*The Annals of Probability*
1987, Vol. 15, No. 1, 40–74

# TIGHTNESS OF PRODUCTS OF RANDOM MATRICES AND STABILITY OF LINEAR STOCHASTIC SYSTEMS

BY PHILIPPE BOUGEROL

*Université Paris 7*

Let $\mu^n$ be the distribution of a product of $n$ independent identically distributed random matrices. We study tightness and convergence of the sequence $\{\mu^n,\ n \geq 1\}$. We apply this to linear stochastic differential (and difference) equations, characterize the stability in probability, in the sense of Hashminski, of the zero solution, and find all their stationary solutions.

This paper is devoted to the study of the tightness of products of i.i.d. random matrices and its applications to stability properties of linear stochastic equations evolving either in discrete or continuous time (mainly when the upper Liapounov exponent is zero).

Let us first give an example of the results we obtain.

Consider the continuous time stochastic differential equation on $\mathbb{R}^d$:

$$dx_t = S_0 x_t\, dt + \sum_{i=1}^{r} S_i x_t \circ db_t^i, \tag{1}$$

where $S_0, S_1, \dots, S_r$ are fixed matrices of order $d$ and $b_t^1, b_t^2, \dots, b_t^r$ $r$ independent real Brownian motions ($\circ\, db_t^i$ is the Stratonovich differential). Let $\mathscr{M}(d)$ be the set of real matrices of order $d$ and consider the solution $(M_t,\ t \geq 0)$ of the following equation on $\mathscr{M}(d)$:

$$dM_t = S_0 M_t\, dt + \sum_{i=1}^{r} S_i M_t \circ db_t^i, \qquad M_0 = Id. \tag{2}$$

For any $x$ in $\mathbb{R}^d$, $x_t := M_t x$ is a solution of (1).

The upper Liapounov exponent $\gamma = \gamma(S_0, \dots, S_r)$ associated with (1) is

$$\gamma = \lim_{t \to \infty} \frac{1}{t} E\left(\log \|M_t\|\right).$$

If $\gamma = 0$, the zero solution (i.e., $x_t = 0$ for any $t \geq 0$) of (1) is not almost surely stable but the question arises (cf. Hashminski [8]) whether it is stable in probability. Recall that:

DEFINITION ([8], page 25). The zero solution of (1) is said to be stable in probability if for every $\varepsilon > 0$ and $\eta > 0$, there exists a $\delta > 0$ such that:

$$\text{for} \quad \|y\| < \delta \quad \text{and} \quad t > 0, \qquad P\left(\|M_t y\| \geq \eta\right) \leq \varepsilon.$$

---

Received January 1985; revised August 1985.

*AMS* 1980 *subject classifications*. Primary 60B15; secondary 60B10, 60H10, 60H25.

*Key words and phrases*. Products of random matrices, linear stochastic systems, stability in probability, stationary solution, convergence in distribution, linear stochastic differential equations.

We shall prove the following (see Theorem 7.3):

The zero solution of (1) is stable in probability if and only if there exists an invertible matrix $Q$ such that for $i = 0, \dots, r$

$$QS_iQ^{-1} = \begin{pmatrix} A_i & * & * \\ 0 & K_i & * \\ 0 & 0 & B_i \end{pmatrix},$$

where each $K_i$ is a skew-symmetric matrix and the upper Liapounov exponents $\gamma(A_0, \dots, A_r)$ and $\gamma(B_0, \dots, B_r)$ are strictly negative.

If $\mu_t$ is the law of $M_t$, $(\mu_t)_{t \geq 0}$ is a convolution semigroup on $\mathscr{M}(d)$. It is easily seen that the stability in probability of the zero solution of (1) is equivalent to the tightness of $\{\mu_n, n \geq 1\}$ on $\mathscr{M}(d)$. This leads us to consider the following more general problem:

Problem (P). Let $Y_1, Y_2, \dots$ be independent matrices on $\mathscr{M}(d)$ with the same arbitrary distribution $\mu$. Denote by $\mu^n$ the distribution of $M_n := Y_nY_{n-1} \cdots Y_1$. When is $\{\mu^n, n \geq 1\}$ a tight sequence of probability measures on $\mathscr{M}(d)$?

The purpose of this paper is to solve it under weak assumptions on $\mu$. This will for instance permit us to study the convergence in distribution of $M_n$, a question raised by Kesten and Spitzer in [13]. Specific results on the solutions of (1) will be immediate applications of general theorems (needless to say some of our arguments can be simplified if one deals only with (1)).

*Organization of the paper.* Consider a probability measure $\mu$ on $\mathscr{M}(d)$.

This paper is organized as follows. In Section 1 we introduce the notation and definitions we shall use. In Sections 2 and 3 we give a necessary condition ensuring the tightness of $\{\mu^n, n \in \mathbb{N}\}$ either if $\mu$ is carried by the invertible matrices or if there exists in the closure of $\bigcup_{n \in \mathbb{N}} \mathrm{Supp}(\mu^n)$ a matrix with at most one (simple) eigenvalue of modulus one. For instance we prove in the first case that if $\{\mu^n, n \in \mathbb{N}\}$ is tight then there exists some matrix $Q$ such that for every $n$ each $M_n$ can be written as

$$(3) \qquad M_n = Q \begin{pmatrix} A_n & C_n & E_n \\ 0 & K_n & D_n \\ 0 & 0 & B_n \end{pmatrix} Q^{-1}, \quad \text{a.s.},$$

where $K_n$ is an orthogonal matrix and the sequences $\{A_n, n \in \mathbb{N}\}$ and $\{B_n, n \in \mathbb{N}\}$ converge in distribution in the Cesàro sense to the zero matrix (see Theorem 3.1). Under some stronger assumptions on $\mu$ the Liapounov exponents associated with $A_n$ and $B_n$ are strictly negative. In Section 2 we suppose that $\mu$ satisfies an irreducibility condition and use a method introduced in Furstenberg [5]. The general case is considered in Section 3.

In Section 4 we prove a converse of these results. For instance if a decomposition such as in (3) holds, if $E(\log^+\|M_1\|)$ is finite, and if the Liapounov exponents associated with $(A_n)$ and $(B_n)$ are strictly negative then $\{\mu^n, n \in \mathbb{N}\}$ is tight on $\mathscr{M}(d)$.

In Section 5 we apply these results to describe the stationary solutions of the following stochastic system on $\mathbb{R}^d$:

$$x_n = Y_n x_{n-1}, \quad \text{for } n \geq 1, \tag{4}$$

where $Y_1, Y_2, \ldots$ are i.i.d. random matrices. In other words we describe the set of the invariant probability measures of the $\mathbb{R}^d$-valued Markov chain $M_n x$.

In Section 6 we study the asymptotic behaviour of the paths of the Markov chain $M_n x$, for $x$ in $\mathbb{R}^d$, when each $M_n$ can be written in the form (3) and when the Liapounov exponents associated with $A_n$ and $B_n$ are strictly negative. This will be a consequence of a result which shows that the following situation may be considered as the model case:

For two integers $d$ and $d'$ let $(L_n)$, $(K_n)$, and $(R_n)$ be three independent sequences of random matrices such that

(i) $L_1, L_2, \ldots$ are $d \times d'$ random i.i.d. matrices;
(ii) $R_1, R_2, \ldots$ are $d' \times d$ random i.i.d. matrices;
(iii) $K_1, K_2, \ldots$ are i.i.d. orthogonal $d' \times d'$ matrices;
(iv) For each integer $m$, $n$, $R_n L_m$ is the identity matrix of order $d'$.

In this case $M_n = Y_n \cdots Y_1 = L_n(K_n \cdots K_1)R_1$ and $Y_1 \cdots Y_n = L_1(K_1 \cdots K_n)R_n$. In general $M_n$ will be written asymptotically as a product of three matrices, the first converging in law (here it is $L_n$), the second being a random walk on the orthogonal group (here it is $K_n \cdots K_1$), and the third converging almost surely (here it is $R_1$). As in the model case the convergence in law of the first component will be a consequence of the a.s. convergence of the first one of $Y_1 \cdots Y_n$.

In Section 7 we apply some of these results to the study of the stability of the zero solution of (1). We describe which equations (1) are stable and their ergodic behaviour. As an immediate application of the results of Section 5 we find the stationary solutions of all the equations of the form (1).

Although we have stated all our results for real matrices, a lot of them are also true for complex matrices (but not the necessary condition in Section 2 under C2).

Finally we must say that these problems have already been considered. First by Furstenberg (see Theorem 1.2 of [5]), who introduced the main tools. Recently (see [13]), Kesten and Spitzer have completely solved them when the matrices are nonnegative.

## 1. Notation and definitions.

1.1. *Matrices.* If $d$ and $d'$ are integers we will denote by $\mathscr{M}(d)$, the set of real $d \times d$ matrices; $\mathrm{Gl}(d)$, the set of real invertible $d \times d$ matrices; $\mathscr{M}(d, d')$, the set of real $d \times d'$ matrices; $O(d)$, the set of orthogonal matrices (i.e., the $M$ in $\mathrm{Gl}(d)$ such that ${}^tM = M^{-1}$). We will frequently use the fact that if $K$ is a compact subgroup of $\mathrm{Gl}(d)$, for some $Q$ in $\mathrm{Gl}(d)$, $QKQ^{-1}$ is contained in $O(d)$; see (22.23) of Hewitt and Ross [10].

DEFINITION 1.1. Let $T$ be a subset of $\mathscr{M}(d)$.

(a) A linear subspace $V$ of $\mathbb{R}^d$ is said to be $T$-invariant if $Mx$ is in $V$ for each $x$ in $V$ and $M$ in $T$. If there is no proper subspace of $V$ which is $T$-invariant we say that $T$ acts irreducibly on $V$.

(b) $T$ is said to be irreducible if $T$ acts irreducibly on $\mathbb{R}^d$.

We introduce two kinds of subsemigroups $T$ of $\mathscr{M}(d)$ which will play a major role in the sequel ($T$ is a subsemigroup if $M, M' \in T$ implies $MM' \in T$).

DEFINITION 1.2. A subsemigroup $T$ of $\mathscr{M}(d)$ is said to be an $F$-semigroup if it satisfies the following:

(i) $T$ is finite and irreducible.

(ii) The spectral radius of each element of $T$ is 1.

(iii) There exists in $T$ a rank-one projection, i.e., a matrix $P$ such that $P^2 = P$ and $\dim(\operatorname{Im} P) = 1$.

For instance,

$$\left\{\pm\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}, \pm\begin{pmatrix}-1 & 0\\ 0 & 0\end{pmatrix}, \pm\begin{pmatrix}1 & 0\\ 2 & 0\end{pmatrix}, \pm\begin{pmatrix}-1 & 1\\ -2 & 2\end{pmatrix}\right\}$$

is an $F$-semigroup in $\mathscr{M}(2)$.

DEFINITION 1.3. Given three nonnegative integers $d_1, d_2, d_3$, $d = d_1 + d_2 + d_3$, we denote by $T(d_1; d_2; d_3)$ the set of matrices $M$ of $\mathscr{M}(d)$ whose entries $M_{i,j}$ satisfy:

$$\begin{aligned} M_{i,j} = 0 \quad & \text{if } d_1 < i \le d_1 + d_2 \text{ and } j \le d_1 \\ & \text{or if } d_1 + d_2 < i \le d \text{ and } j \le d_1 + d_2. \end{aligned}$$

We will write such a matrix as:

$$M = \begin{pmatrix} a(M) & c(M) & e(M) \\ 0 & k(M) & d(M) \\ 0 & 0 & b(M) \end{pmatrix} \tag{5}$$

with $a(M)$ in $\mathscr{M}(d_1)$, $k(M)$ in $\mathscr{M}(d_2)$, $b(M)$ in $\mathscr{M}(d_3)$, $c(M)$ in $\mathscr{M}(d_1, d_2)$, $e(M)$ in $\mathscr{M}(d_1, d_3)$ and $d(M)$ in $\mathscr{M}(d_2, d_3)$.

We shall usually write 0 for any zero matrix, the context making clear the dimension of that matrix. In the same way "$I$" will represent any identity matrix.

We choose on $\mathscr{M}(d)$ the supremum norm defined by

$$\|M\| = \sup\{\|Mx\|;\ x \in \mathbb{R}^d, \|x\| = 1\}.$$

1.2. *Measures.*

DEFINITION 1.4. Given a probability measure $\mu$ on $\mathscr{M}(d)$, $T(\mu)$ denotes the smallest closed semigroup $T$ in $\mathscr{M}(d)$ such that $\mu(T) = 1$. We say that $\mu$ is irreducible if $T(\mu)$ is irreducible.

If $\mu$ and $\mu'$ are probability measures on $\mathscr{M}(d)$ we denote by $\mu * \mu'$ the convolution product of $\mu$ and $\mu'$, i.e., the image of $\mu \otimes \mu'$ under the mapping $\psi$: $\mathscr{M}(d) \times \mathscr{M}(d) \to \mathscr{M}(d)$ defined by $\psi(M, M') = MM'$. For any integer $n$, $\mu^n$ will be the $n$th power of convolution of $\mu$, for instance $\mu^1 = \mu$, $\mu^2 = \mu * \mu$.

DEFINITION 1.5. Given a probability measure $\mu$ on $\mathscr{M}(d)$ and a probability measure $\nu$ on $\mathbb{R}^d$, we say that $\nu$ is $\mu$-invariant if for any bounded Borel function $f$ on $\mathbb{R}^d$

$$\iint f(Mx)\, d\mu(M)\, d\nu(x) = \int f(x)\, d\nu(x).$$

Notice that $\nu$ is a $\mu$-invariant probability measure on $\mathbb{R}^d$ if and only if it is an invariant probability measure for the Markov chain $x_n$ solution of (4).

We shall often make use of:

DEFINITION 1.6. Let $\mu$ be a probability measure on $\mathscr{M}(d)$. We say that

(i) $\mu$ satisfies Condition C1 if $\mu(\mathrm{Gl}(d)) = 1$.

(ii) $\mu$ satisfies Condition C2 if there exists in $T(\mu)$ a matrix with at most one eigenvalue of modulus one, this eigenvalue being simple.

(Recall that the eigenvalue $\lambda$ of a matrix $M$ is simple if for all $p \in \mathbb{N} \setminus 0$ the null space of $(M - \lambda I)^p$ is one-dimensional.)

Recall that if $X$ is a complete separable metric space, a family $\mathscr{F}$ of probability measures on $X$ is tight (i.e., of compact closure for the weak topology) if and only if for each $\varepsilon > 0$ there exists a compact set $K$ in $X$ such that $\nu(K) \geq 1 - \varepsilon$ for any $\mu$ in $\mathscr{F}$.

1.3. *Liapounov exponents.* Consider a probability measure $\mu$ on $\mathscr{M}(d)$. Given a sequence $Y_1, Y_2, \ldots$ of independent matrices with distribution $\mu$ we set $M_n = Y_n Y_{n-1} \cdots Y_1$. We define (see for instance Ledrappier [14]):

DEFINITION 1.7. If $E(\log^+ \|Y_1\|)$ is finite, the upper Liapounov exponent associated with $\mu$ is

$$\gamma(\mu) = \lim_{n \to \infty} \frac{1}{n} \log \|M_n\| \quad \text{a.s.}$$

We shall often call $\gamma(\mu)$ "the Liapounov exponent of $\mu$."

**2. The irreducible case.** In this part we consider an irreducible probability measure $\mu$ on $\mathscr{M}(d)$. We study the set of $\mu$-invariant probability measures on

$\mathbb{R}^d$. As a consequence we obtain a necessary condition on $\mu$ ensuring the tightness of $\{\mu^n, n \in \mathbb{N}\}$ on $\mathscr{M}(d)$, under one of the conditions C1 or C2 (see Definition 1.6).

We begin with a general result. The idea of the proof is borrowed from Furstenberg [5].

PROPOSITION 2.1. *Let $Y_1, Y_2, \dots$ be i.i.d. random matrices with distribution $\mu$ on $\mathscr{M}(d)$ and $U_n = Y_1 Y_2 \cdots Y_n$.*

*If there exists on $\mathbb{R}^d$ a $\mu$-invariant probability measure whose support is not contained in a hyperplane then the sequence $\{U_n(\omega), n \in \mathbb{N}\}$ is for almost all $\omega$ bounded and for each $M$ in $T(\mu)$, $\|M\| \geq 1$.*

*If, moreover, $\mu$ is irreducible then $T(\mu)$ is a bounded set.*

PROOF. Let $\nu$ be a $\mu$-invariant probability measure on $\mathbb{R}^d$ such that for any hyperplane $H$ of $\mathbb{R}^d$, $\nu(H) \neq 1$. If $M$ is in $\mathscr{M}(d)$ we define the measure $M\nu$ on $\mathbb{R}^d$ by

$$\int f(x)\, d(M\nu)(x) = \int f(Mx)\, d\nu(x)$$

for any Borel bounded function $f$ on $\mathbb{R}^d$.

Furstenberg has pointed out that $U_n\nu$ is a measure-valued martingale and thus converges almost surely to a random probability measure $\lambda_\omega$ on $\mathbb{R}^d$ (see Lemma 1.3 of [5]). Suppose that for a fixed $\omega$, $U_n(\omega)\nu$ converges to $\lambda_\omega$ but that $\sup\{\|U_n(\omega)\|, n \geq 1\}$ is not finite. We can find a subsequence $n(i)$ such that $U_{n(i)}(\omega)\nu$ converges to $\lambda_\omega$ and $\|U_{n(i)}(\omega)\|^{-1}U_{n(i)}(\omega)$ converges to a nonzero matrix $H(\omega)$. If $x$ is in $\mathbb{R}^d$ and $H(\omega)x \neq 0$, $\|U_{n(i)}(\omega)x\| \to +\infty$ when $i \to +\infty$. This implies that for any continuous $f$ with compact support on $\mathbb{R}^d$,

$$\begin{aligned} \int f\, d\lambda_\omega &= \lim_{i\to\infty} \int f\big(U_{n(i)}(\omega)x\big)\, d\nu(x) \\ &= \lim_{i\to\infty} \int 1_{\ker H(\omega)}(x) f\big(U_{n(i)}(\omega)x\big)\, d\nu(x), \end{aligned}$$

hence $\nu(\ker H(\omega)) = 1$. This contradicts the assumption on $\nu$ and the sequence $U_n(\omega)$ must be almost surely bounded.

It follows from Lemma 2.13 in Guivarc'h and Raugi [7] that for almost all $M$ with respect to $\sum_{n=1}^{\infty} 2^{-n}\mu^n$ and almost all $\omega$, $U_n(\omega)M\nu$ converges weakly to $\lambda_\omega$. If $U(\omega)$ is a limit point of the bounded sequence $U_n(\omega)$ we thus have

$$U(\omega)M\nu = \lambda_\omega \quad \text{a.s.} \tag{6}$$

This equality remains true for any $M$ in the support of $\sum_{n=1}^{\infty} 2^{-n}\mu^n$, i.e., for any $M$ in $T(\mu)$. If we could find a matrix $M$ in $T(\mu)$ such that $\|M\| < 1$, the zero matrix would be in $T(\mu)$. By (6) this would imply that $\lambda_\omega$ is the Dirac measure $\delta_0$. But this cannot hold since, $U_n\nu$ being a bounded martingale, $\nu = E(\lambda_\omega)$.

We now prove that if moreover $\mu$ is irreducible, then $T(\mu)$ is bounded. Suppose there exists a sequence $\{M_n, n \in \mathbb{N}\}$ in $T(\mu)$ such that $\lim\|M_n\| = +\infty$. We may assume that $\|M_n\|^{-1}M_n$ converges to some nonzero matrix $A$. By (6) we have for each integer $p$ and matrix $M$ in $T(\mu)$

$$U(\omega)MM_p\nu = \lambda_\omega \quad \text{a.s.}$$

For such an $\omega$ fixed let $V = \{x \in \mathbb{R}^d \colon U(\omega)MAx = 0\}$. If $x$ is not in $V$, $\|U(\omega)MM_p x\| \to \infty$ when $p \to \infty$. Therefore if $f$ is a continuous function with compact support on $\mathbb{R}^d$,

$$\int f(x)\, d\lambda_\omega(x) = \int f\big(U(\omega)MM_p x\big)\, d\nu(x)$$

$$= \lim_{p \to \infty} \int 1_V(x) f\big(U(\omega)MM_p x\big)\, d\nu(x).$$

This implies that $\nu(V) = 1$ and $V$ must be $\mathbb{R}^d$. Choose some $x$ such that $Ax$ is not zero. For all $M$ in $T(\mu)$, $U(\omega)MAx = 0$. So the subspace $W$ of $\mathbb{R}^d$ spanned by $\{MAx, M \in T(\mu)\}$ is contained in the null space of $U(\omega)$. Since $W$ is $T(\mu)$-invariant the irreducibility of $\mu$ implies that $U(\omega)$ is the zero matrix. This cannot hold since the zero matrix is not in $T(\mu)$. □

We need the following algebraic result:

PROPOSITION 2.2. *Let $T$ be an irreducible subsemigroup of $\mathscr{M}(d)$ such that for some $c > 0$, $1 \le \|M\| \le c$ for each $M$ in $T$. Then*:

(a) *If $T$ contains a matrix with a unique eigenvalue of modulus one, this eigenvalue being simple, then $T$ must be an $F$-semigroup* (*see Definition* 1.2).

(b) *If $T$ is contained in* $\mathrm{Gl}(d)$, *$T$ is contained in a compact subgroup.*

PROOF OF (a). It is easily seen using the Jordan decomposition that under the hypothesis of (a), the closure $S$ of $T$ in $\mathscr{M}(d)$ contains a rank one projection $P$. We put on $\mathbb{R}^d$ a scalar product for which $\operatorname{Im} P$ is orthogonal to $\ker P$, and choose some unit vector $y$ in $\operatorname{Im} P$. Then for any $x$, $Px = \langle x, y\rangle y$. Each $M$ in $S$ has an eigenvalue of modulus one (because $1 \le \|M^p\| \le c$ for each integer $p$). So, for $M_1$, $M_2$, and $M$ in $S$, $PM_1MM_2P$ has such an eigenvalue and the relation

$$PM_1MM_2Px = \langle M_1MM_2y, y\rangle\langle x, y\rangle y \tag{7}$$

leads to

$$\langle M_1MM_2y, y\rangle = \pm 1. \tag{8}$$

By irreducibility we can find $M_1, \dots, M_{2d}$ in $T$ s.t. $\{u_1 = {}^tM_1y, \dots, u_d = {}^tM_dy\}$ and $\{v_1 = M_{d+1}y, \dots, v_d = M_{2d}y\}$ are two bases of $\mathbb{R}^d$. By (8) the set $\{\langle Mv_i, u_j\rangle;\ M \in T, 1 \le i, j \le d\}$ is finite so $T$ must be finite. Since this implies that $T = S$, $P$ is in $T$, proving the (a) of the proposition. □

PROOF OF (b). We now suppose that each element of $T$ is invertible. Let $S$ be the closure of $T$ in $\mathscr{M}(d)$ and

$$p := \inf\{\operatorname{rank}(M);\ M \in S\}.$$

For any $M$ in $\mathscr{M}(d)$ we write $\Lambda^p M$ for the endomorphism of $\Lambda^p\mathbb{R}^d$ defined by, if

$x_1, \dots, x_p$ are in $\mathbb{R}^d$,

$$(\Lambda^p M)(x_1 \wedge x_2 \wedge \cdots \wedge x_p) = Mx_1 \wedge Mx_2 \wedge \cdots \wedge Mx_p.$$

Let $\{e_1, \dots, e_d\}$ be the canonical basis of $\mathbb{R}^d$. We endow $\Lambda^p \mathbb{R}^d$ with the scalar product for which $e_{i(1)} \wedge \cdots \wedge e_{i(p)}$, $i(1) < i(2) < \cdots < i(p)$, are orthonormal vectors. If we work with the associated norm, then

$$1 \leq \|\Lambda^p M\| \leq c^p \tag{9}$$

for any $M$ in $S$. The right-hand side inequality is a consequence of the fact that $\|\Lambda^p M\| \leq \|M\|^p$; and if for some $M$ in $S$, $\|\Lambda^p M\| < 1$, then any limit point $\tilde{M}$ of the sequence $\{M^n, n \in \mathbb{N}\}$ satisfies $\Lambda^p \tilde{M} = 0$. But this is equivalent to $\operatorname{rank}(\tilde{M}) < p$ and thus cannot hold.

Let $Q$ be an element of $S$ with $\operatorname{rank}(Q) = p$. If $v_1, \dots, v_p$ is a basis of $\operatorname{Im} Q$, $\operatorname{Im} \Lambda^p Q$ is the one dimensional subspace $L(w_0)$ of $\Lambda^p \mathbb{R}^d$ spanned by $w_0 := v_1 \wedge v_2 \wedge \cdots \wedge v_p$. The spectral radius of $\Lambda^p Q$ is one by (9); thus if $P = Q^2$, $\Lambda^p P$ is a projection on $L(w_0)$. We know (see Chevalley [2], Chapter IV, Section 5) that since $T$ is irreducible, its action on $\Lambda^p \mathbb{R}^d$ is semisimple, i.e., there exists a direct sum decomposition $\Lambda^p \mathbb{R}^d = W_1 \oplus W_2 \oplus \cdots \oplus W_r$ such that, for each $i$, $1 \leq i \leq r$,

(i) $(\Lambda^p M)(W_i) \subset W_i$, for any $M$ in $T$.

(ii) There is no proper subspace $W$ of $W_i$ such that $(\Lambda^p M)(W) \subset W$, for all $M$ in $T$.

Since for each $i$ $(\Lambda^p P)(W_i)$ is contained in $W_i$ we may choose $W_1$ such that $w_0$ is in $W_1$.

As in (8), for some scalar product on $\Lambda^p \mathbb{R}^d$, if $M$, $M_1$, and $M_2$ are in $T$

$$\langle (\Lambda^p M_1)(\Lambda^p M)(\Lambda^p M_2) w_0, w_0 \rangle = \pm 1$$

and, as above, by the irreducibility property (ii), the set whose elements are the restrictions of $\Lambda^p M$, $M \in T$, to $W_1$ is finite. Since $P$ is in the closure of $T$, there exists some $M$ in $T$ such that

$$\Lambda^p P w = \Lambda^p M w, \quad \text{for all } w \text{ in } W_1.$$

But $M$ and $\Lambda^p M$ are invertible and $\operatorname{Im} \Lambda^p P = L(w_0)$, so $W_1 = L(w_0)$. By (i) the linear span of $v_1, \dots, v_p$ in $\mathbb{R}^d$ is thus $T$-invariant. The irreducibility of $T$ implies that $p$ is equal to $d$. Therefore each element of $S$ is invertible, and $S$ is a compact cancellative semigroup. $S$ is thus a compact group (see (9.16) of Hewitt and Ross [10]). Since $T$ is contained in $S$ the proposition is proved. □

REMARK 2.3. It is not difficult to modify the proof in order to obtain that the same proposition holds if, instead of requiring that $1 \leq \|M\| \leq c$ for $M$ in $T$, we only suppose that the spectral radius of each element of $T$ is one. The first modification is to define $p$ as the least integer $n$ such that for some $M$ in $T$ the dimension of the direct sum of the generalized eigenspaces associated with an eigenvalue of modulus one is $n$. As above one shows that $p = d$, so that all the eigenvalues of each element in $T$ have modulus one. It thus follows from the

Lemma 1 in Conze and Guivarc'h [3] (stated for groups but actually valid for semigroups as well with the same proof) that $T$ is contained in a compact group.

REMARK 2.4. The proof does not work for matrices with complex entries but part (b) remains true in this case. To verify this, write each matrix $M$ in $\mathrm{Gl}(d, \mathbb{C})$ as $M = A + iB$ with $A$ and $B$ in $\mathrm{Gl}(d)$.

Define $\varphi$: $\mathrm{Gl}(d, \mathbb{C}) \to \mathrm{Gl}(2d)$ by

$$\varphi(A + iB) = \begin{pmatrix} A & -B \\ B & A \end{pmatrix}$$

and put $J = \varphi(iI)$. If $T$ is a semigroup contained in $\mathrm{Gl}(d, \mathbb{C})$ which acts irreducibly on $\mathbb{C}^d$, the semigroup $T'$ of $\mathrm{Gl}(2d)$ generated by $\varphi(T)$ and $J$ is $\varphi(T) \cup -\varphi(T) \cup J\varphi(T) \cup -J\varphi(T)$. This implies that $T$ is irreducible. If $1 \leq \|M\| \leq c$ for each $M$ in $T$, the same holds for each $M$ in $T'$. From the proposition we conclude that $T'$ is in a compact subgroup of $\mathrm{Gl}(2d)$. Therefore $T$ is contained in a compact subgroup of $\mathrm{Gl}(d, \mathbb{C})$.

REMARK 2.5. If $T$ is an $F$-subsemigroup of $\mathscr{M}(d)$, it contains by definition a rank-one projection $P$. Unless $d = 1$, $-P$ is also in $T$. (By (7), for $M$, $M_1$, and $M_2$ in $T$, $PM_1MM_2P$ is equal to $P$ or $-P$. If $-P$ is not in $T$, we have instead of (8), $\langle M_1MM_2y, y\rangle = 1$ and, as above, we can find two bases $(u_i)$ and $(v_j)$ such that $\langle u_i, v_j\rangle = 1$ for $1 \leq i, j \leq d$. This of course can hold only if $d = 1$.)

Notice that if 1 is an eigenvalue of each matrix in $T$ then $d$ must be equal to 1.

We can now prove our main result on irreducible probability measures.

THEOREM 2.6. *Let $\mu$ be an irreducible probability measure on $\mathscr{M}(d)$ such that one of the following assumptions is true*:

(i) *$\mu$ satisfies* C1 *and $T(\mu)$ is not contained in a compact subgroup of* $\mathrm{Gl}(d)$;

*or*

(ii) *$\mu$ satisfies* C2 *and $T(\mu)$ is not an $F$-semigroup.*

*Then*:

(a) *The unique $\mu$-invariant measure on $\mathbb{R}^d$ is the Dirac measure at* 0.

(b) *If $\{\mu^n, n \in \mathbb{N}\}$ is tight on $\mathscr{M}(d)$, then $(1/n)\Sigma_{i=1}^n \mu^i$ converges to the Dirac measure at the zero matrix.*

For unimodular matrices (a) is proved in Furstenberg [5].

PROOF. Let $\nu$ be an $\mu$-invariant probability measure on $\mathbb{R}^d$ and $H$ the linear subspace spanned by the support of $\nu$. By definition

$$\iint 1_H(Yx)\, d\mu(Y)\, d\nu(x) = \nu(H) = 1;$$

hence, for $\mu$-almost all $Y$, $\nu\{x \in \mathbb{R}^d;\ Yx \in H\} = 1$ and $Y(H)$ is contained in $H$. The irreducibility of $\mu$ implies that $H = \{0\}$ or $\mathbb{R}^d$. It follows from Propositions 2.1 and 2.2 that $H$ cannot be $\mathbb{R}^d$, so $H = \{0\}$ and $\nu$ is the Dirac measure at 0, proving (a).

It is clear that if $\{\mu^n, n \in \mathbb{N}\}$ is tight, $\{(1/n)\sum_{i=1}^n \mu^i, n \in \mathbb{N}\}$ is tight too. Under this hypothesis consider a limit point $m$ of $(1/n)\sum_{i=1}^n \mu^i$, we have

$$\mu * m = \lim \mu * \frac{1}{n}\sum_{i=1}^{n} \mu^i = \lim \frac{1}{n}\sum_{i=1}^{n} \mu^i + \lim \frac{1}{n}\{\mu^{n+1} - \mu\} = m.$$

For any $x$ in $\mathbb{R}^d$, denote by $m_x$ the distribution of $Mx$ if $M$ is a random matrix with distribution $m$. Since $\mu * m = m$, $m_x$ is a $\mu$-invariant probability measure on $\mathbb{R}^d$. From (a) we deduce that $m_x = \delta_0$ for all $x$, so $m$ is the Dirac measure at the zero matrix. □

Let $M_n = Y_n \cdots Y_1$, where $Y_1, Y_2, \dots$ are i.i.d. matrices with distribution $\mu$. The following corollary is in particular useful in the study of the central limit theorem for products of random matrices (see for instance the proof of the proposition below). It implies that $\sup_n(\log^2\|M_n x\|) = \infty$ for any $x \neq 0$ of $\mathbb{R}^d$, which settles a question of Hashminski ([8], page 244).

COROLLARY 2.7. *Under the hypothesis* (i) *or* (ii) *of Theorem* 2.6, *for any* $x \neq 0$ *in* $\mathbb{R}^d$, *the sequence of the distributions of* $\log\|M_n x\|$, $n = 1, 2, \dots$, *is not tight on* $\mathbb{R}$.

PROOF. Suppose that this sequence is tight. If $\mu_x^n$ is the distribution of $M_n x$, the sequence $\{(1/n)\sum_{i=1}^n \mu_x^i, n \in \mathbb{N}\}$ is tight on $\mathbb{R}^d$. It is easy to see that any limit point of this sequence is a $\mu$-invariant probability measure; therefore by (a) of the theorem $(1/n)\sum_{i=1}^n \mu_x^i$ converges to $\delta_0$. This contradicts the tightness of the sequence $\{\log\|M_n x\|, n \in \mathbb{N}\}$. □

The condition "$(1/n)\sum_{i=1}^n \mu^i$ converges to the Dirac mass at the zero matrix," which appears in Theorem 2.6, is not easy to handle. It would be nice to replace it, under moments conditions, by: "the upper Liapounov exponent $\gamma(\mu)$ is strictly negative." Since $(1/n)\log\|M_n\|$ converges a.s. to $\gamma(\mu)$ it is clear that $\gamma(\mu) \leq 0$. But even for $1 \times 1$ matrices one can have $(1/n)\sum_{i=1}^n \mu^i \to \delta_0$ and $\gamma(\mu) = 0$. There exists a sequence $(X_n)$ of i.i.d. real random variables with $E(\|X_1\|)$ finite and $E(X_1) = 0$ such that, for any real $c$,

$$P(X_1 + X_2 + \cdots + X_n < c) \to 1 \quad \text{as } n \to +\infty$$

(see the unfavorable fair game in Feller [4], Example 15, page 262). If we set $Y_n = \exp(X_n)$, $Y_n$ is a $1 \times 1$ matrix and if $\mu$ is its distribution, $\gamma(\mu) = 0$ but $\mu^n$ converges to $\delta_0$.

Nevertheless we have the following result which settles in particular the case of linear SDE (see Section 7).

**Proposition 2.8.** *Let $\mu$ be a probability measure on $\mathscr{M}(d)$ such that:*

(i) $\mu(\mathrm{Gl}(d)) = 1$.

(ii) *There does not exist a finite union $U$ of proper subspaces of $\mathbb{R}^d$ such that $M(U) \subset U$ for each $M$ in $T(\mu)$.*

(iii) *For some $a > 0$, $\int \|M\|^a \, d\mu(M)$ and $\int \|M^{-1}\|^a \, d\mu(M)$ are finite.*

*Then $(1/n)\Sigma_{i=1}^n \mu^i$ converges to the Dirac mass at the zero matrix if and only if $\gamma(\mu)$ is strictly negative.*

Proof. Since $(1/n)\log\|M_n\|$ converges to $\gamma(\mu)$ a.s. we have only to prove that if $\gamma(\mu) = 0$, $(1/n)\Sigma_{i=1}^n \mu^i$ cannot converge to the Dirac measure at the zero matrix. Let us introduce all the Liapounov exponents associated with $\mu$, i.e., the reals $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d$ defined by, for $1 \leq p \leq d$:

$$\lambda_1 + \lambda_2 + \cdots + \lambda_p = \lim_{n \to \infty} \frac{1}{n} \log\|\Lambda^p M_n\|$$

(see Ledrappier [14]). We shall suppose that $\gamma(\mu) = 0$, i.e., $\lambda_1 = 0$. We consider three cases.

(a) *First case:* $\lambda_1 \neq \lambda_2$. Since $\mu$ satisfies (ii) and (iii) it follows from the central limit theorem of Le Page ([15], Theorem 2) that in this case, for each $x \neq 0$, the sequence $n^{-1/2}\log\|M_n x\|$ converges in distribution to some normal law $\mathscr{N}(0, \sigma^2)$.

If $\sigma^2 = 0$, then by Propriété 1 of ([15], page 278) there exists some $x \neq 0$ and $C > 0$ such that

$$|\log\|M_n x\|| < C \quad \text{a.s. for each integer } n.$$

By Corollary 2.7 this can hold only if $T(\mu)$ is contained in a compact subgroup of $\mathrm{Gl}(d)$, which would imply that $\lambda_1 = \lambda_2$.

Therefore $\sigma^2 \neq 0$ and for each $x \neq 0$

$$\liminf_{n \to \infty} P(\|M_n\| \geq 1) \geq \lim_{n \to \infty} P(\log\|M_n x\| \geq 0) = \tfrac{1}{2}.$$

It is thus clear that $(1/n)\Sigma_{i=1}^n \mu^i$ does not converge to the Dirac measure at the zero matrix.

(b) *Second case:* $\lambda_1 = \lambda_2 = \cdots = \lambda_p \neq \lambda_{p+1}$ for some $p \in \{2, \ldots, d-1\}$. Let $\tilde{\mu}$ be the image of $\mu$ under the mapping $\psi$: $\mathrm{Gl}(d) \to \mathrm{Aut}(\Lambda^p \mathbb{R}^d)$ defined by $\psi(M) = \Lambda^p M$. If $\tilde{\lambda}_1$ and $\tilde{\lambda}_2$ are the two largest Liapounov exponents associated with $\tilde{\mu}$, $\tilde{\lambda}_1 = \lambda_1 + \lambda_2 + \cdots + \lambda_p = 0$ and $\tilde{\lambda}_2 = \lambda_1 + \lambda_2 + \cdots + \lambda_{p-1} + \lambda_{p+1} < 0$. As in the proof of Proposition 2.2 we may write $\Lambda^p \mathbb{R}^d = W_1 \oplus W_2 \oplus \cdots \oplus W_r$ where each $W_j$ is a subspace invariant under $\{\Lambda^p M, M \in T(\mu)\}$ and where, if $\mu_j$ is the image of $\tilde{\mu}$ under the restriction to $W_j$, each $\mu_j$ is an irreducible probability measure on $\mathrm{End}(W_j)$. It is clear that for some $k$ in $\{1, \ldots, r\}$ the upper Liapounov exponent of $\mu_k$ is $\tilde{\lambda}_1 = 0$. If $(1/n)\Sigma_{i=1}^n \mu^i$ converges to the Dirac measure at the zero matrix, $(1/n)\Sigma_{i=1}^n \mu_k^i$ converges to the Dirac measure at the zero element of $\mathrm{End}(W_k)$. Either $W_k$ is one dimensional and we are led to a contradiction using the usual central limit theorem, or $\dim(W_k) \geq 2$ and the second upper exponent associated with $\mu_k$ is no larger than $\tilde{\lambda}_2$, hence, strictly negative. We arrive at a contradiction by applying (a) to $\mu_k$.

(c) *Third case:* $\lambda_1 = \lambda_2 = \cdots = \lambda_d$. Define $\psi$: $\mathrm{Gl}(d) \to \mathrm{Gl}(d)$ by $\psi(M) = |\det M|^{-1/d}M$. By Theorem 8.6 of Furstenberg [5] there exists a compact subgroup $K$ of $\mathrm{Gl}(d)$ such that $\psi(Y)$ is in $K$ for $\mu$-almost all $Y$. This implies that for a suitable norm on $\mathscr{M}(d)$, $\|\psi(M_n)\| = 1$ a.s. for all $n \geq 1$. Therefore $\|M_n\|^d = |\det M_n|$ and $n^{-1/2}d$. $\mathrm{Log}\|M_n\| = n^{-1/2}\sum_{i=1}^{n} \log|\det Y_i|$. We conclude immediately with the usual central limit theorem. □

**3. A necessary condition for tightness of $\{\mu^n, n \geq 1\}$.** In this section we consider a general probability measure on $\mathscr{M}(d)$ satisfying C1 or C2 (see Definition 1.6). We shall show that if $\{\mu^n, n \geq 1\}$ is tight, $\mu$ must have a very particular form. Namely we have, with the notation introduced in Definition 1.3:

THEOREM 3.1. *Let $\mu$ be a probability measure on $\mathscr{M}(d)$ for which either condition* C1 *or condition* C2 *holds. If $\{\mu^n, n \geq 1\}$ is tight on $\mathscr{M}(d)$ there exist three nonnegative integers $d_1, d_2, d_3$, $d_1 + d_2 + d_3 = d$, such that*:

(i) *For some invertible matrix $Q$, $QT(\mu)Q^{-1}$ is contained in $T(d_1; d_2; d_3)$.*

(ii) *If $\mu_a$ (resp. $\mu_b$) is the image of $\mu$ on $\mathscr{M}(d_1)$ (resp. $\mathscr{M}(d_3)$) under the mapping $f_a$: $\mathscr{M}(d) \to \mathscr{M}(d_1)$ defined by $f_a(M) = a(QMQ^{-1})$ (resp. under $f_b$: $\mathscr{M}(d) \to \mathscr{M}(d_3)$ defined by $f_b(M) = b(QMQ^{-1})$), then $(1/n)\sum_{i=1}^{n}\mu_a^i$ (resp. $(1/n)\sum_{i=1}^{n}\mu_b^i$) converges to the Dirac measure at the zero matrix of $\mathscr{M}(d_1)$ (resp. $\mathscr{M}(d_3)$).*

(iii) *Under* C1, *$\{k(QMQ^{-1}), M \in T(\mu)\}$ is contained in the orthogonal group $O(d_2)$.*

(iv) *Under* C2, *$\{k(QMQ^{-1}), M \in T(\mu)\}$ is an F-subsemigroup of $\mathscr{M}(d_2)$.*

The proof of this theorem under C2 is easy and will be given at the end of this section. Under C1 the proof is more intricate and we begin with two lemmas.

LEMMA 3.2. *Let $\mu$ be a probability measure on $\mathscr{M}(d)$ such that $\{(1/n)\sum_{i=1}^{n}\mu^i, n \geq 1\}$ is tight on $\mathscr{M}(d)$. The sequence $(1/n)\sum_{i=1}^{n}\mu^i$ converges to a probability measure $m$ such that*

$$\mu * m = m * \mu = m * m = m. \tag{10}$$

*Moreover the linear space $\{x \in \mathbb{R}^d;\ Mx = 0$ for $m$-almost all $M\}$ is $T(\mu)$-invariant.*

PROOF OF THE LEMMA. This result is well known and easy to verify: If $m$ and $m'$ are two limit points of $\{(1/n)\sum_{i=1}^{n}\mu^i, n \geq 1\}$ we have (see the proof of Theorem 2.6) $\mu * m = m$, hence $m' * m = m$, and $m' * \mu = m'$, hence $m' * m = m'$. So $m' = m$ and (10) holds. For $W = \{x \in \mathbb{R}^d;\ Mx = 0$ for $m$-almost all $M\}$, the relation $m * \mu = m$ gives that if $x$ is in $W$, $MYx = 0$ for $\mu \otimes m$-almost all $(Y, M)$. Therefore $Yx$ is in $W$ for $\mu$-almost all $Y$ and $\{M \in \mathscr{M}(d);\ MW \subset W\}$ is a closed semigroup in $\mathscr{M}(d)$ of $\mu$-probability one. It must contain $T(\mu)$. □

Before stating the next lemma, let us introduce some notation. It will be more convenient to deal with endomorphisms rather than matrices.

NOTATION 3.3. Let $E$ be a finite dimensional vector space. If $M$ is an endomorphism of $E$ and $F$ a subspace invariant under $M$ (i.e., $MF \subset F$) we denote by $M_F$ the endomorphism of $F$ defined as the restriction of $M$ to $F$ (i.e., $M_F x = Mx$ for any $x$ in $F$).

As usual $\mathrm{End}(E)$ is the set of endomorphisms of $E$, $\mathrm{Aut}(E)$ the set of automorphisms. The next lemma is the key for the proof of the theorem under C1.

LEMMA 3.4. *Let $\mu$ be a probability measure on $\mathrm{End}(E)$ such that $\mu(\mathrm{Aut}(E)) = 1$ and such that $(1/n)\sum_{i=1}^{n}\mu^i$ converges to a probability measure $m$ on $\mathrm{End}(E)$. We suppose that for any $x \neq 0$ in $E$,*

$$m\{M \in \mathrm{End}(E);\ Mx = 0\} \neq 1. \tag{11}$$

*Let $F_1$ be a $T(\mu)$-invariant subspace of $E$ with maximal dimension such that $\{M_{F_1};\ M \in T(\mu)\}$ is contained in a compact subgroup of $\mathrm{Aut}(F_1)$. Then, for any $x$ in $E$, $m\{M \in \mathrm{End}(E);\ Mx \in F_1\} = 1$.*

PROOF. (a) Let $G_1$ be a subspace of $E$ such that $F_1 \oplus G_1 = E$ and $F_2 = \{x \in G_1;\ Mx \in F_1$ for $m$-almost all $M\}$. We want to prove that $F_2 = G_1$. If this does not hold we may consider a subspace $G_2$ such that $F_1 \oplus F_2 \oplus G_2 = E$ and a subspace $F_3 \neq \{0\}$ of $G_2$ of minimal dimension s.t.

$$MF_3 \subset F_1 \oplus F_2 \oplus F_3 \quad \text{for all } M \text{ in } T(\mu).$$

Notice that $F = F_1 \oplus F_2 \oplus F_3$ is $T(\mu)$-invariant. If $\mu$ satisfies the hypotheses of the lemma, then so does its image under the map which sends each $M$ in $T(\mu)$ to $M_F$. Looking for a contradiction we can (and shall) reduce the study to the case where $E = F$. Now if $E = F$, in a basis compatible with the direct sum decomposition $E = F_1 \oplus F_2 \oplus F_3$ we can write each $M$ in $T(\mu)$ as

$$M = \begin{pmatrix} a_{11}(M) & a_{12}(M) & a_{13}(M) \\ 0 & a_{22}(M) & a_{23}(M) \\ 0 & 0 & a_{33}(M) \end{pmatrix}$$

where, for $d_i = \dim F_i$, $a_{ij}(M)$ is a $d_i \times d_j$ matrix. I claim that $\{a_{33}(M),\ M \in T(\mu)\}$ is contained in a compact subgroup of $\mathrm{Gl}(d_3)$. The image $\mu_3$ of $\mu$ under $a_{33}$ is an irreducible probability measure on $\mathscr{M}(d_3)$, carried by $\mathrm{Gl}(d_3)$, such that $(1/n)\sum_{i=1}^{n}\mu_3^i$ converges to the image $m_3$ of $m$ under $a_{33}$. If the claim were not true, by Theorem 2.6, $m_3$ would be carried by the zero matrix. Since $m * m = m$ this would imply that, for any $x$ in $F_3$,

$$m\{M \in \mathrm{End}(E);\ Mx = 0\} = m \otimes m\{(M_1, M_2);\ M_1 M_2 x = 0\} = 1,$$

which contradicts the definition of $F_2$. We can of course suppose that the bases of $F_1$ and $F_3$ are chosen in such a way that $a_{11}(M)$ is in $O(d_1)$ and $a_{33}(M)$ in $O(d_3)$ for each $M$ in $T(\mu)$.

(b) The next step is to prove that, for each $r > 0$,

$$S_r = \{M \in T(\mu);\ \|a_{22}(M)\| \le \tfrac{1}{2},\ \|a_{12}(M)\| \le r,\ \|a_{23}(M)\| \le r\}$$

is a bounded subset of End($E$). An easy way to show this is to mimick the proof of Proposition 2.1: Let $Y_1, Y_2, \dots$ be independent matrices with distribution $\mu$ and $M_n = Y_n \cdots Y_1$. For any $M$ in End($E$) we denote by $mM$ the image of $m$ under the right multiplication by $M$. For almost all $\omega$, the sequence $\{mM_n(\omega), n \ge 1\}$ converges to a probability measure $m_\omega$ on End($E$). This implies that $\{M_n(\omega), n \ge 1\}$ is a.s. bounded (if not we could find, as in Proposition 2.1, a nonzero matrix $H(\omega)$ s.t. $m\{M;\ MH(\omega) = 0\} = 1$, which contradicts (11)). For any limit point $M(\omega)$ of $\{M_n(\omega), n \ge 1\}$ and each $M$ in $S_r$

$$mMM(\omega) = m_\omega \tag{12}$$

(see the analogue (6)). If $S_r$ is not bounded we can find a sequence $M_p$ in $S_r$ such that $\|M_p\|^{-1}M_p$ converges to a nonzero matrix $A$ with $a_{ij}(A) = 0$ if $(i, j) \ne (1,3)$. As in Proposition 2.1, (12) implies that $m\{M;\ MAM(\omega) = 0\} = 1$ and $AM(\omega) = 0$. Carrying out the matrix multiplication we find that $A$ is equal to zero, which is not true.

(c) Choose some matrix $A$ in the support of $m$. Since $a_{22}(A) = 0$, there is an $r$ such that $S := \{A^n, n \ge 1\}$ is contained in $S_r$. As $S_r$ is bounded the closure of $S$ is a compact semigroup contained in $T(\mu)$. Therefore $S$, and thus $T(\mu)$, contains a projection $P$ (see, e.g., Hewitt and Ross [10], (9.18)). Notice that $a_{22}(P)$ is the zero matrix. Consider the semigroup $T = PT(\mu)P$. It is easy to see that $T$ is closed and contained in $S_r$, hence compact. Let $V$ be the range of $P$ and let $T' = \{M_V;\ M \in T\}$. For each $M$ in $T(\mu)$ the eigenvalues of $a_{11}(PMP)$ and $a_{33}(PMP)$ are nonzero, whence dim(ker $PMP$) is equal to $d_2$. Since ker $P$ is contained in ker $PMP$ this yields that ker $P$ = ker $PMP$, so that $M_V$ is one to one. Thus $T'$ is a compact semigroup in Aut($V$), and (see Hewitt and Ross [10], (9.16) and (22.23)) we can find a scalar product $\langle \cdot, \cdot \rangle$ on $E$ such that:

(i) $V$ is orthogonal to ker $P$.
(ii) For $M$ in $T(\mu)$, $PMP$ acts on $V$ as an isometry.

(d) Let $W = \{y \in E;\ \langle My, x\rangle = 0,\ \forall\ M \in T(\mu),\ \forall\ x \in F_1\}$. This subspace is $T(\mu)$-invariant. If $U$ is the orthogonal of $F_1$ in $V$ (note that $F_1 \subset V$), $\dim(U) = d_3$ is not zero and by (ii), $PMP(U) \subset U$ for $M$ in $T(\mu)$. This implies that if $x \in F_1$, $y \in U$, and $M \in T(\mu)$,

$$\langle My, x\rangle = \langle MPy, Px\rangle = \langle PMPy, x\rangle = 0.$$

Thus $W$ contains $U$ and $W \ne \{0\}$. We easily see, as at the end of (a), that $\{M_W;\ M \in T(\mu)\}$ is in a compact subgroup of Aut($W$). Since $F_1$ and $W$ are orthogonal and $T(\mu)$-invariant this implies that $\{M_{F_1 \oplus W};\ M \in T(\mu)\}$ is in a compact subgroup of Aut($F_1 \oplus W$) which contradicts the maximality assumption on $F_1$. □

Proof of Theorem 3.1 under Condition C1. Since $\{\mu^n;\ n \ge 1\}$ is tight, $(1/n)\sum_{i=1}^n \mu^i$ converges to a distribution $m$. Let $V = \{x \in \mathbb{R}^d;\ Mx = 0$ for $m$-

almost all $M\}$. For any $M$ in $T(\mu)$ let $\overline{M}$ be the endomorphism of $E := \mathbb{R}^d/V$ defined by, if $x$ is in $\mathbb{R}^d$ and $\bar{x}$ is its canonical image in $E$,

$$\overline{M}\bar{x} = \overline{Mx}.$$

The image $\bar{\mu}$ of $\mu$ under $M \to \overline{M}$ has the properties required in Lemma 3.4. (By C1 $\bar{\mu}$ is carried by $\mathrm{Aut}(E)$ and if $\overline{m}$ is the image of $m$ under $M \to \overline{M}$, $1/n\Sigma_{i=1}^n \bar{\mu}^i \to \overline{m}$. If for some $\bar{x} \neq 0$ in $E$,

$$\overline{m}\{\overline{M} \in \mathrm{End}(E);\ \overline{M}\bar{x} = 0\} = 1,$$

then

$$m\{M \in \mathscr{M}(d);\ Mx \in V\} = 1$$

and, since $m * m = m$,

$$m\{M;\ Mx = 0\} = m \otimes m\{(M_1, M_2);\ M_1M_2x = 0\} = 1,$$

which contradicts the definition of $V$.)

By Lemma 3.4 we can write $E = F_1 \oplus F_2$, where:

($\alpha$) $F_1$ is invariant under $\{\overline{M} \in \mathrm{End}(E);\ M \in T(\mu)\}$ and $\{\overline{M}_{F_1} \in \mathrm{End}(F_1);\ M \in T(\mu)\}$ is in a compact subgroup of $\mathrm{Aut}(F_1)$.

($\beta$) For $\bar{x}$ in $F_2$, $\overline{M}\bar{x}$ is in $F_1$ for $m$-almost all $M$.

Consider now a basis $\{f_i, 1 \le i \le d\}$ of $\mathbb{R}^d$ such that $\{f_i, 1 \le i \le d_1\}$ is a basis of $V$, $\{f_i, d_1 < i \le d_1 + d_2\}$ a basis of $F_1$, and $\{f_i, d_1 + d_2 < i \le d\}$ is a basis of $F_2$. If $\{e_i, 1 \le i \le d\}$ is the canonical basis of $\mathbb{R}^d$ and $Q$ the matrix defined by $Q(f_i) = e_i$, $1 \le i \le d$, it is clear that $QT(\mu)Q^{-1}$ is contained in $T(d_1; d_2; d_3)$. The property (ii) is a consequence of the definition of $V$ and of ($\beta$) above. By ($\alpha$) we may choose the basis $\{f_i, d_1 < i \le d_1 + d_2\}$ in such a way that (iii) holds. □

**Proof of Theorem 3.1 under Condition C2.** Let $\{0\} = E_0 \subset E_1 \subset \cdots \subset E_r = \mathbb{R}^d$ be a sequence of $T(\mu)$-invariant subspaces of $\mathbb{R}^d$ such that $E_p$ contains strictly $E_{p-1}$, for $p = 1, \dots, r$. For $M$ in $T(\mu)$ denote by $s_p(M)$ the endomorphism of $F_p := E_p/E_{p-1}$ defined by $s_p(M)\bar{x} = \overline{Mx}$ (if $x$ is in $E_p$ and $\bar{x}$ is the class of $x$ in $F_p$). Consider $m = \lim (1/n)\Sigma_{i=1}^n \mu^i$ and $\mu_p$ (resp. $m_p$) the image of $\mu$ (resp. $m$) under $s_p$. By choosing inductively each $E_p$ as a $T(\mu)$-invariant subspace of minimal dimension, we may suppose that each $\mu_p$ is an irreducible probability measure on $\mathrm{End}(F_p)$; it satisfies, of course, C2. Therefore by Theorem 2.6, either $m_p$ is the Dirac measure at the null endomorphism of $\mathrm{End}(F_p)$ (and this case occurs if and only if some $s_p(M)$, $M \in T(\mu)$, has no eigenvalue of modulus one) or $\{s_p(M);\ M \in T(\mu)\}$ is a finite $F$-semigroup. By Condition C2 we know that there exists a matrix $M_0$ in $T(\mu)$ with at most one eigenvalue of modulus one, this eigenvalue being simple. Hence, there exists at most one integer $q$, $1 \le q \le r$, such that $m_q$ is not carried by the zero matrix. It follows from the relation $m * m = m$ that if $Q$ is an invertible matrix sending $E_{q-1}$ onto the subspace generated by $e_1, \dots, e_{d_1}$ (where $e_i$, $1 \le i \le d$, is the canonical basis of $\mathbb{R}^d$ and $d_1$ the dimension of $E_{q-1}$) and $E_q$ onto the subspace generated by $e_1, \dots, e_{d_1+d_2}$ (where $d_1 + d_2$ is the dimension of $E_q$), the theorem holds with this $Q$. □

REMARK 3.5. From the above proof it is clear that if there exists in $T(\mu)$ a matrix with no eigenvalue of modulus one, $\{(1/n)\Sigma_{i=1}^{n}\mu^{i},\ n \in \mathbb{N}\}$ is tight on $\mathscr{M}(d)$ if and only if this sequence converges to the Dirac measure at the zero matrix.

REMARK 3.6. Suppose that the spectral radius of each matrix in $T(\mu)$ is an eigenvalue (this is true for instance if $T(\mu)$ is contained in the set of nonnegative matrices). Under C2 if $\{(1/n)\Sigma_{i=1}^{n}\mu^{i},\ n \in \mathbb{N}\}$ is tight and $d_2 \neq 0$, then $d_2 = 1$ and $k(QMQ^{-1}) = 1$ for each $M$ in $T(\mu)$. To prove this, consider a matrix $M_0$ in $T(\mu)$ such that $k(QM_0Q^{-1})$ is the projection $P$ introduced in Definition 1.2. By the Remark 2.5 it suffices to prove that $-P$ is not in $\{k(QMQ^{-1});\ M \in T(\mu)\}$. Suppose that for some $M_1$ in $T(\mu)$, $k(QM_1Q^{-1}) = -P$ and consider a matrix $M_2$ in the support of $\lim(1/n)\Sigma_{i=1}^{n}\mu^{i}$. Se have $a(QM_2Q^{-1}) = 0$, $b(QM_2Q^{-1}) = 0$, and either $k(QM_1M_2M_0Q^{-1})$ or $k(QM_0M_2M_1Q^{-1})$ is equal to $-P$ (use (7) and (8)). Therefore one of the matrices $M_1M_2M_0$ or $M_0M_2M_1$ has the only eigenvalues 0 and $-1$, which contradicts the hypothesis.

If $T(\mu)$ is included in the set of nonnegative matrices and contains a positive one, this result and much more has been proved by Kesten and Spitzer in [13].

COROLLARY 3.7. *If* $\mu$ *fulfills the conditions of Theorem* 3.1, $\lim_{n\to\infty}(1/n)\Sigma_{i=1}^{n}\mu^{i}$ *is the law of a random matrix*

$$Q^{-1}\begin{pmatrix} 0 & C_1K_2K_3 & C_1K_2D_3 \\ 0 & K_1K_2K_3 & K_1K_2D_3 \\ 0 & 0 & 0 \end{pmatrix}Q, \tag{13}$$

*where* $(C_1, K_1)$, $K_2$, *and* $(K_3, D_3)$ *are independent and* $C_1 \in \mathscr{M}(d_1, d_2)$, $K_1, K_2, K_3 \in \mathscr{M}(d_2)$, $D_3 \in \mathscr{M}(d_2, d_3)$.

*If moreover* $K := \{k(QMQ^{-1});\ M \in T(\mu)\}$ *is contained in a compact subgroup of* $\mathrm{Gl}(d_2)$ *then we may choose* $K_1$ *and* $K_3$ *above as the identity matrix, and the distribution of* $K_2$ *is then the Haar measure on* $K$.

PROOF. Let $m = \lim(1/n)\Sigma_{i=1}^{n}\mu^{i}$. Since $m * m * m = m$ (see (10)), $m$ is the distribution of $M_1M_2M_3$ where $M_1$, $M_2$ and $M_3$ are three independent matrices with distribution $m$. Using the relations

$$a\big(QM_iQ^{-1}\big) = 0, \qquad b\big(QM_iQ^{-1}\big) = 0 \quad \text{for } i = 1,2,3,$$

and carrying out the matrix multiplication we find (13) with

$$C_1 = c\big(QM_1Q^{-1}\big), \qquad K_i = k\big(QM_iQ^{-1}\big), \qquad D_3 = d\big(QM_3Q^{-1}\big).$$

If $K$ is in a compact subgroup of $\mathrm{Gl}(d_2)$, $K$ is itself a compact group (being a compact cancellative semigroup). Since the distribution of the $K_i$'s is $\lim_{n\to\infty}(1/n)\Sigma_{i=1}^{n}\mu_k^{i}$, where $\mu_k$ is the image of $\mu$ under $M \mapsto k(QMQ^{-1})$ it is the Haar measure on $K$ (see [16]). By invariance of the Haar measure, $(C_1K_2K_3,\ K_1K_2K_3,\ C_1K_2D_3,\ K_1K_2D_3)$ has the same distribution as $(C_1K_1^{-1}K_2,\ K_2,\ C_1K_1^{-1}K_2K_3^{-1}D_3,\ K_2K_3^{-1}D_3)$. So it is clear that $m$ is the distribution of a matrix of the form (13) with $K_1 = K_2 = I$ (one takes new $C_1$ and $D_3$ as being the old $C_1 K_1^{-1}$ and $K_3^{-1} D_3$). $\square$

**4. Sufficient conditions for tightness of $\{\mu^n, n \geq 1\}$.** We may consider the following as a converse of Theorem 3.1 (see the end of Section 2).

THEOREM 4.1. *Suppose that for a probability measure $\mu$ on $\mathscr{M}(d)$ the following holds:*

(i) $\int \log^+ \|M\| \, d\mu(M) < \infty$.

(ii) *For some $d_1 \geq 0$, $d_2 \geq 0$, $d_3 \geq 0$, $T(\mu)$ is contained in $T(d_1; d_2; d_3)$. So we can write each $M$ in $T(\mu)$ as (see Definition 1.3):*

$$M = \begin{pmatrix} a(M) & c(M) & e(M) \\ 0 & k(M) & d(M) \\ 0 & 0 & b(M) \end{pmatrix}.$$

(iii) *If $\mu_a$ (resp. $\mu_b$) is the image of $\mu$ under $a$ (resp. $b$), the Liapounov exponents $\gamma(\mu_a)$ and $\gamma(\mu_b)$ are strictly negative.*

(iv) *$\{k(M);\ M \in T(\mu)\}$ is bounded.*

*Then the sequence $\{\mu^n, n \geq 1\}$ is tight on $\mathscr{M}(d)$. This sequence converges if and only if $\mu_k^n$, $n \geq 1$ converges (where $\mu_k$ is the image of $\mu$ under the mapping $k$).*

The following lemma is easily proved by induction.

LEMMA 4.2. *For $M$ in $T(d_1; d_2; d_3)$ let*

$$s(M) = \begin{pmatrix} a(M) & 0 & e(M) \\ 0 & 0 & 0 \\ 0 & 0 & b(M) \end{pmatrix}.$$

*If $Y_1, Y_2, \ldots, Y_n$ are in $T(d_1; d_2; d_3)$, then*

$$d(Y_n \cdots Y_1) = \sum_{i=1}^{n} k(Y_n \cdots Y_{i+1}) d(Y_i) b(Y_{i-1}) \cdots b(Y_1),$$

$$c(Y_n \cdots Y_1) = \sum_{i=1}^{n} a(Y_n) \cdots a(Y_{i+1}) c(Y_i) k(Y_{i-1} \cdots Y_1),$$

$$e(Y_n \cdots Y_1) = e\{s(Y_n) \cdots s(Y_1)\} + \sum_{i=1}^{n} a(Y_n) \cdots a(Y_{i+2}) c(Y_{i+1}) d(Y_i \cdots Y_1).$$

From Hennion ([9], Proposition 1) or Furstenberg and Kifer ([6], Lemma 3.6) we deduce:

LEMMA 4.3. *Under the hypotheses of Theorem 4.1, the upper Liapounov exponent of the image $\mu_s$ of $\mu$ under $s$ is strictly negative.*

PROOF OF THE THEOREM. We want to prove that $\{\mu^n, n \geq 1\}$ is tight on $\mathscr{M}(d)$. Let $(Y_n)$ be i.i.d. matrices with distribution $\mu$, $M_n := Y_n \cdots Y_1$ and $S_n := Y_1 \cdots Y_n$ have distribution $\mu^n$. If $\alpha = \sup\{\|k(M)\|; M \in T(\mu)\}$, by Lemma 4.2,

$$\|d(M_n)\| \leq \alpha \sum_{i=1}^{n} \|d(Y_i)\| \, \|b(Y_{i-1}) \cdots b(Y_1)\|.$$

Since $E\{\log^+\|d(Y_i)\|\} < \infty$, the Borel–Cantelli lemma yields that

$$\limsup \|d(Y_i)\|^{1/i} \leq 1 \quad \text{a.s.},$$

and since $\lim\|b(Y_n) \cdots b(Y_1)\|^{1/n} = \exp \gamma(\mu_b) < 1$, $d(M_n)$ is bounded with probability 1. In the same way we prove that $c(S_n)$ is a.s. bounded. So, the exponent $\gamma(\mu_s)$ being negative, we have only to check that

$$\sum_{i=1}^{n-1} a(Y_n) \cdots a(Y_{i+2}) c(Y_{i+1}) d(Y_i \cdots Y_1)$$

is tight. This sum has a norm smaller than

$$\sum_{i=1}^{n-1} \|a(Y_n) \cdots a(Y_{i+2})\| \, \|c(Y_{i+1})\| \, \|d(M_n)\|$$

and since $d(M_n)$ is a.s. bounded it suffices to show the tightness of $U_n = \sum_{i=1}^{n-1}\|a(Y_n) \cdots a(Y_{i+2})\| \, \|c(Y_{i+1})\|$. This is clear since $U_n$ has the same law as $V_n = \sum_{i=1}^{n-1}\|a(Y_1) \cdots a(Y_{i-1})\| \, \|c(Y_i)\|$ which is bounded a.s.

We shall now verify that if $\{\mu_k^n, n \geq 1\}$ converges, then $\{\mu^n, n \geq 1\}$ converges too (the converse is obvious). Consider a subsequence $\{\mu^{n(i)}, i \geq 1\}$ which converges to some probability measure $m_1$ on $\mathscr{M}(d)$. For each integer $p$ let $i_p = \max\{i;\ 3n(i) \leq p\}$. Note that if $p \to \infty$, $i_p \to \infty$, and for $m(p) = n(i_p)$, $l(p) := p - 2m(p) \to \infty$. Writing $\mu^p = \mu^{m(p)} * \mu^{l(p)} * \mu^{m(p)}$ we see that for any limit point $m$ of $\{\mu^p, p \geq 1\}$, for some limit point $m'$ of $\{\mu^{l(p)}, p \geq 1\}$ we have

$$m = m_1 * m' * m_1.$$

As in the proof of Corollary 3.7 we see that $m$ is the distribution of a random matrix of the form (13), with $Q = I$, where $(C_1, K_1)$ and $(K_3, D_3)$ depend only on $m_1$ and $K_2$ on $m'$. If $\{\mu_k^n, n \geq 1\}$ converges to some law, $K_2$ has it for distribution and $m$ does not depend on a particular convergent subsequence. This implies that $\{\mu^n, n \geq 1\}$ converges. □

In order to apply the last part of the theorem under Conditions C1 or C2 we must give criteria ensuring the convergence of $\{\mu_k^n, n \geq 1\}$. If $\{k(M);\ m \in T(\mu)\}$ is in a compact subgroup of $\mathrm{Gl}(d_2)$ this is well known: This sequence converges if and only if the support of $\mu_k$ is not contained in a coset of a proper normal closed subgroup of $K = \{k(M);\ M \in T(\mu)\}$ and the limit is the Haar measure on $K$ (see, e.g., [16]). To study this problem under C2 one may use:

PROPOSITION 4.4. *Let $\mu$ be a probability measure on $\mathscr{M}(d)$ such that $T(\mu)$ is an F-semigroup. Unless $d = 1$ and $\mu = \delta_{-1}$, the sequence $\{\mu^n, n \geq 1\}$ converges.*

PROOF. The statement is obvious when $d = 1$. We thus suppose $d \neq 1$ and set $T = T(\mu)$. Since we shall make use of Theorem 4.13 of Mukherjea and Tserpes [16] we first have to explicate the so-called standard representation of the kernel $K$ of $S$. By assumption there exists a rank-one projection $P$ in $T$. For any $M$ in $T$, $PMP$ is a scalar multiple of $P$. Actually $PMP$ is equal to $P$ or $-P$, since the spectral radius of $PMP$ is equal to one (see Definition 1.2). This implies that the kernel $K$ (i.e., the smallest two-sided ideal) of $T$ is equal to $TPT$. It follows from Theorem 2.14 of [16] that if $G = PKP$ and if $X$ (resp. $Y$) is the set of idempotents of $KP$ (resp. $PK$) then $X \times G \times Y$ is the standard representation of $K$. Notice that $KP = TP$, that $PK = PT$, and that, since $-P$ is in $T$ by Remark 2.5, $G = \{P, -P\}$. By Theorem 4.13 of [16], if $\{\mu^n, n \geq 1\}$ does not converge then there exists a proper subgroup $G'$ of $G$ such that $YX$ is in $G'$. Since the only proper subgroup of $G$ is $\{P\}$, this yields that $YX = \{P\}$.

Let now $\langle \cdot , \cdot \rangle$ be a scalar product on $\mathbb{R}^d$ for which $\operatorname{Im} P$ is orthogonal to $\ker P$ and $v$ be a unit vector in $\operatorname{Im} P$. Note that since $PTP = \{P, -P\}$, for any $M$ in $T$, $MP$, or $-MP$ is in $X$ and $PM$ or $-PM$ is in $Y$. Making use of the irreducibility of $T$ we thus can find $M_i$ in $Y$ and $N_i$ in $X$, $1 \leq i \leq d$, such that $\{{}^tM_i v, 1 \leq i \leq d\}$ and $\{N_i v, 1 \leq i \leq d\}$ are two basis of $\mathbb{R}^d$. Since $YX = \{P\}$,

$$\langle N_i v, M_j v\rangle = \langle {}^tM_j N_i v, v\rangle = 1, \qquad 1 \leq i, j \leq d.$$

This can hold only when $d = 1$. □

**5. Stationary solutions of $x_n = Y_n x_{n-1}$.** Consider the following linear equation on $\mathbb{R}^d$:

$$x_n = Y_n x_{n-1}, \qquad n \geq 1,$$

where $Y_1, Y_2, \dots$ are independent random matrices with a common distribution $\mu$ and $x_0$ is a random vector with law $\nu$ independent of the sequence $\{Y_n, n \geq 1\}$. Then the process $\{x_n, n \geq 0\}$ is stationary if and only if $\nu$ is a $\mu$-invariant probability measure on $\mathbb{R}^d$. We shall determine the set of such invariant distributions when the $Y_i$'s are invertible.

If $\nu$ is a measure on $\mathbb{R}^d$ we write $E(\nu)$ for the linear span of its support. We have already noticed that $E(\nu)$ is $T(\mu)$-invariant when $\nu$ is $\mu$-invariant (see the beginning of the proof of Theorem 2.6). Thus the main step in the description of all the $\mu$-invariant distributions is

THEOREM 5.1. *Let $\mu$ be a probability measure on* $\mathrm{Gl}(d)$ *such that $\int \log^+\|Y\|\, d\mu(Y)$ is finite. A necessary and sufficient condition for the existence of a $\mu$-invariant probability measure $\nu$ on $\mathbb{R}^d$ such that $E(\nu) = \mathbb{R}^d$ is the following*:

(i) *For some $d_1 \geq 0$, $d_2 \geq 0$, $d_1 + d_2 = d$, and some invertible matrix $Q$, $QT(\mu)Q^{-1}$ is contained in $T(d_1; d_2; 0)$. In a convenient basis of $\mathbb{R}^d$ we can write each $M \in T(\mu)$ as*

$$M = \begin{pmatrix} a(M) & c(M) \\ 0 & k(M) \end{pmatrix}$$

*with $a(M) \in \mathcal{M}(d_1)$, $c(M) \in \mathcal{M}(d_1, d_2)$, $k(M) \in \mathcal{M}(d_2)$.*

(ii) *The upper Liapounov exponent of* $\mu_a$, *the image of* $\mu$ *under* $a$, *is strictly negative.*

(iii) *Each* $k(M)$, $M \in T(\mu)$, *is orthogonal.*

(iv) *If* $P$ *is a projection in* $T(\mu)$ *such that* $a(P) = 0$ *and* $P \neq 0$, *the only* $T(\mu)$*-invariant subspace which contains* Im $P$ *is* $\mathbb{R}^d$.

*In this case, if* $m = \lim_{n\to\infty}(1/n)\sum_{i=1}^{n}\mu^i$, *then*

$$\{m \circledast \rho;\ \rho \text{ being a probability measure on } \mathbb{R}^d\}$$

*is the set of* $\mu$*-invariant probability measures on* $\mathbb{R}^d$.

We have used the following notation: If $m$ (resp. $\rho$) is a probability measure on $\mathscr{M}(d)$ (resp. on $\mathbb{R}^d$), $m \circledast \rho$ is the measure on $\mathbb{R}^d$ which satisfies for every bounded Borel function $f$ on $\mathbb{R}^d$,

$$\int f(x)\, d(m \circledast \rho)(x) = \iint f(Mx)\, dm(M)\, d\rho(x).$$

PROOF OF THE NECESSITY. We assume that $\nu$ exists with $E(\nu) = \mathbb{R}^d$. From Proposition 2.1 we know that if $(Y_n)$ are i.i.d. matrices with distribution $\mu$, then the sequence $Y_1Y_2 \cdots Y_n$ is bounded with probability one. Therefore $\{\mu^n, n \geq 1\}$ is tight on $\mathscr{M}(d)$ and we may apply Theorem 3.1. Let $Q$, $d_1$, $d_2$, and $d_3$ be as in that theorem; for convenience we assume that $Q = I$. Consider a random variable $(A, X)$ with values in $\mathscr{M}(d) \times \mathbb{R}^d$ and distribution $m \otimes \nu$. By the $\mu$-invariance of $\nu$ the distribution of $AX$ is $\nu$, but $b(A) = 0$ a.s. and $E(\nu) = \mathbb{R}^d$, thus $d_3 = 0$. We now prove that $a(Y_1 \cdots Y_n)$ converges to 0 almost surely. Since the support of $m$ is in $T(\mu)$ there exists a matrix $P$ in $T(\mu)$ such that $a(P) = 0$ and $k(P) = I$. Considering the image of $\mu$ under the conjugation by

$$\begin{pmatrix} I & c(P) \\ 0 & I \end{pmatrix}$$

we can moreover suppose $c(P) = 0$. As in the proof of Proposition 2.1 for almost all $\omega$,

$$\lim_{n\to\infty} Y_1(\omega) \cdots Y_n(\omega)Pm = \lim_{n\to\infty} Y_1(\omega) \cdots Y_n(\omega)m.$$

We shall prove that for each $\omega$ such that this equality holds $a(Y_1(\omega) \cdots Y_n(\omega))$ converges to 0. Any limit point $M$ of the sequence $Y_1(\omega) \cdots Y_n(\omega)$ satisfies $MPm = Mm$. Let us see that this implies $a(M) = 0$. If $H$ is a random matrix with distribution $m$, we deduce from $MPm = Mm$ carrying out the matrix multiplication that the random variables $(X, Y) = (k(M)k(H), a(M)c(H) + c(M)k(H))$ and $(X', Y') = (k(M)k(H), c(M)k(H))$ have the same distribution. But $Y' = c(M)k(M)^{-1}X'$ so $Y = c(M)k(M)^{-1}X$ and $a(M)c(H) = 0$ a.s. This implies that for $V = \{(x, y) \in \mathbb{R}^d;\ x \in \mathbb{R}^{d_1},\ y \in \mathbb{R}^{d_2},\ a(M)x = 0\}$,

$$m\{H \in \mathscr{M}(d);\ \operatorname{Im} H \subset V\} = 1.$$

By invariance of $\nu$, $m \otimes \nu\{(H, x);\ Hx \in V\} = \nu(V)$ so $\nu(V) = 1$. This holds only if $V = \mathbb{R}^d$, i.e., $a(M) = 0$.

We have thus proved that $\|a(Y_1 \cdots Y_n)\|$ converges to 0 with probability one. If $E(\log^+\|a(Y_1)\|)$ is finite this implies that the Liapounov exponent $\gamma(\mu_a)$ is strictly negative (see Lemma 5.2).

To verify (iv) we use the fact that for any $P$ in $T(\mu)$, $Y_1 \cdots Y_n P\nu$ converges to some probability measure $\nu_\omega$ such that $\nu = \int \nu_\omega \, dP(\omega)$ (see the proof of Proposition 2.1). Since each $\nu_\omega$ is carried by the smallest $T(\mu)$-invariant subspace which contains Im $P$, this subspace carries $\nu$ and is thus equal to $\mathbb{R}^d$. □

PROOF OF THE SUFFICIENCY. If the conditions hold $\{\mu^n, n \geq 1\}$ is tight and $(1/n)\sum_{i=1}^n \mu^i$ converges to some $m$ (see Theorem 4.1). For a probability measure $\rho$ on $\mathbb{R}^d$ it is clear that $m \circledast \rho$ is $\mu$-invariant. Consider a projection $P$ in $T(\mu)$ such that $a(P) = 0$. As above we may suppose if $P \neq 0$ that $k(P) = I$ and $c(P) = 0$. We choose a probability measure $\rho$ on Im $P$ whose support spans Im $P$ and set $\nu = m \circledast \rho$. Since $E(\nu)$ is $T(\mu)$-invariant $PE(\nu) \subset E(\nu)$ and since the projection of $E(\nu)$ into Im $P$ is all Im $P$, Im $P \subset E(\nu)$. By (iv), $E(\nu) = \mathbb{R}^d$. This proves the sufficiency.

The last claim of the theorem is clear: If $\lambda$ is an $\mu$-invariant probability measure on $\mathbb{R}^d$, $\mu \circledast \lambda = \lambda$ and $\lambda = m \circledast \lambda$. Conversely each $m \circledast \rho$ is $\mu$-invariant. □

In the course of the preceding proof we have used:

LEMMA 5.2. *Let $X_1, X_2, \ldots$ be independent random elements of* Gl$(d)$ *with a common distribution $\mu$. Assume that $\int \log^+\|Y\| \, d\mu(Y)$ is finite and that a.s.*

$$\lim_{n \to \infty} \|X_n X_{n-1} \cdots X_1\| = 0.$$

*Then the upper Liapounov exponent associated with $(X_n)$ is strictly negative.*

PROOF. Let $\lambda$ be a probability measure on the unit sphere $S$ of $\mathbb{R}^d$ such that

$$\iint f\left(\frac{Yu}{\|Yu\|}\right) d\mu(Y) \, d\lambda(u) = \int f(u) \, d\lambda(u)$$

for any bounded Borel function $f$ on $S$. If $U_0$ is a random variable with law $\lambda$, independent of the sequence $(X_n)$, then

$$Z_n = \left(X_n, \frac{X_{n-1} \cdots X_1 U_0}{\|X_{n-1} \cdots X_1 U_0\|}\right)$$

is a stationary Markov chain. If we set $F(Y, u) = \log\|Yu\|$, then

$$\frac{1}{n}\log\|X_n \cdots X_1 U_0\| = \frac{1}{n} \sum_{i=1}^n F(Z_i).$$

By Lemma 3.3 of Furstenberg and Kifer [6] we can choose $\lambda$ in such a way that $Z_n$ is ergodic and $(1/n)\sum_{i=1}^n F(Z_i)$ converges to the exponent $\gamma(\mu)$. Since

$$\frac{1}{n} \sum_{i=1}^n F(Z_i) \leq \log\|X_{n-1} \cdots X_1\| \to -\infty \quad \text{as } n \to \infty$$

it thus follows from Lemma 3.6 of Guivarc'h and Raugi [7] that $\gamma(\mu) < 0$. □

To describe all the invariant measures on $\mathbb{R}^d$ for any $\mu$ on $\mathrm{Gl}(d)$ such that $\int \log^+\|Y\|\, d\mu(Y)$ is finite we choose a maximal $T(\mu)$ invariant subspace $F$ with the following property:

In a suitable basis of $F$, for each $M$ in $T(\mu)$ the restriction $M_F$ of $M$ to $F$ can be written

$$M_F = \begin{pmatrix} a(M) & c(M) \\ 0 & k(M) \end{pmatrix},$$

where the exponent $\gamma(\mu_a)$ is strictly negative and each $k(M)$ is orthogonal.

Such a subspace is in fact unique. Each $\mu$-invariant probability measure on $\mathbb{R}^d$ is carried by $F$ and is invariant under the image $\mu_F$ of $\mu$ under the restriction to $F$. Therefore, for $m = \lim (1/n)\Sigma_{i=1}^n \mu_F^i$, each $\mu$-invariant probability measure $\nu$ can be written $\nu = m \circledast \rho$ for some measure $\rho$ on $F$.

**6. Ergodic properties of stable linear stochastic equations.** In this part we consider a discrete time linear stochastic equation on $\mathbb{R}^d$:

$$x_n = Y_n x_{n-1}, \qquad n \geq 1.$$

We want to describe the ergodic properties of the Markov chain $(x_n)$ under the following assumption (A), related to the stability in probability of this process:

CONDITION (A). We say that a probability measure $\mu$ on $\mathscr{M}(d)$ satisfies (A) if the following holds:

(i) $\int \log^+\|Y\|\, d\mu(Y) < +\infty$.

(ii) *For some* $d_1 \geq 0$, $d_2 \geq 0$, $d_3 \geq 0$, *and* $d_1 + d_2 + d_3 = d$, $T(\mu)$ *is contained in* $T(d_1; d_2; d_3)$. *As in Definition* 1.3 *we write for $M$ in* $T(\mu)$,

$$M = \begin{pmatrix} a(M) & c(M) & e(M) \\ 0 & k(M) & d(M) \\ 0 & 0 & b(M) \end{pmatrix}.$$

(iii) *The upper Liapounov exponents of the image of* $\mu$ *under* $a(\cdot)$ *and* $b(\cdot)$ *are strictly negative.*

(iv) *Each* $k(M)$, $M \in T(\mu)$, *is orthogonal.*

We shall write $K = \{k(M);\ M \in T(\mu)\}$; it is a compact group. We have met this condition (A) either under C1 or under C2 (see Remark 3.6 for instance).

As before we set $M_n = Y_n \cdots Y_1$, where $Y_1, Y_2, \ldots$ are i.i.d. matrices with distribution $\mu$. The following decomposition shows that asymptotically $Y_n$ can be written as $L_n K_n R_n$ where $L_n \in \mathscr{M}(d, d_2)$, $K_n \in O(d_2)$, and $R_n \in \mathscr{M}(d_2, d)$ are independent matrices such that, for all $m$ and $n$, $R_n L_m$ is the identity. If this were exactly true we would have

$$M_n = L_n K_n \cdots K_1 R_1$$

and the proposition would be obvious. This kind of decomposition already appears in Kesten and Spitzer [13].

PROPOSITION 6.1. *Let $M_n := Y_n \cdots Y_1$, where the $Y_i$'s are i.i.d. matrices in $\mathcal{M}(d)$ with distribution $\mu$. We suppose that $\mu$ satisfies* (A).

*For any $M$ in $T(\mu)$ let*

$$l(M) = \begin{pmatrix} c(M)k(M)^{-1} \\ I \\ 0 \end{pmatrix} \quad \text{and} \quad r(M) = \begin{pmatrix} 0 & I & k(M)^{-1}d(M) \end{pmatrix}.$$

*Then, for each $n$,*

$$M_n = l(M_n)k(M_n)r(M_n) + P_n,$$

*where*

(i) *$r(M_n)$ converges almost surely to a random matrix $R$.*
(ii) *$l(Y_1 \cdots Y_n)$ converges almost surely.*
(iii) *$P_n$ converges to 0 a.s.*
(iv) *$l(M_n)$, $k(M_n)$, and $r(M_n)$ are asymptotically independent.*

In this proposition $l(M)$ is a $d \times d_2$ matrix, $r(M)$ a $d_2 \times d$ matrix, $I$ is the identity matrix of order $d_2$, and 0 in $l(M)$ (resp. $r(M)$) is the zero matrix of order $d_3$ (resp. $d_1$).

PROOF. We first verify (i). For $M$ in $T(\mu)$ we set

$$r'(M) = k(M)^{-1}d(M).$$

By Lemma 4.2 for $n > p$,

$$r'(M_n) - r'(M_p) = \sum_{j=p+1}^{n} k(M_j)^{-1}d(Y_j)b(M_{j-1});$$

hence,

$$\|r'(M_n) - r'(M_p)\| \le \sum_{j=p+1}^{n} \|d(Y_j)\|\,\|b(M_{j-1})\|.$$

Let $\gamma$ denote the upper Liapounov exponent associated with the distribution $\mu_s$ which is defined in Lemma 4.3. It follows from this lemma that $\gamma < 0$.

Now $\|b(M_n)\| \le \exp(\frac{3}{4}\gamma_n)$ for $n$ large enough, a.s. Since $E(\log^+\|d(Y_1)\|) < \infty$ we obtain with the Borel–Cantelli lemma $\|d(Y_n)\| \le \exp(-\frac{1}{4}\gamma_n)$ for $n$ large enough. This yields that

$$\|r'(M_p) - r'(M_n)\| \le \sum_{j=p+1}^{n} \exp\left(-\tfrac{1}{4}\gamma_j + \tfrac{3}{4}\gamma(j-1)\right) \le C\exp\tfrac{1}{2}\gamma p \tag{14}$$

for a suitable $C$ and $p, n$ large enough, entailing that $r'(M_n)$ is almost surely a Cauchy sequence and that $r(M_n)$ converges a.s. The proof of (ii) follows the same line.

In order to prove (iii) it suffices to check that a.s.

$$\lim_{n\to\infty} e(M_n) - c(M_n)k(M_n)^{-1}d(M_n) = 0.$$

The proof we give is well suited for generalization to the continuous time model.

Fix an $\varepsilon > 0$ sufficiently small. Since $(1/n)E(\log\|s(M_n)\|)$ converges to $\gamma$ we can find an integer $k$ such that

$$E(\log\|s(M_k)\|) \le k(\gamma - \varepsilon). \tag{15}$$

For each integer $m$ we set

$$Z_m = Y_{mk} \cdots Y_{(m-1)k+1}.$$

For convenience we will write, for $n > p$, $M_n M_p^{-1}$ instead of $Y_n \cdots Y_{p+1}$ even if the matrices are not invertible.

If $mk < n \le (m+1)k$ we have, by Lemma 4.2,

$$\begin{aligned} e(M_n) &- c(M_n)k(M_n)^{-1}d(M_n) \\ &= e(s(Y_n)\cdots s(Y_1)) - a(M_n M_{mk}^{-1})a(Z_m)\cdots a(Z_2)c(Z_1)r'(M_n) \\ &\quad + \sum_{j=1}^{m-1} a(M_n M_{mk}^{-1})a(Z_m)\cdots a(Z_{j+2})c(Z_{j+1})k(M_{kj})\{r'(M_{kj}) - r'(M_n)\} \\ &\quad + c(M_n M_{mk}^{-1})k(M_{km})\{r'(M_{km}) - r'(M_n)\}. \end{aligned}$$

In order to show that this quantity converges to zero it is enough to prove, since $\gamma < 0$, that for each $p$ fixed

$$\sum_{j=p}^{m-1} \|a(M_n M_{mk}^{-1})\|\,\|a(Z_m)\cdots a(Z_{j+2})\|\,\|c(Z_{j+1})\|\,\|r'(M_{kj}) - r'(M_n)\| \tag{16}$$

and

$$\|c(M_n M_{mk}^{-1})\|\,\|r'(M_{mk}) - r'(M_n)\| \tag{17}$$

converge to 0 as $n \to \infty$.

By the Borel–Cantelli lemma, for almost all $\omega$ and $m$, $j$ sufficiently large,

$$\sup\{\|a(M_n M_{mk}^{-1})\|;\ mk < n \le (m+1)k\} \le e^{m\varepsilon} \tag{18}$$

and

$$\sup\{\|c(M_n M_{mk}^{-1})\|;\ mk < n \le (m+1)k\} \le e^{m\varepsilon}. \tag{19}$$

By (14) and (19), (17) converges to zero. Now by the law of large numbers, (15) yields that

$$\limsup_{n\to\infty}\left(\sup_{j\le n}\frac{1}{n}\sum_{i=j}^{n}\{\log\|a(Z_i)\| - k(\gamma-\varepsilon)\}\right) \le 0 \quad \text{a.s.}$$

Hence for $m$ large enough and every $j \le m$

$$\|a(Z_m)\cdots a(Z_j)\| \le \|a(Z_m)\|\cdots\|a(Z_j)\| \le \exp k(m-j)(\gamma-\varepsilon). \tag{20}$$

Using (14), (18), (19), and (20) we obtain that a.s. for $m$ and $r$ large enough, (16) is smaller than

$$Ce^{2m\varepsilon}\sum_{j=r}^{m-1} e^{k(m-2-j)(\gamma-\varepsilon)}e^{(j+1)\varepsilon}e^{\gamma kj/2},$$

which is easily seen to converge to 0 when $m \to \infty$, for $\varepsilon$ small enough.

We now prove (iv). Recall that by definition, $l(M_n)$, $k(M_n)$, and $r(M_n)$ are *asymptotically independent* if for every bounded continuous functions $f$: $\mathscr{M}(d, d_2) \to \mathbb{R}$, $g$: $\mathscr{M}(d_2) \to \mathbb{R}$, $h$: $\mathscr{M}(d_2, d) \to \mathbb{R}$,

$$E\{f(l(M_n))g(k(M_n))h(r(M_n))\} - E\{f(l(M_n))\}E\{g(k(M_n))\}E\{h(r(M_n))\}$$

converges to 0 as $n \to \infty$.

We know that $\{\mu^n, n \geq 1\}$ is tight (see Theorem 4.1). As in the proof of this theorem if $m$ is the limit of a convergent subsequence $(\mu^{n(i)})$, $m$ is the distribution of a random matrix which can be written

$$\begin{pmatrix} 0 & C_1K_2K_3 & C_1K_2D_3 \\ 0 & K_1K_2K_3 & K_1K_2D_3 \\ 0 & 0 & 0 \end{pmatrix},$$

where $(C_1, K_1)$, $K_2$, and $(K_3, D_3)$ are independent and the distributions of $K_1, K_2, K_3$ are limit points of $\{\mu_k^n, n \geq 1\}$, where $\mu_k$ is the image of $\mu$ under $k(\cdot)$. Since $\mu_k$ is a probability measure on the compact group $K$, we know that there is some $b$ in $K$ such that $\delta_{b^n} * \mu_k^n$ converges to the Haar measure on a compact normal subgroup $K'$ of $K$ (see Mukherjea and Tserpes [16], Theorem 4.15). We thus may write $K_1 = k_1\tilde{K}_1$, $K_2 = k_2\tilde{K}_2$, and $K_3 = k_3\tilde{K}_3$, where $k_1, k_2, k_3$ are in $K$ and $\tilde{K}_1$, $\tilde{K}_2$, and $\tilde{K}_3$ are three independent random matrices whose distribution is the Haar measure on $K'$. Since $K'$ is normal $k\tilde{K}_i$ has the same law as $\tilde{K}_ik$ for each $k$ in $K$ and $i = 1, 2, 3$. Using the invariance of the Haar measure and independence we get that the following random variables have the same distribution:

(a) $$\big(C_1K_1^{-1}, K_1K_2K_3, K_3^{-1}D_3\big) = \big(C_1\tilde{K}_1^{-1}k_1^{-1}, k_1\tilde{K}_1k_2\tilde{K}_2k_3\tilde{K}_3, \tilde{K}_3^{-1}k_3^{-1}D_3\big),$$

(b) $$\big(C_1k_2\tilde{K}_1^{-1}k_2^{-1}k_1^{-1}, k_1k_2\tilde{K}_1\tilde{K}_2\tilde{K}_3k_3, k_3^{-1}\tilde{K}_3^{-1}D_3\big),$$

(c) $$\big(C_1k_2\tilde{K}_1^{-1}k_2^{-1}k_1^{-1}, k_1k_2\tilde{K}_2k_3, k_3^{-1}\tilde{K}_3^{-1}D_3\big),$$

(d) $$\big(C_1K_1^{-1}, k_1K_2k_3, K_3^{-1}D_3\big).$$

We obtain, if $f(l(M)) = f'(c(M)k(M)^{-1})$ and $h(r(M)) = h'(k(M)^{-1}d(M))$,

$$\begin{aligned}
&\lim_{i\to\infty} E\big\{f\big(l(M_{n(i)})\big)g\big(k(M_{n(i)})\big)h\big(r(M_{n(i)})\big)\big\} \\
&= E\big\{f'\big(C_1K_1^{-1}\big)g(K_1K_2K_3)h'\big(K_3^{-1}D_3\big)\big\} \\
&= E\big\{f'\big(C_1K_1^{-1}\big)g(k_1K_2k_3)h'\big(K_3^{-1}D_3\big)\big\} \\
&= E\big\{f'\big(C_1K_1^{-1}\big)\big\}E\{g(k_1K_2k_3)\}E\big\{h'\big(K_3^{-1}D_3\big)\big\} \\
&= \lim_{i\to\infty} E\big\{f\big(l(M_{n(i)})\big)\big\}E\big\{g\big(k(M_{n(i)})\big)\big\}E\big\{h\big(r(M_{n(i)})\big)\big\}.
\end{aligned}$$

This being true for every convergent subsequence, the result is proved. □

To study the pathwise behaviour of the solutions of $x_n = Y_n x_{n-1}$, i.e., $x_n = M_n x_0$ for $M_n = Y_n \cdots Y_1$, we shall prove an ergodic result. We first state it in a particular case as a lemma.

LEMMA 6.2. *Suppose that $\mu$ is a distribution on $\mathscr{M}(d)$ satisfying* (A) *with $d_3 = 0$. If $m = \lim(1/n)\Sigma_{i=1}^n \mu^i$, for each bounded continuous function $f$ on $\mathscr{M}(d)$*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} f(M_i(\omega)) = \int f\,dm \quad a.s.$$

PROOF. Let $X = \{M \in T(d_1; d_2; 0);\ k(M) \in K\}$. For any $M_0$ in $X$ $\{M_n(\omega)M_0, n \geq 1\}$ is a Markov chain on $X$ starting from $M_0$ with transition probability $P$ defined, for any bounded Borel function $f$ on $X$ and $A$ in $X$, by

$$Pf(A) = \int f(MA)\,d\mu(M).$$

If $\lambda$ is an invariant distribution for this Markov chain, $\lambda P = \lambda$ entailing $\mu * \lambda = \lambda$ and $m * \lambda = \lambda$. Consider two independent random matrices, $Z_1$ with distribution $m$ and $Z_2$ with distribution $\lambda$. Carrying out the matrix multiplication we obtain

$$a(Z_1 Z_2) = 0, \qquad c(Z_1 Z_2) = c(Z_1)k(Z_2), \qquad k(Z_1 Z_2) = k(Z_1)k(Z_2).$$

From the description of $m$ given in Corollary 3.7 it is clear that $\lambda$, which is the distribution of $Z_1 Z_2$, must be equal to $m$. So $m$ is the unique invariant distribution of this Markov chain. By the ergodic theorem applied to a countable dense set of continuous functions on $X$ with compact support we know that

For almost all $\omega$ and $m$-almost all $M_0$,

$$\frac{1}{n} \sum_{i=1}^{n} f(M_i(\omega)M_0) \to \int f\,dm \tag{21}$$

for all $f$ in this dense set. But for $\omega$ and $M_0$ fixed this is in fact a "convergence in law" statement; therefore, it holds for each bounded continuous function $f$ on $\mathscr{M}(d)$. We want to prove that (21) is true when $M_0$ is the identity matrix. Fix an $M_0$ for which (21) holds for almost all $\omega$. If $A$ is the matrix in $X$ such that $a(A) = I$, $c(A) = 0$, $k(A) = k(M_0)^{-1}$ we define $f'\colon X \to \mathbb{R}$ by $f'(M) = f(MA)$ for $M$ in $X$. Since $a(M_n) \to 0$ a.s.

$$f(M_n(\omega)) - f'(M_n(\omega)M_0) \to 0 \quad \text{a.s. when } n \to \infty.$$

Applying (21) to $f'$ we find that a.s.

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} f(M_i(\omega)) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} f'(M_i(\omega)M_0) = \int f'\,dm.$$

As the mean of the left side is $\int f\,dm$, we have $\int f'\,dm = \int f\,dm$ and the lemma is proved. □

In the next theorem we use the notation and the results of Proposition 6.1.

THEOREM 6.3. *Let $M_n = Y_n \cdots Y_1$ where $Y_1, Y_2, \ldots$ are i.i.d. matrices on $\mathcal{M}(d)$ whose distribution satisfies Condition* (A). *Let $\lambda$ be the distribution of $\lim l(Y_1 \cdots Y_n)$, $\rho$ be the Haar measure on $K = \{k(M);\ M \in T(\mu)\}$, and $R = \lim r(M_n)$. For almost all $\omega$ and any bounded continuous function $f$ on $\mathcal{M}(d)$,*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} f(M_i(\omega)) = \iint f(LUR(\omega))\, d\lambda(L)\, d\rho(U).$$

PROOF. Since this is for each $\omega$ a "convergence in law" statement it suffices to prove the theorem for functions $f$ on $T(d_1; d_2; d_3)$ which can be written as

$$f(M) = g(l(M)k(M))h(r(M)),$$

where $g$ and $h$ are bounded continuous functions on $T(d_1; d_2; d_3)$.

If we apply Lemma 6.2 to the image of $\mu$ under the map $\tau$: $T(d_1; d_2; d_3) \to T(d_1; d_2; 0)$ given by, for $M$ in $T(d_1; d_2; d_3)$,

$$\tau(M) = \begin{pmatrix} a(M) & c(M) \\ 0 & k(M) \end{pmatrix},$$

it is easily seen that, a.s.,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} g(l(M_i)k(M_i)) = \iint g(LU)\, d\lambda(L)\, d\rho(U).$$

Since $\lim h(r(M_n)) = h(R)$ a.s. the theorem follows. □

COROLLARY 6.4. *Under the above hypotheses, a.s. for each bounded continuous function $\varphi$ on $\mathbb{R}^d$ and each $x$ in $\mathbb{R}^d$,*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \varphi(M_i(\omega)x) = \iint \varphi(LUR(\omega)x)\, d\lambda(L)\, d\rho(U).$$

From this we deduce the asymptotic behaviour of the Markov chain $x_n = M_n x$ on $\mathbb{R}^d$:

Let $\{e_i, 1 \le i \le d\}$ be the canonical basis of $\mathbb{R}^d$ and let $E_1$, $E_2$, and $E_3$ be the linear span of $\{e_1, \ldots, e_{d_1}\}$, $\{e_{d_1+1}, \ldots, e_{d_1+d_2}\}$, and $\{e_{d_1+d_2+1}, \ldots, e_d\}$. Write each $x$ in $\mathbb{R}^d$ as $x = u + v + w$ where $u \in E_1$, $v \in E_2$, and $w \in E_3$. For any $v$ in $E_2$ consider the cylinder $A(v) = E_1 \times Kv \times \{0\}$.

(i) If $v = w = 0$, $M_n x$ converges to 0 exponentially fast.

(ii) If $w = 0$, $v \neq 0$, $A(v)$ is an invariant set for the Markov chain: starting at time 0 in $A(v)$ the process remains in this set. Moreover, on $A(v)$ this Markov chain has a unique invariant distribution $m_v$ defined by, for each Borel set $B$ in $A(v)$

$$m_v(B) = \iint 1_B(LUv)\, d\lambda(L)\, d\rho(U) = \int 1_B(Mx)\, dm(M),$$

where $m = \lim(1/n)\Sigma_{i=1}^n \mu^i$, and is recurrent in each open set of nonzero $m_v$ measure.

(iii) If $w \neq 0$, the Markov chain chooses a random cylinder, $A(Rx)$, and goes to it exponentially fast.

REMARK 6.5. Suppose that Condition (A) holds and that there exists a fixed $x$ in $\mathbb{R}^d$ so that $Mx = x$ for each $M$ in $T(\mu)$ (consider for instance stochastic matrices). It is easy to see that in this case one can suppose $d_1 = 0$, $d_2 = 1$, and $k(M) = 1$. Therefore, by Proposition 6.1, $M_n$ converges almost surely.

**7. Linear stochastic differential equations.** Consider as in (1) the linear stochastic differential equation (linear SDE) on $\mathbb{R}^d$

$$dx_t = S_0 x_t\, dt + \sum_{i=1}^{r} S_i x_t \circ db_t^i, \tag{22}$$

where $S_0, S_1, \dots, S_r$ are $d \times d$ matrices and $(b^1, \dots, b^r)$ the usual $\mathbb{R}^r$-valued Brownian motion. Let $\{e_i, 1 \le i \le d\}$ be the canonical basis of $\mathbb{R}^d$ and let $M_t$ be the matrix whose $i$th column is the solution of (22) starting from $e_i$ at time 0. We have

$$dM_t = S_0 M_t\, dt + \sum_{i=1}^{r} S_i M_t \circ db_t^i, \qquad M_0 = I. \tag{23}$$

Since if $x_0 = x$, $x_t = M_t x$ is a solution of (22), $(M_t,\ t \ge 0)$ is the flow associated with (22) (see Ikeda and Watanabe [12]). As such, if $\mu_t$ is the distribution of $M_t$, $\mu_t * \mu_s = \mu_{t+s}$ and $M_t$ is in $\mathrm{Gl}(d)$.

We define for $i = 0, 1, \dots, r$ the right invariant vector field $\tilde{S}_i$ on the manifold $\mathrm{Gl}(d)$ by: if $f$ is a smooth real function on $\mathrm{Gl}(d)$

$$[\tilde{S}_i(M)](f) = \frac{d}{dt} f((\exp tS_i)M)\Big|_{t=0}, \qquad M \in \mathrm{Gl}(d).$$

With the usual notation for SDE on manifolds (see Ikeda and Watanabe [12], Chapter V.1), (23) can be written as

$$dM_t = \tilde{S}_0(M_t)\, dt + \sum_{i=1}^{r} \tilde{S}_i(M_t) \circ db_t^i, \qquad M_0 = I. \tag{24}$$

By Theorem 1.2, Chapter V of [12], the infinitesimal generator associated with $M_t$ is

$$A = \frac{1}{2}\left( \sum_{i=1}^{r} \tilde{S}_i^2 + \tilde{S}_0 \right).$$

LEMMA 7.1. *Let $G$ be the connected Lie subgroup of* $\mathrm{Gl}(d)$ *whose Lie algebra is generated by $S_0, S_1, \dots, S_r$. For any $t > 0$, $\mu_t(G) = 1$. For $\lambda = \int_0^\infty e^{-t} \mu_t\, dt$, if $H$ is a closed subgroup of* $\mathrm{Gl}(d)$ *such that $\lambda(H) = 1$, $G$ is included in $H$.*

PROOF. Since each $\tilde{S}_i$ is a vector field on the manifold $G$ the equation (24) can be considered as a SDE on $G$. By unicity $M_t$ is in $G$ a.s. and $\mu_t(G) = 1$. We consider also $A$ as a differential operator on $G$. Since $S_0, \dots, S_r$ generate the Lie algebra of $G$, the Hörmander's theorem (see Hörmander [11]) implies that $A$ is hypoelliptic. Therefore, $\lambda$ being the solution of $(A^* - 1)\lambda = -\delta_0$, the measure $\lambda$ has a density on $G$. If $H$ is a closed subgroup of $\mathrm{Gl}(d)$ such that $\lambda(H) = 1$, $H \cap G$ must contain an open set. The connectedness of $G$ implies that $H \cap G = G$ (see Hewitt and Ross [10], 7.9). □

LEMMA 7.2. *Suppose that*:

(i) *There is no proper subspace $V$ of $\mathbb{R}^d$ such that $S_i(V) \subset V$ for $i = 0, 1, \dots, r$.*

(ii) *There is no invertible matrix $Q$ such that $QS_iQ^{-1}$ is skew-symmetric for $i = 0, 1, \dots, r$.*

*Then the family $\{\mu_t, t > 0\}$ is tight on $\mathscr{M}(d)$ if and only if $\gamma(\mu_t) < 0$.*

This result is a generalization of Theorem 7.2 of Hashminski [8]. Recall that the set of all skew-symmetric matrices (i.e., matrices $M$ with $M + {}^tM = 0$) is the Lie algebra of the orthogonal group.

PROOF. We first prove that if $\{\mu_t, t > 0\}$ is tight $\gamma(\lambda) < 0$, where $\lambda = \int_0^\infty \mu_t \, dt$. We shall apply Proposition 2.8 to $\lambda$. We first verify that there does not exist a finite set $V_1, \dots, V_p$, where each $V_i$ is a proper subspace of $\mathbb{R}^d$ such that

$$M(V_1 \cup \cdots \cup V_p) \subset V_1 \cup \cdots \cup V_p \quad \text{for any } M \text{ in } T(\lambda).$$

We can suppose that $\dim V_1 \leq \dim V_i$, $1 \leq i \leq p$. For each $M$ of $\mathrm{Gl}(d) \cap T(\lambda)$, $MV_1$ must be one of the $V_i$. Therefore, after reordering terms, we may suppose that for some $q \leq p$, $\dim V_1 = \dim V_i$, $1 \leq i \leq q$ and that

$$MV_1 \in \{V_1, V_2, \dots, V_q\} \quad \text{for any } M \text{ in } \mathrm{Gl}(d) \cap T(\lambda).$$

Let $H = \{M \in \mathrm{Gl}(d);\ MV_1 \in \{V_1, \dots, V_q\}\}$. $H$ is a closed subgroup of $\mathrm{Gl}(d)$ and $\lambda(H) = 1$. The group $G$ defined in Lemma 7.1 is contained in $H$ and its action on the space of $\dim V_1$-dimensional subspaces of $\mathbb{R}^d$ is continuous. $G$ being connected, the orbit of $V_1$ under $G$ is finite only if it is $V_1$ alone. So $MV_1 = V_1$ for each $M$ in $G$. This in turn implies $S_i(V_1) \subset V_1$ for $i = 0, 1, \dots, r$ in contradiction with (i).

For some $c > 0$, $E(\|M_t\|^2) \leq e^{ct}$ (see [12], page 164). Thus for $a = 1/c$, $E(\|M_t\|^a) \leq e^{t/2}$ and $\int \|Y\|^a \, d\lambda(Y) = \int_0^\infty E(\|M_t\|^a) e^{-t} \, dt$ is finite. Since $M_t^{-1}$ also satisfies a linear SDE, namely

$$dM_t^{-1} = -M_t^{-1} S_0 \, dt - \sum_{i=1}^{r} M_t^{-1} S_i \circ db_t^i,$$

we may suppose that $\int \|Y^{-1}\|^a \, d\lambda(Y)$ is also finite. Therefore the hypotheses of Proposition 2.8 are verified.

If $\{\mu_t,\ t \geq 0\}$ is tight, $\{\lambda^n,\ n \geq 0\}$ is also tight and by Proposition 2.8 and Theorem 2.6 either $\gamma(\lambda) < 0$ or $T(\lambda)$ is in a compact subgroup of $\mathrm{Gl}(d)$. In the second case $G$ would be, by Lemma 7.1, contained in a conjugate of $O(d)$ in contradiction with (ii).

Thus $\gamma(\lambda) < 0$ but

$$\gamma(\lambda) = \lim_{n\to\infty} \frac{1}{n} \int \log \|M\|\, d\lambda^n(M) = \lim \frac{1}{n} \iint_0^\infty \log\|M\| e^{-s} \frac{s^{n-1}}{(n-1)!}\, d\mu_s(M)\, ds$$

$$= \lim_{s\to\infty} \frac{1}{s} \int \log\|M\|\, d\mu_s(M) = \gamma(\mu_1).$$

The theorem follows immediately. □

We can now characterize the stability in probability of the zero solution of (1). By definition it is equivalent to the tightness of $\{\mu_t,\ t > 0\}$ on $\mathscr{M}(d)$.

Given $T_0, T_1, \dots, T_r$ being $p \times p$ matrices we denote by $\gamma(T_0, T_1, \dots, T_r)$ the upper Liapounov exponent associated with the linear equation on $\mathbb{R}^p$

$$dy_t = T_0 y_t\, dt + \sum_{i=1}^r T_i y_t \circ db_t^i$$

defined in the introduction. For instance, with our notation

$$\gamma(S_0, S_1, \dots, S_r) = \gamma(\mu_1) = \frac{1}{t}\gamma(\mu_t) \quad \text{if } t > 0.$$

**Theorem 7.3.** *Consider on $\mathbb{R}^d$ the linear SDE*

$$dx_t = S_0 x_t\, dt + \sum_{i=1}^r S_i x_t \circ db_t^i.$$

*The solution $x_t \equiv 0$ is stable in probability if and only if there exists some invertible matrix $Q$ and integers $d_1 \geq 0$, $d_2 \geq 0$, $d_3 \geq 0$, and $d = d_1 + d_2 + d_3$, such that each $S_i' = QS_iQ^{-1}$, $0 \leq i \leq r$, is in $T(d_1; d_2; d_3)$ and*

(i) *$k(S_i')$ is skew-symmetric.*

(ii) *The Liapounov exponents $\gamma(a(S_0'), \dots, a(S_r'))$ and $\gamma(b(S_0'), \dots, b(S_r'))$ are strictly negative.*

We have used Definition 1.3 and written

$$S_i' = \begin{pmatrix} a(S_i') & c(S_i') & e(S_i') \\ 0 & k(S_i') & d(S_i') \\ 0 & 0 & b(S_i') \end{pmatrix}.$$

**Proof.** We first suppose that $x_t \equiv 0$ is stable in probability. In this case $\{\mu_t,\ t > 0\}$ is tight and for $\lambda = \int_0^\infty e^{-s}\mu_s\, ds$, $\{\lambda^n,\ n \geq 1\}$ is tight. We may thus apply Theorem 3.1 to $\lambda$. Let $Q$ and $d_1, d_2, d_3$ be the quantities introduced in this theorem, and $G$ the group defined in Lemma 7.1. We can suppose that $Q = I$. The verification of (i) is easy: if $H$ is the smallest closed subgroup of $\mathrm{Gl}(d)$ which contains $\{k(M);\ M \in T(\lambda)\}$, $H$ is compact by (iii) of Theorem 3.1 and $\{k(M);\, M \in G\}$ is contained in $H$ by Lemma 7.1. So $\{k(M);\ M \in G\}$ is contained in a conjugate of $O(d_2)$ and we may choose $Q$ in such a way that it is in fact in $O(d_2)$. In this case each $k(S_i') = k(S_i)$ is in the Lie algebra of $O(d_2)$ hence is skew-symmetric (recall that we have supposed that $Q = I$).

We verify (ii) using Lemma 7.2. Since the proof is the same for both exponents we only consider $\gamma(a(S_0'), \dots, a(S_r'))$. If we set $a(S_i) = T_i$, $a(\mu_t) = \rho_t$, and $a(\lambda) = \tau$ we have to show that if $\{\rho_t, t > 0\}$ is tight and $(1/n)\Sigma_{i=1}^n \tau^i$ converges to the Dirac mass at the zero matrix then the upper Liapounov exponent associated with

$$dy_t = T_0 y_t\, dt + \sum_{i=1}^{r} T_i y_t \circ db_t^i, \qquad y_t \in \mathbb{R}^{d_1}, \tag{25}$$

is strictly negative. Consider a strictly increasing sequence $\{0\} = E_0 \subset E_1 \subset \cdots \subset E_p = \mathbb{R}^{d_1}$ such that for $j = 1, \dots, p$ $E_j$ is a minimal subspace which contains $E_{j-1}$ and such that $T_i(E_j) \subset E_j$ for $i = 0, \dots, r$. Let $\alpha_j(T_i)$ denote the endomorphism of $F_j := E_j/E_{j-1}$ defined by

$$\alpha_j(T_i)\bar{x} = \overline{T_i x}, \qquad x \in E_j, \tag{26}$$

where $\bar{x}$ is the class of $x$ in $F_j$. By Hennion [9] or Furstenberg and Kifer [6], the exponent associated with (25) is the supremum of the exponents associated with the following SDE on $F_j$:

$$d\bar{x}_t = \alpha_j(T_0)\bar{x}_t\, dt + \sum_{i=1}^{r} \alpha_j(T_i)\bar{x}_t \circ db_t^i, \qquad \bar{x}_t \in F_j,$$

for $j = 1, \dots, p$.

Making use of the minimality assumption on $E_j$ and of the convergence of $1/n\Sigma_{i=1}^n \tau^i$ to $\delta_0$, it follows from Lemma 7.2 that each of these exponents is strictly negative. This shows that (ii) holds.

We now verify the converse. We can suppose $Q = I$ too. The assumption (i) implies that $\{k(M);\ M \in G\}$ is in $O(d_2)$, so for $\mu = \mu_1$, $\{k(M);\ M \in T_\mu\}$ is bounded. By (ii) the Liapounov exponents of the image of $\mu$ under $a(\cdot)$ and $b(\cdot)$ are strictly negative. Therefore, Theorem 4.1 implies that $\{\mu_n, n \geq 1\}$ is tight on $\mathscr{M}(d)$. Since by continuity $\{\mu_s, 0 \leq s \leq 1\}$ is tight and since each $\mu_t$, $t \geq 0$, can be written $\mu_t = \mu_n * \mu_s$ with $n$ in $\mathbb{N}$ and $0 \leq s \leq 1$, $\{\mu_t, t \geq 0\}$ is tight. □

This theorem shows clearly that when the upper Liapounov exponent is 0 instability is the rule. For instance $S_0$ must have at least one eigenvalue with real part 0. We also have:

COROLLARY 7.4. *If* $\mathrm{trace}(S_0) \geq 0$, *the solution* $x_t \equiv 0$ *is stable in probability if and only if for some invertible* $Q$, *all the matrices* $QS_iQ^{-1}$, $0 \leq i \leq r$, *are skew-symmetric.*

PROOF. If $x_t \equiv 0$ is stable in probability we can apply Theorem 7.3. We know from Theorem 2.2 of Arnold, Crauel, and Wihstutz [1] that

$$\operatorname{tr} a(S_0') \le d_1 \gamma(a(S_0'), \dots, a(S_r')) < 0$$

and

$$\operatorname{tr} b(S_0') \le d_3 \gamma(b(S_0'), \dots, b(S_r')) < 0.$$

Therefore if $d_1$ or $d_2$ is nonzero then

$$\operatorname{tr}(S_0') = \operatorname{tr} a(S_0') + \operatorname{tr} b(S_0') + \operatorname{tr} k(S_0') < \operatorname{tr} k(S_0').$$

But $k(S_0')$ is skew-symmetric and thus has trace 0. Hence if $\operatorname{tr}(S_0) = \operatorname{tr}(S_0') \ge 0$, then $d_1 = d_3 = 0$ and all the matrices $QS_iQ^{-1}$ are skew-symmetric. The converse is obvious. □

We may paraphrase this corollary by saying that if $\operatorname{tr}(S_0) \ge 0$ (for instance if $S_0 = 0$) the only linear diffusions on $\mathbb{R}^d$ which are stable in probability are (possibly degenerate) Brownian motions on spheres (for a convenient scalar product).

To study the ergodic properties of stable linear SDE we have the following analogue of Proposition 6.1 and Theorem 6.3:

PROPOSITION 7.5. *Suppose that the solution $x_t \equiv 0$ of (1) is stable in probability. By Theorem 7.3, if we suppose for convenience that $Q = I$, we can write the solution $M_t$ of (23) as*

$$M_t = \begin{pmatrix} a(M_t) & c(M_t) & e(M_t) \\ 0 & k(M_t) & d(M_t) \\ 0 & 0 & b(M_t) \end{pmatrix}.$$

*Set*

$$R_t = \begin{pmatrix} 0 & I & k(M_t)^{-1} d(M_t) \end{pmatrix} \in \mathscr{M}(d_2, d), \qquad K_t = k(M_t) \in O(d_2)$$

*and*

$$L_t = \begin{pmatrix} c(M_t) k(M_t)^{-1} \\ I \\ 0 \end{pmatrix} \in \mathscr{M}(d, d_2).$$

*Then*

(i) *$M_t - L_t K_t R_t$ converges to 0 almost surely.*
(ii) *$R_t$ converges a.s. to a random matrix $R$.*
(iii) *$L_t$ converges in law to some distribution $\lambda$ on $\mathscr{M}(d, d_2)$.*
(iv) *$K_t$ is a possibly degenerate Brownian motion on $SO(d_2)$.*
(v) *$R_t$, $K_t$, and $L_t$ are asymptotically independent.*

OUTLINE OF THE PROOF. The proof is an easy adaptation of the proof of Proposition 6.1. We just indicate the main modifications. As there, we first verify that $r'(M_t) = k(M_t)^{-1}d(M_t)$ converges a.s. It is easily seen that if $[t]$ is the integral part of $t$,

$$\left\| r'(M_t) - r'\left(M_{[t]}\right) \right\| \le X_{[t]} \left\| b\left(M_{[t]}\right) \right\| \tag{27}$$

if for each integer we define

$$X_n = \sup\left\{ \left\| d\left(M_t M_n^{-1}\right) \right\|,\ n \le t < n+1 \right\}.$$

Since (see for instance Ikeda and Watanabe [12], page 240)

$$E\left\{ \sup_{0 \le t \le 1} \|M_t\| \right\} < \infty \tag{28}$$

the Borel–Cantelli lemma implies that a.s. for $t$ large enough $X_{[t]} \le [t]$. We know that $r'(M_{[t]})$ converges a.s. Since the exponent of $b(M_t)$ is strictly negative, $r'(M_t)$ has the same limit by (27).

To prove (iii) we remark that if $N_t$ is the solution of

$$dN_t = N_t S_0\, dt + \sum_{i=1}^{r} N_t S_i \circ db_t^i, \qquad N_0 = I,$$

and if $L'_t$ is associated with $N_t$ (in the same way as $L_t$ is associated with $M_t$), $L_t$ and $L'_t$ have the same distributions for each $t$. One proves as above that $L'_t$ converges a.s. The other points are proved exactly as in Section 6 using (28). □

As in Section 6 we have:

COROLLARY 7.6. *If $x_t(y)$ is the solution of* (1) *such that $x_0(y) = y$, under the notation and hypotheses of Proposition* 7.5, *for any bounded continuous function $f$ on $\mathbb{R}^d$ and every $y$,*

$$\frac{1}{t}\int_0^t f(x_s(y)(\omega))\, ds \to \int f(LUR(\omega)y)\, d\lambda(L)\, d\rho(U) \quad a.s.,$$

*if $\rho$ is the Haar measure on the closure of the Lie subgroup of $SO(d_2)$ whose Lie algebra is generated by $k(S_0), \dots, k(S_r)$.*

We have of course the same description of the behaviour of the path of the Markov process $x_t$ as at the end of Section 6. Let us give a typical example in $\mathbb{R}^3$.

Consider the matrices $S_0, S_1, \dots, S_r$ for which

$$S_i = \begin{pmatrix} a_i & * & * \\ 0 & 0 & * \\ 0 & 0 & b_i \end{pmatrix}, \qquad a_i \in \mathbb{R},\ b_i \in \mathbb{R},\ i = 0, \dots, r,$$

with $a_0 < 0$ and $b_0 < 0$. In this case $x_t \equiv 0$ is stable in probability and for each $y$ the distribution of $x_t(y)$ converges. In law $M_t$ converges to some random matrix which can be written

$$\begin{pmatrix} 0 & C & CD \\ 0 & 1 & D \\ 0 & 0 & 0 \end{pmatrix}, \qquad C, D \in \mathbb{R},$$

and $d(M_t)$ converges to $D$ almost surely.

If $y = (u_0, v_0, w_0)$ is in $\mathbb{R}^3$:

(i) for $v_0 = w_0 = 0$, $x_t(y)$ converges to 0 exponentially fast a.s.;
(ii) for $w_0 = 0$, $x_t(y)$ remains on the line $\{(u, v_0, 0),\ u \in \mathbb{R}\}$ and is recurrent;
(iii) for $w_0 \neq 0$, $x_t(y)$ is attracted by the random line $\{(u, D, 0),\ u \in \mathbb{R}\}$.

Finally we can describe the set of invariant distributions of the Markov process solution of (1), and this gives the stationary solutions in the general case (i.e., without any assumption of stability). Using the results of Section 5 and the fact that such an invariant distribution is $\int_0^\infty e^{-s}\mu_s\, ds$-invariant we have:

PROPOSITION 7.7. *Consider the linear SDE* (1). *Let $E$ be the maximal subspace of $\mathbb{R}^d$ invariant under each $S_i$, $0 \le i \le r$, such that if $\tilde{S}_i$ is the restriction of $S_i$ to $E$ we can write in a convenient basis of $E$,*

$$\tilde{S}_i = \begin{pmatrix} a(S_i) & c(S_i) \\ 0 & k(S_i) \end{pmatrix},$$

*where $\gamma(a(S_0), \dots, a(S_r)) < 0$ and $k(S_i)$ is skew-symmetric, $0 \le i \le r$. Then every stationary solution of* (1) *is carried by $E$.*

On $E$ the invariant distributions are given as in Proposition 7.5. If for instance some $S_i$ has no eigenvalue with null real part, the only stationary solution of (1) is $x_t \equiv 0$.

## REFERENCES


[1] ARNOLD, L., CRAUEL, H. and WIHSTUTZ, V. (1983). Stabilization of linear systems by noise. *SIAM J. Control Optim.* **21** 451–461.
[2] CHEVALLEY, C. (1951). *Théorie des Groups de Lie.* **2**, *Groupes Algébriques*. Hermann, Paris.
[3] CONZE, J. P. and GUIVARC'H, Y. (1974). Remarques sur la distalité dans les espaces vectoriels. *C.R. Acad. Sci. Paris Sér. A* **278** 1083–1086.
[4] FELLER, W. (1968). *An Introduction to Probability Theory and Its Applications* **1**, 3rd ed. Wiley, New York.
[5] FURSTENBERG, H. (1963). Non commuting random products. *Trans. Amer. Math. Soc.* **108** 377–428.
[6] FURSTENBERG, H. and KIFER, Y. (1983). Random matrix products and measures on projective spaces. *Israel J. Math.* **46** 12–32.
[7] GUIVARC'H, Y. and RAUGI, A. (1985). Frontière de Furstenberg, propriétés de contraction et théorèmes de convergences. *Z. Wahrsch. verw. Gebiete* **69** 187–242.
[8] HASHMINSKI, R. Z. (1980). *Stochastic Stability of Differential Equations*. Sijthoff and Noordhoff, Alphen aan den Rijn.
[9] HENNION, H. (1984). Loi des grandes nombres et perturbations pour des produits réductibles de matrices aléatoires indépendantes. *Z. Wahrsch. verw. Gebiete* **67** 265–278.
[10] HEWITT, E. and ROSS, K. (1963). *Abstract Harmonic Analysis*. Springer, Berlin.

[11] HÖRMANDER, L. (1967). Hypoelliptic second order differential equations. *Acta Math.* **119** 147–171.

[12] IKEDA, N. and WATANABE, S. (1981). *Stochastic Differential Equations and Diffusion Processes*. Kodanska, Tokyo/North-Holland, Amsterdam.

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

[14] LEDRAPPIER, F. (1985). Quelques propriétés des exposants caractéristiques. *Ecole d'Été de Probabilités de Saint Flour. Lecture Notes in Math.* **1097** 306–396. Springer, Berlin.

[15] LE PAGE, E. (1982). Théorèmes limites pour les produits de matrices aléatoires. *Lecture Notes in Math.* **928** 258–303. Springer, Berlin.

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

U.E.R. DE MATHÉMATIQUES
UNIVERSITÉ PARIS 7
2, PLACE JUSSIEU
75251 PARIS CEDEX
FRANCE