d(x, y)\colon x, y \in \overline{B}\} = 1$;
(iii) *if $x \in B$ and $y \in \overline{B}$, then $d(x, y) = 1$ if and only if $y \in \overline{B}\setminus B$.*

PROOF. Since $\varphi'(s) = -2/(1+s)^2$, $\varphi$ is decreasing.

Moreover, the function $F(s) = \varphi(s) + \varphi(t) - \varphi(st)$ satisfies $F(1) = 0$ and

$$F'(s) = -\frac{2(1-t)}{(1+s)^2(1+st)^2}(1 - s^2 t),$$

so that, for $s, t \in [0, 1]$, $\varphi(st) \le \varphi(s) + \varphi(t)$.

These two properties of $\varphi$ and the preceding lemma allow us to conclude. □

To study $d$, it is more convenient to use the following formula.

LEMMA 10.3. *Let $x, y \in \overline{B}$, $x \ne y$. We set*

$$a = (1-\lambda_1)x + \lambda_1 y, \qquad \lambda_1 = \inf\{\lambda\colon (1-\lambda)x + \lambda y \in \overline{B}\},$$

$$b = (1-\lambda_2)x + \lambda_2 y, \qquad \lambda_2 = \sup\{\lambda\colon (1-\lambda)x + \lambda y \in \overline{B}\},$$

*where $a$ and $b$ are the end points of the segment which is the intersection of $\overline{B}$ with the line through $x$ and $y$.*

*Writing $x = u_1 a + u_2 b$ and $y = v_1 a + v_2 b$, we have*

$$d(x, y) = \frac{|u_1 v_2 - u_2 v_1|}{u_1 v_2 + u_2 v_1}.$$

PROOF. Let us denote $a=(a_1,\dots,a_q)$, $b=(b_1,\dots,b_q)$ and set $I=\{i: i=1,\dots,q,\ a_i>0\}$, $J=\{i: i=1,\dots,q,\ b_i>0\}$.

There is no inclusion relation between $I$ and $J$. In particular, $I$ and $J$ are different from $\{1,\dots,q\}$. In fact, suppose $J\subset I$. Then it is possible to choose $\varepsilon>0$ such that $(1+\varepsilon)a-\varepsilon b\in\overline{B}$. This vector belongs to the line through $x, y$ but does not belong to the segment $[a,b]$, which gives a contradiction.

We first verify that the formula giving $d(x,y)$ is correct if one of these points coincides with $a$ or $b$. Suppose, for example, $x=a$. From the preceding discussion and point (iv) of Lemma 10.1, $m(x,y)=0$, so that we get $d(x,y)=1$. This is also the value of the right-hand side when $u_1=1$, $u_2=0$ and $v_2\neq 0$.

Consider now $x=(x_1,\dots,x_q)$, $y=(y_1,\dots,y_q)$, both different from $a$ and $b$. Set $r=\min\{u_i/v_i: i=1,2\}$. As $ry_i=r(v_1a_i+v_2b_i)\le u_1a_i+u_2b_i=x_i$, we have $r\le m(x,y)$. Conversely, writing $m(x,y)(v_1a_i+v_2b_i)\le u_1a_i+u_2b_i$ for $i_0\in J\backslash I$ and $j_0\in I\backslash J$, we get $m(x,y)v_2b_{i_0}\le u_2b_{i_0}$ and $m(x,y)v_1a_{j_0}\le u_1a_{j_0}$. Hence $m(x,y)\le r$. At last $m(x,y)=r$. Since $x\neq y$, $m(x,y)m(y,x)<1$. Consequently,

$$m(x,y)m(y,x)=\min\left\{\frac{u_1}{v_1}\frac{v_2}{u_2},\frac{u_2}{v_2}\frac{v_1}{u_1}\right\}$$

and $d(x,y)$ is written as stated. □

REMARK. The cross-ratio of $(a,b,x,y)$ is the element of the extended real line defined by $[a,b,x,y]=(u_1/v_1)/(u_2/v_2)$. We now consider $\varphi$ as a function on this line, $d(x,y)=\varphi([a,b,x,y])$, while $d_H(x,y)=|\ln[a,b,x,y]|$.

LEMMA 10.4. *For $x,y\in\overline{B}$, $d(x,y)\ge\frac{1}{2}||x-y||$.*

PROOF. We use the notation and the formula of the preceding lemma. Noticing that $u_1+u_2=v_1+v_2=1$, we get

$$|u_1v_2-u_2v_1|=|u_1(1-v_1)-(1-u_1)v_1|=|u_1-v_1|=\tfrac{1}{2}||x-y||.$$

Moreover,

$$0<u_1v_2+u_2v_1\le(u_1^2+u_2^2)^{1/2}(v_1^2+v_2^2)^{1/2}\le(u_1+u_2)^{1/2}(v_1+v_2)^{1/2}=1,$$

hence the stated inequality. □

LEMMA 10.5. *Set $d_1$ for the distance on $\overline{B}$ associated with $||\cdot||$.*

*Let $y\in B$, $x\in\overline{B}$ and $(x^{(n)})_n$ be a sequence in $\overline{B}$ such that $\lim_n d_1(x^{(n)},x)=0$. Then $\lim_n d(x^{(n)},y)=d(x,y)$.*

*The spaces $(B,d)$ and $(B,d_1)$ are homeomorphic.*

PROOF. Since, for all $i$, $\langle y, e_i\rangle > 0$, if $n$ is large enough

$$m(y, x^{(n)}) = \min\left\{\frac{\langle y, e_i\rangle}{\langle x^{(n)}, e_i\rangle}: \forall i,\ \langle x^{(n)}, e_i\rangle > 0\right\}$$
$$= \min\left\{\frac{\langle y, e_i\rangle}{\langle x^{(n)}, e_i\rangle}: \forall i,\ \langle x, e_i\rangle > 0\right\},$$

so that $\lim_n m(y, x^{(n)}) = m(y, x)$. This property is not true when $y \in \overline{B}\backslash B$. For example, set $x = y = e_1$ and $x^{(n)} = (1 - 1/n)e_1 + (1/n)\chi$. We have $m(x, x^{(n)}) = 0$ but $m(x, x) = 1$. On the contrary, it is clear that the convergence $\lim_n m(x^{(n)}, y) = m(x, y)$ holds for every $y \in \overline{B}$. Hence we have the first assertion.

Restricting the preceding to $x = y \in B$, we conclude that the identity is a continuous function from $(B, d_1)$ to $(B, d)$.

Conversely, Lemma 10.4 shows that, in $B$, the $d$-convergence implies $d_1$-convergence. □

REMARK. The spaces $(\overline{B}, d)$ and $(\overline{B}, d_1)$ are not homeomorphic. Indeed, point (iii) of Lemma 10.2 shows that $(\overline{B}, d)$ is not connected.

We come now to the contraction properties.

LEMMA 10.6. *Let $g \in S$ and set $c(g) = \sup\{d(g\cdot x, g\cdot y): x, y \in \overline{B}\}$. Then:*

(i) *for $x, y \in \overline{B}$, $d(g\cdot x, g\cdot y) \le c(g)d(x, y)$;*
(ii) *if $g' \in S$, $c(gg') \le c(g)c(g')$;*
(iii) *$c(g) \le 1$, moreover, $c(g) < 1$ if and only if $g \in S^\circ$.*

PROOF. Let $x, y \in \overline{B}$, $x \neq y$. If $g\cdot x = g\cdot y$, inequality (i) is established. Suppose $g\cdot x \neq g\cdot y$. Let $a, b$ and $a_1, b_1$ be the extreme points of the segments obtained as the intersections with $\overline{B}$ of the lines through $x, y$ and through $g\cdot x, g\cdot y$. Consider $g$ as an isomorphism between the two-dimensional spaces whose bases are $(a, b)$ and $(a_1, b_1)$. With respect to these bases, $g$ has matrix $\left(\begin{smallmatrix}\alpha & \beta\\ \gamma & \delta\end{smallmatrix}\right)$. Clearly, $\alpha, \beta, \gamma, \delta \ge 0$; moreover, $\alpha\delta + \beta\gamma > 0$, since otherwise the matrix has a zero lign or a zero column. Set $x = u_1 a + u_2 b$ and $y = v_1 a + v_2 b$, $gx = (\alpha u_1 + \beta u_2)a_1 + (\gamma u_1 + \delta u_2)b_1$ and a similar formula is available for $gy$, so that, from Lemma 10.3,

$$d(g\cdot x, g\cdot y) = \frac{|(\alpha u_1 + \beta u_2)(\gamma v_1 + \delta v_2) - (\gamma u_1 + \delta u_2)(\alpha v_1 + \beta v_2)|}{(\alpha u_1 + \beta u_2)(\gamma v_1 + \delta v_2) + (\gamma u_1 + \delta u_2)(\alpha v_1 + \beta v_2)}$$
$$= \frac{|\alpha\delta - \beta\gamma||u_1 v_2 - u_2 v_1|}{2\alpha\gamma u_1 v_1 + (\alpha\delta + \beta\gamma)(u_1 v_2 + u_2 v_1) + 2\beta\delta u_2 v_2}$$
$$\le \frac{|\alpha\delta - \beta\gamma||u_1 v_2 - u_2 v_1|}{(\alpha\delta + \beta\gamma)(u_1 v_2 + u_2 v_1)}$$
$$= d(g\cdot a, g\cdot b)d(x, y) \le c(g)d(x, y).$$

If $g' \in S$, for $x, y \in \overline{B}$ we have

$$d\big((gg')\cdot x, (gg')\cdot y\big) \le c(g) d(g'\cdot x, g'\cdot y),$$

and hence (ii), owing to the definitions of $c(gg')$ and $c(g')$.

It is clear that $c(g) \le 1$.

If $g \in S \backslash S^{\circ}$ there exists $i$ such that $g \cdot e_i \in \overline{B} \backslash B$, but $g \cdot \chi \in B$, so that $1 = d(g\cdot e_i, g\cdot \chi) \le c(g)$ and $c(g) = 1$.

Suppose $g \in S^{\circ}$. We have $g \cdot \overline{B} \subset B$, so $g\cdot \overline{B}$ is a compact of $(B, d_1)$ and hence, by Lemma 10.5, of $(B, d)$. Consequently, there exist $x_0, y_0 \in \overline{B}$ such that $c(g) = d(g\cdot x_0, g\cdot y_0) < 1$. □

It is possible to give an explicit formula for $c(g)$.

LEMMA 10.7. *If* $g = [g_{ij}]_{i,\, j=1,\dots,q} \in S$,

$$c(g) = \max\{d(g\cdot e_i, g\cdot e_j)\colon i, j = 1, \dots, q\} = \max_{i,\, j,\, k,\, \ell=1,\dots,q} \frac{|g_{ki}g_{\ell j} - g_{kj}g_{\ell i}|}{g_{ki}g_{\ell j} + g_{kj}g_{\ell i}},$$

*so that* $c(g^*) = c(g)$.

PROOF. The equality is straightforward if $g \in S\backslash S^{\circ}$.

Let $g \in S^{\circ}$. We are going to show that

$$\sup\big\{d(g\cdot x, g\cdot y)\colon x, y \in \overline{B}\big\} = \max\big\{d(g\cdot e_i, g\cdot e_j)\colon i, j\big\}$$

and that this last term is calculated by means of the stated formula.

Let $x = (x_1, \dots, x_q)$, $y = (y_1, \dots, y_q)$ be two elements of $g\cdot \overline{B} \subset B$. We may write

$$x = \sum_{i=1}^{q} \alpha_i \; g\cdot e_i, \qquad y = \sum_{j=1}^{q} \beta_j \; g\cdot e_j,$$

where $(\alpha_i)_{i=1}^q$ and $(\beta_j)_{j=1}^q$ are sequences of positive numbers with sum 1. We have

$$m(x, y)m(y, x) = \min_k \frac{x_k}{y_k} \min_\ell \frac{y_\ell}{x_\ell} = \min_{k,\ell} \frac{x_k y_\ell}{y_k x_\ell}$$

and

$$\frac{x_k y_\ell}{y_k x_\ell} = \frac{x_k}{x_\ell}\frac{y_\ell}{y_k} = \frac{\sum_i \alpha_i g_{ki}}{\sum_i \alpha_i g_{\ell i}} \frac{\sum_j \beta_j g_{\ell j}}{\sum_j \beta_j g_{kj}} \ge \min_i \frac{g_{ki}}{g_{\ell i}} \min_j \frac{g_{\ell j}}{g_{kj}} = \min_{i,j} \frac{g_{ki}g_{\ell j}}{g_{\ell i}g_{kj}}.$$

Using the first relations, we get

$$\min_{k,\ell} \frac{g_{ki}g_{\ell j}}{g_{\ell i}g_{kj}} = m(g\cdot e_i, g\cdot e_j)m(g\cdot e_j, g\cdot e_i).$$

Consequently, the second relations lead to

$$m(x, y)m(y, x) \ge \min_{i,\, j,\, k,\, \ell} \frac{g_{ki}g_{\ell j}}{g_{\ell i}g_{kj}} = \min_{i,\, j} m(g\cdot e_i, g\cdot e_j)m(g\cdot e_j, g\cdot e_i).$$

Applying the decreasing function $\varphi$, we conclude that the $d$-diameter of $g \cdot \overline{B}$ and the $d$-diameter of $\{g \cdot e_i\colon i\}$ are equal and that, considering the function $\varphi$ as defined on $\mathbb{R}_+$, this number is equal to

$$\max_{i,\,j,\,k,\,\ell} \varphi\left(\frac{g_{ki}g_{\ell j}}{g_{\ell i}g_{kj}}\right) = \max_{i,\,j,\,k,\,\ell} \frac{g_{\ell i}g_{kj} - g_{ki}g_{\ell j}}{g_{\ell i}g_{kj} + g_{ki}g_{\ell j}} = \max_{i,\,j,\,k,\,\ell} \frac{|g_{\ell i}g_{kj} - g_{ki}g_{\ell j}|}{g_{\ell i}g_{kj} + g_{ki}g_{\ell j}}. \qquad \square$$

Let us complete this section with the answer to a legitimate question.

LEMMA 10.8. *We have that $c$ is a continuous function on $S$.*

PROOF. Let $g \in S$ and let $(g_n)_n$ be a sequence in $S$ such that $\lim_n \|g_n - g\| = 0$.

Suppose $g \in S^\circ$. For all $i$, we have $\lim_n d_1(g_n \cdot e_i, g \cdot e_i) = 0$ and $g \cdot e_i \in \overline{B}\backslash B$, so by Lemma 10.5, for all $i$ and $j$, $\lim_n d(g_n \cdot e_i, g_n \cdot e_j) = d(g \cdot e_i, g \cdot e_j)$. Consequently,

$$\lim_n c(g_n) = \lim_n \max\{d(g_n \cdot e_i, g_n \cdot e_j)\colon i, j\} = \max\{d(g \cdot e_i, g \cdot e_j)\colon i, j\} = c(g).$$

Suppose $g \in S\backslash S^\circ$. Let $i$ be such that $g \cdot e_i \in \overline{B}\backslash B$. Then

$$c(g_n) \geq d(g_n \cdot \chi, g_n \cdot e_i) \geq d(g \cdot \chi, g_n \cdot e_i) - d(g \cdot \chi, g_n \cdot \chi).$$

By Lemma 10.5, $\liminf_n c(g_n) \geq d(g \cdot \chi, g \cdot e_i) = 1$. $\square$

## REFERENCES


[1] ARNOLD, L., DEMETRIUS, L. and GUNDLACH, V. M. (1994). Evolutionary formalism for products of positive random matrices. *Ann. Appl. Probab.* 4 859–901.

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

[3] BOUGEROL, P. and LACROIX, J. (1985). *Products of Random Matrices with Applications to Schrödinger Operators*. Birkhäuser, Boston.

[4] BUSHELL, P. J. (1973). Hilbert's metric and positive contraction mappings in a Banach space. *Arch. Rational Mech. Anal.* 52 330–338.

[5] COHEN, J. (1986). Products of random matrices and related topics in mathematics and science—a bibliography. *Contemp. Math.* 50 337–358.

[6] COHN, H., NERMAN, O. and PELIGRAD, M. (1993). Weak ergodicity and products of random matrices. *J. Theoret. Probab.* 6 389–405.

[7] CORNFELD, I. P., FOMIN, S. V. and SINAI, YA. (1982). *Ergodic Theory*. Springer, New York.

[8] DOUKHAN, P. (1994). *Mixing: Properties and Examples. Lecture Notes in Statist.* 85. Springer, New York.

[9] DÜRR, D. and GOLDSTEIN, S. (1984). Remarks on the CLT for weakly dependent random variables. *Lecture Notes in Math.* 1158. Springer, New York.

[10] DURRETT, R. (1991). *Probability: Theory and Examples.* Wadsworth and Brooks/Cole, Monterey, CA.

[11] ESSEEN, C. G. and JANSON, S. (1985). On moment conditions for sums of independent variables and martingale differences. *Stochastic Process. Appl.* 19 173–182.

[12] EVSTIGNEEV, C. G. I.V. (1974). Positive matrix-valued cocycles over dynamical systems. *Uspehi Mat. Nauk* 29 219–226.

[13] FERRERO, P. and SCHMITT, B. (1988). Produits aléatoires d'opérateurs matrices de transfert. *Probab. Theory Related Fields* 79 227–248.

[14] FURSTENBERG, H. (1963). Non-commuting random products. *Trans. Amer. Math. Soc.* 108 377–428.
[15] FURSTENBERG, H. and KESTEN, H. (1960). Products of random matrices. *Ann. Math. Statist.* 31 457–469.
[16] GORDIN, M. I. (1969). The central limit theorem for stationary processes. *Soviet Math. Dokl* 10 1174–1176.
[17] GUIVARC'H, Y. and RAUGI, A. (1986). Products of random matrices: convergence theorems. *Contemp. Math.* 50 31–54.
[18] HALL, P. and HEYDE, C. C. (1980). *Martingale Limit Theory and Its Application.* Academic Press, New York.
[19] HENNION, H. (1991). Dérivabilité du plus grand exposant caractéristique des produits de matrices aléatoires indépendantes à coefficients positifs. *Ann. Inst. H. Poincaré Probab. Statist.* 27 27–59.
[20] KESTEN, H. and SPITZER, F. (1984). Convergence in distribution of products of random matrices. *Z. Wahrsch. Verw. Gebiete* 67 363–386.
[21] KINGMAN, J. F. C. (1973). Subadditive ergodic theory. *Ann. Probab.* 1 883–899.
[22] KRANOSEL'SKIJ, M. A., LIFSHITS, JE. A. and SOBOLEV, A. V. (1980). *Positive Linear Systems—The Method of Positive Operators. Sigma Ser. Appl. Math.* Heldermann, Berlin.
[23] LEDRAPPIER, F. (1982). Quelques propriétés des exposants caractéristiques. *Ecole d'été de Saint-Flour 12. Lecture Notes in Math.* 1097. Springer, Berlin.
[24] LE PAGE, E. (1982). Théorèmes limites pour les produits de matrices aléatoires. *Probability Measures on Groups. Lecture Notes in Math.* 928 258–303. Springer, Berlin.
[25] LIVERANI, C. (1996). Central limit theorem for deterministic systems. In *International Conference on Dynamical Systems, Montevideo 1995: a Tribute to Ricardo Mané.*
[26] MUKHERJEA, A. (1987). Convergence in distribution of products of random matrices: a semigroup approach. *Trans. Amer. Math. Soc.* 303 395–411.
[27] PELLAUMAIL, J. (1990). Graphes et algorithme de calcul de probabilités stationnaires d'un processus markovien discret. *Ann. Inst. H. Poincaré Probab. Statist.* 26 121–143.
[28] PERES, Y. (1992). Domain of analytic continuation for the top Lyapunov exponent. *Ann. Inst. H. Poincaré Probab. Statist.* 28 131–148.
[29] RAUGI, A. (1983). Une démonstration du théorème de Choquet–Deny par les martingales. *Ann. Inst. H. Poincaré Statist. Probab.* 19 101–109.
[30] SENETA, E. (1981). *Nonnegative Matrices and Markov Chains.* Springer, New York.
[31] VOLNY, D. (1993). Approximating martingales and the central limit theorem for strictly stationary processes. *Stochastic Process. Appl.* 44 41–74.

INSTITUT MATHÉMATIQUES DE RENNES
UNIVERSITÉ DE RENNES I
CAMPUS DE BEAULIEU
35042 RENNES-CEDEX
FRANCE
E-MAIL: hubert.hennion@univ-rennes1.fr