The Annals of Probability
1984, Vol. 12, No. 2, 283–310

SPECIAL INVITED PAPER

THE STABILITY OF LARGE RANDOM MATRICES AND THEIR PRODUCTS

BY JOEL E. COHEN AND CHARLES M. NEWMAN

Rockefeller University and University of Arizona

Let $A(1), A(2), \cdots$ be a sequence of independent identically distributed (i.i.d.) random real $n \times n$ matrices and let $x(t) = A(t)x(t-1)$, $t = 1, 2, \cdots$. Define $\overline{\lambda}_n = \sup\{\lim_{t\uparrow\infty} \| x(t) \|^{1/t}: 0 \neq x(0) \in R^n\}$ where $\| \cdot \|$ denotes, e.g. the Euclidean norm, providing the limit exists almost surely (a.s.) and is nonrandom, and define $\underline{\lambda}_n$ analogously with sup replaced by inf. If all n^2 entries of each $A(t)$ are i.i.d. standard symmetric stable random variables of exponent α, then $\underline{\lambda}_n = \overline{\lambda}_n = \lambda_n(\alpha)$. In the standard normal case ($\alpha = 2$), $\lambda_n(2) = (2 \exp[\psi(n/2)])^{1/2}$, where ψ is the digamma function, and $n^{-1/2}\lambda_n(2) \to 1$; for $0 < \alpha < 2$, $(n \log n)^{-1/\alpha}\lambda_n(\alpha)$ converges to $\{2\,\Gamma(\alpha) \sin(\alpha\pi/2)/[\alpha\pi]\}^{1/2}$.

Criteria for stability ($\overline{\lambda}_n < 1$) and instability ($\underline{\lambda}_n > 1$) are investigated for more general distributions of $A(t)$. We obtain, for example, the general bound, $\overline{\lambda}_n \leq \{r[E(A(1)^T A(1))]\}^{1/2}$, where A^T is the transpose of A and r denotes the spectral radius. In the case of independent entries of mean zero and common variance s^2/n, this leads to $\lim\sup_n \overline{\lambda}_n \leq s$. If the entries of $A(t)$ are i.i.d. and distributed as $W/n^{1/2}$ where W is independent of n, has mean zero, variance s^2 and satisfies $E(\exp[iuW]) = O(|u|^{-\delta})$ as $|u| \uparrow \infty$ for some $\delta > 0$, then $\lim\inf_n \underline{\lambda}_n \geq s$.

These conditions for the asymptotic stability or instability of a product of random matrices are of the form originally proposed by May for differential equations governed by a single random matrix. We give counterexamples to show that May's criteria for the system of linear ordinary differential equations that he considered are not valid in the generality originally proposed, nor are they valid for the related system of difference equations considered by Hastings. The validity of May's criteria for these systems under more restrictive hypotheses remains an open question.

1. Introduction. Ecologists have long been concerned with the question of whether ecological communities or ecosystems that are more complex, in some sense, are also more stable, in some sense. In the past decade, much theoretical and empirical investigation has been stimulated by May's (1972) proposal of a specific quantitative relationship between complexity and stability within the framework of a mathematically explicit ecological model. May's model is described in Section 4. The results presented here arose from the desire to generalize May's model from a randomly constructed dynamical system with coefficients that are fixed in time to a randomly constructed dynamical system with coefficients that vary randomly in time.

Received July 1983; revised December 1983.

[1] This paper is based on a special invited lecture given at the Arcata, California meeting of the Institute of Mathematical Statistics, June 15, 1983.

AMS 1980 *subject classification.* Primary 60B15; secondary 60H25, 92A17.

Key words and phrases: Products of random matrices, stability of ecosystems, normal random matrices, symmetric stable random matrices, strong law, central limit theorem, stability of random linear ordinary differential equations, Liapunov exponent.

For each positive integer n, let $A(t)$, $t = 1, 2, \cdots$, be an infinite sequence of random real $n \times n$ matrices with elements $A(t)_{ij}$, $i, j = 1, \cdots, n$, and let $x(0)$ be a nonzero vector in R^n. Define

$$(1.1) \qquad x(t) = A(t) \cdots A(1)x(0), \quad t = 1, 2, \cdots$$

and letting $\| x \|$ denote $(\sum_{j=1}^{n} | x_j |^p)^{1/p}$ for some fixed $p \in (0, \infty)$, define

$$(1.2) \qquad \log \lambda(x(0)) = \lim_{t\uparrow\infty} t^{-1} \log \| x(t) \|,$$

$$(1.3) \qquad \underline{\lambda} = \inf\{\lambda(x(0)): 0 \neq x(0) \in R^n\}, \quad \overline{\lambda} = \sup\{\lambda(x(0)): 0 \neq x(0) \in R^n\}.$$

We assume in (1.2) that for each $x(0)$, the limit exists almost surely (a.s.) and is nonrandom. The inf and sup in (1.3) are taken over these nonrandom values. For a fixed n, we shall say that the system $\{A(t)\}_{t=1}^{\infty}$ is *strongly stable* if $\log \overline{\lambda} < 0$ and is *strongly unstable* if $\log \underline{\lambda} > 0$. As $n \uparrow \infty$, we shall say that the sequence of systems is *asymptotically strongly stable* (resp. *unstable*) if there exists a positive integer N such that strong stability (resp. instability) is valid for each $n > N$.

If the limit in (1.2) exists a.s., it is independent of p for all $0 < p < \infty$ because $\max_j | x_j | \leq \| x \| \leq n^{1/p} \max_j | x_j |$ and $\lim_{t\uparrow\infty} n^{1/(pt)} = 1$.

In many situations, it can be shown that the limit of (1.2) exists (a.s.) and it is nonrandom (see Furstenberg and Kesten, 1960, Furstenberg, 1963). We will be concerned entirely with such situations in this paper and particularly with those where, for each n, the $A(t)$'s are independent and identically distributed. The first purpose of this paper is to point out a number of examples of that type where $\underline{\lambda} = \overline{\lambda}$ and an explicit expression for $\lambda = \underline{\lambda} = \overline{\lambda}$ can be obtained. In certain cases, λ can be calculated in closed form, e.g., when the entries of each $A(t)$ are i.i.d. normal with mean zero. To our knowledge, these are the first nontrivial (although simple) cases in which $\log \lambda$ has been computed explicitly for general n. In all the examples, conditions can be determined which imply asymptotic strong stability (resp. instability).

After having obtained our results concerning these examples, we discovered that they had been partially anticipated by Girko (1976) who considered essentially the same set of examples, but with a different motivation. He gave neither the explicit expressions for λ (see (2.7) and (2.16) below) nor their asymptotic behavior (see (2.10)–(2.11) and (2.19)–(2.21) below).

The second purpose of this paper is to extend the conditions for asymptotic strong stability and instability to more general products of i.i.d. random matrices where $\lambda(x(0))$ may not be explicitly expressible. These conditions are of the form proposed by May (1972) for a linear system of n ordinary differential equations governed by a single randomly chosen matrix of coefficients fixed in time.

The third purpose of this paper is to show by a number of counterexamples that the criteria proposed by May, in the generality with which he stated them, may fail for the system he considered. The related but slightly different "theorems" of Hastings (1982a, 1982b) are similarly false. One of the systems considered by Hastings is the degenerate special case of (1.1) in which $A(t) = A$, a fixed t-independent random $n \times n$ matrix. In this case, $\overline{\lambda}$ is replaced by the (random) spectral radius $r(A)$ (maximum modulus of the eigenvalues) of A. Rather than considering asymptotic strong stability or instability as defined above, one is interested in criteria which imply that lim sup $r(A) < 1$ or lim inf $r(A) > 1$ (either in probability or a.s.). The May criteria may yet turn out to be correct under hypotheses which are not excluded by our counterexamples.

2. Some explicit examples. We begin with a general proposition which leads to all our examples. We define for $v \geq 0$,

$$\text{(2.1)} \qquad \log_+(v) = \begin{cases} \log v, & v \geq 1 \\ 0, & 0 \leq v < 1, \end{cases} \qquad \log_-(v) = \begin{cases} 0, & v \geq 1 \\ -\log v, & 0 < v < 1 \\ +\infty, & v = 0. \end{cases}$$

Throughout this section $A(t)$, $t = 1, 2, \cdots$ will denote a sequence of i.i.d. $n \times n$ real matrices.

PROPOSITION 2.1. *Suppose that* $V_x = \| A(1)x \| / \| x \|$ *has a distribution which does not depend on* x *in* $R^n \backslash \{0\}$. *Then the limit in* (1.2) *exists, is nonrandom, is independent of* $x(0)$, *and is given by*

$$\text{(2.2)} \qquad \log \lambda = E[\log(\| A(1)x \| / \| x \|)]$$

providing $\log_+(V_x)$ *or* $\log_-(V_x)$ *or both have finite expectation. If* $E(\log^2(V_x)) < \infty$, *then for any* $x(0) \neq 0$, $t^{-1/2}(\log \| x(t) \| - t \log \lambda)$ *converges in distribution as* $t \uparrow \infty$ *to a normal random variable of mean zero and variance equal to the variance of* $\log V_x$.

PROOF. Either $V_x > 0$ w.p. 1 or $V_x > 0$ w.p. < 1. In the second case the proposition is true with $\log \lambda = -\infty$. Henceforth we assume the first case. It follows from the lack of dependence of V_x on x that $\| A(t)x(t-1) \| / \| x(t-1) \|$ is independent of $\{x(t'): t' < t\}$ and hence of $\{\| x(t') \| / \| x(t'-1) \|: t' < t\}$. Expressing

$$\text{(2.3)} \qquad t^{-1}\log \|x(t)\| - t^{-1}\log \| x(0) \| = t^{-1} \textstyle\sum_{u=1}^{t} \log[\| x(u) \| / \| x(u-1) \|],$$

we see that the summands on the right side of (2.3) are i.i.d. The desired result now follows from the standard law of large numbers and central limit theorem. Note that (2.2) also follows immediately from Corollary 3.2 below.

In order to apply Proposition 2.1 when $A(1)$ is normally distributed, we note the following.

LEMMA 2.2. *Suppose* Y *is a mean zero normal vector in* R^n *with covariance matrix* C. *Let* $\| \cdot \|$ *denote the Euclidean* ℓ_2 *norm. Then the distribution of* $\| Y \|$ *depends only on the eigenvalues of* C *(and their multiplicities). Moreover* $\log(\| Y \|)$ *has finite mean and variance unless* $C = 0$, *and*

$$\text{(2.4)} \qquad \begin{aligned} E[\log(\| Y \|)] &= \tfrac{1}{2} E[\log(\textstyle\sum_{i=1}^{n} c_i Z_i^2)] \\ \mathrm{Var}[\log(\| Y \|)] &= \tfrac{1}{4} \mathrm{Var}[\log(\textstyle\sum_{i=1}^{n} c_i Z_i^2)], \end{aligned}$$

where the c_i*'s are the eigenvalues of* C *and the* Z_i*'s are* i.i.d. *standard normal random variables.*

Proof. By an orthogonal transformation leaving $\|\cdot\|$ invariant, Y can be transformed to a mean zero normal vector with a diagonal covariance matrix whose diagonal entries are the c_i's. Assume $C \neq 0$. To see that $\log(\|Y\|)$ has finite absolute expectation, we may bound $\log_+(\|Y\|)$ above by $\frac{1}{2}\log_+(c_*(Z_1^2 + \cdots + Z_n^2))$ where $c_* = \max\{c_i: i = 1, \cdots, n\}$. We may bound $\log_-(\|Y\|)$ above by $\frac{1}{2}\log_-(c_* Z_*^2)$ where $Z_* = Z_j$ with j such that $c_* = c_j$. The finiteness of the moments of these quantities follows from the next proposition.

Lemma 2.3. *Suppose W has a chi squared distribution with $n > 0$ degrees of freedom ($W = Z_1^2 + \cdots + Z_n^2$). Then $\log W$ has finite mean and variance with*

$$(2.5)\qquad \begin{aligned} E(\log W) &= \log 2 + \psi(n/2) = \log 2 + \Gamma'(n/2)/\Gamma(n/2), \\ \mathrm{Var}(\log W) &= \psi'(n/2) = [\Gamma''(n/2)/\Gamma(n/2)] - [\Gamma'(n/2)/\Gamma(n/2)]^2, \end{aligned}$$

where $\Gamma(r)$ is the gamma function and $\psi(r) = \Gamma'(r)/\Gamma(r)$ is the digamma function.

Proof. Using the standard formula for the density of W, we have

$$\begin{aligned} E(W^b) &= [2^{n/2}\Gamma(n/2)]^{-1} \int_0^\infty w^b w^{n/2-1} e^{-w/2}\, dw \\ &= [\Gamma(n/2)]^{-1} 2^b \Gamma(n/2 + b) = \exp[(\log 2)b + \phi(n/2 + b) - \phi(n/2)], \end{aligned}$$

where $\phi(b) = \log \Gamma(b)$. Since $E(W^b) = E(\exp[b \log W])$ is the moment generating function for $\log W$, we know from the finiteness of $E(W^b)$ in a neighborhood of $b = 0$ that all moments of $\log W$ exist and in particular that

$$\begin{aligned} E(\log W) &= (d/db)E(W^b)\,|_{b=0} = (d/db) \log E(W^b)\,|_{b=0} \\ &= \log 2 + \phi'(n/2) = \log 2 + \psi(n/2), \\ \mathrm{Var}(\log W) &= (d^2/db^2) \log E(W^b)\,|_{b=0} = \phi''(n/2) = \psi'(n/2) \end{aligned}$$

as desired.

Theorem 2.4. *Suppose $\{A(1)_{ij}: i, j = 1, \cdots, n\}$ are jointly normal mean zero random variables. Define the $n \times n$ matrix $C^{k\ell}$ by*

$$(C^{k\ell})_{ij} = [\mathrm{Cov}(A(1)_{ik}, A(1)_{j\ell}) + \mathrm{Cov}(A(1)_{i\ell}, A(1)_{jk})]/2.$$

Suppose that the matrices

$$(2.6)\qquad C(x) \equiv \textstyle\sum_{k,\ell=1}^n x_k x_\ell C^{k\ell}, \quad x \in R^n, \quad \sum_{k=1}^n (x_k)^2 = 1,$$

are isospectral, i.e. have the same eigenvalues and multiplicities, $c_1, \cdots, c_n$, independent of x. ($C(x)$ is the covariance matrix of $A(1)x$.) Then Proposition 2.1 applies with $\|\cdot\|$ taken as the ℓ_2 norm, and $\log \lambda$ and $\mathrm{Var}(\log V_x)$ are given by (2.4).

Proof. This is an immediate consequence of Proposition 2.1 and Lemma

2.2 and the fact that

$$\begin{aligned}\mathrm{Cov}((A(1)x)_i, (A(1)x)_j) &= \textstyle\sum_{k,\ell=1}^n x_k x_\ell \mathrm{Cov}(A(1)_{ik}, A(1)_{j\ell}) \\ &= C(x)_{ij}.\end{aligned}$$

We remark that a mean zero normal $A(1)$ that does not directly satisfy the hypotheses of Theorem 2.4 can sometimes be transformed to one that does by applying a nonsingular transformation T to R^n which transforms $A(1)$ to $TA(1)T^{-1}$. Equivalently one can replace the norm $\| x \|$ by the norm $\| Tx \|$. This extension should be kept in mind when assessing the generality of Theorem 2.4. A simple situation in which the $C(x)$'s are isospectral occurs when the *columns* of $A(1)$ are independent (so that $C^{k\ell} = 0$ for $k \neq \ell$) and identically distributed. A somewhat more complicated situation can be constructed by replacing such an $A(1)$ by $TA(1)T^{-1}$ with T orthogonal. The next theorem describes the simplest of all situations.

THEOREM 2.5. *Suppose* $\{A(1)_{ij}: i, j = 1, \cdots, n\}$ *are* i.i.d. $N(0, s^2)$. *Then the limit in* (1.2) *exists, is nonrandom, is independent of* $x(0)$, *and is given by*

(2.7) $$\log \lambda = \tfrac{1}{2}[\log(s^2) + \log 2 + \psi(n/2)].$$

Moreover, for any $x(0) \neq 0$, $t^{-1/2} \log(\lambda^{-t} \| x(t) \|)$ *converges in distribution to* $N(0, \sigma^2)$ *with*

(2.8) $$\sigma^2 = \tfrac{1}{4}\psi'(n/2).$$

Thus $\{A(t)\}$ *is strongly stable* (*resp. unstable*) *if*

(2.9) $$s^2 < (\text{resp.} >)[e^{-\psi(n/2)}]/2.$$

A sequence of such $n \times n$ *systems with* $s^2 = s_n^2$ *is asymptotically strongly stable if*

(2.10) $$\lim\sup_{n\uparrow\infty} n^{1/2} s_n < 1$$

and asymptotically strongly unstable if

(2.11) $$\lim\inf_{n\uparrow\infty} n^{1/2} s_n > 1.$$

PROOF. This is an immediate consequence of Theorem 2.4, Lemma 2.3 and the asymptotic expansion (see Abramowitz and Stegun, 1964, page 259)

(2.12) $$\psi(n/2) = \log(n/2) - 1/n + O(1/n^2).$$

REMARK. The exact finite n criterion (2.9) can be combined with (2.12) to give finer asymptotic criteria than (2.10)–(2.11). For example, if $ns_n^2 = 1 + (g/n) + o(1/n)$, then $g < 1$ (resp. $g > 1$) implies asymptotic strong stability (resp. instability).

The proof of Theorem 2.5 (especially the analysis leading to Proposition 2.1) yields not only $\log \lambda$, which describes the asymptotic behavior of $\| x(t) \|$ for large t, but also the explicit distribution of $\| x(t) \|$ for finite t. In particular, if

$r > -n/2$,

$$E([\| x(t) \| / \| x(0) \|]^r) = \{2^{r/2}\Gamma([n + r]/2)/\Gamma(n/2)\}^t.$$

For the model of Theorem 2.5, we define the critical variance s_c^2 as $\frac{1}{2}\exp[-\psi(n/2)]$, since $\{A(t)\}$ is strongly stable for $s^2 < s_c^2$ and strongly unstable for $s^2 > s_c^2$. Here are some illustrative values.

n	s_c^2	ns_c^2
1	3.5621	3.562
2	0.8905	1.781
10	0.1109	1.109
100	0.0101	1.010

The criteria (2.10)–(2.11) are of the form given, for a different random system, by May (1972). In the next section, we extend these criteria to a much larger class of distributions for which no explicit expression for λ is obtainable for finite n. Before that, however, we give a further example where an expression for λ can be obtained which leads to quite different criteria for asymptotic stability than (2.10)–(2.11). In this example, the normal distribution of Theorem 2.5 is replaced by a symmetric stable distribution of exponent α in (0, 2] (see e.g. Feller, 1971). The normal distribution corresponds to $\alpha = 2$.

We say that W is a *standard symmetric stable random variable of exponent* α if

$$E(\exp(irW)) = \exp[-|r|^\alpha/\alpha]. \tag{2.13}$$

For the rest of this section $\| x \|$ denotes $(\sum_{j=1}^n |x_j|^\alpha)^{1/\alpha}$.

LEMMA 2.6. *Suppose* $W_0, W_1, W_2, \cdots$ *are* i.i.d. *standard symmetric stable random variables of exponent* α *and* $U = (W_1, \cdots, W_n)$. *Then* $\sum_{j=1}^n y_j W_j$ *has the same distribution as* $\| y \| W_0$, *and* $\log(\| U \|)$ *has finite mean and variance.*

PROOF. The first statement is a standard fact, proved thus:

$$\begin{aligned} E(\exp[ir \textstyle\sum_{j=1}^n y_j W_j]) &= \textstyle\prod_{j=1}^n E(\exp[iry_j W_j]) = \prod_{j=1}^n \exp[-|ry_j|^\alpha/\alpha] \\ &= \exp[-\| y \|^\alpha |r|^\alpha/\alpha] = \exp[-(|r| \, \| y \|)^\alpha/\alpha] \\ &= E(\exp[ir \| y \| W_0]). \end{aligned}$$

To prove the second statement, we write

$$\log(\| U \|) = \alpha^{-1}\{\log([1/n] \textstyle\sum_{j=1}^n |W_j|^\alpha) + \log n\}. \tag{2.14}$$

As in the proof of Lemma 2.3, it suffices to show that the moment generating function for $\log(\sum_{j=1}^n |W_j|^\alpha/n)$,

$$E((n^{-1} \textstyle\sum_{j=1}^n |W_j|^\alpha)^b), \tag{2.15}$$

is finite in an interval about $b = 0$. For $b \le 0$, we may use the convexity of

$f_b(u) = u^b$ to bound $f_b(n^{-1} \sum_{j=1}^n |W_j|^\alpha)$ by $n^{-1} \sum_{j=1}^n f_b(|W_j|^\alpha)$. For $0 < b \le 1$, $f_b(n^{-1} \sum_{j=1}^n |W_j|^\alpha) \le n^{-b} \sum_{j=1}^n f_b(|W_j|^\alpha)$ since $|x + x'|^b \le \max(1, 2^{b-1})\{|x| + |x'|^b\}$. Together these imply that it suffices to show that $E(|W_1|^{\alpha b}) < \infty$ for b in $(-\varepsilon, 1)$ for some $\varepsilon > 0$. It will be shown in Lemma 3.9 that for $0 < c < 1$, $E(|W_1|^{-c}) < \infty$ if

$$\int_{-\infty}^{\infty} |t|^{c-1} |E(e^{itW_1})| \, dt < \infty.$$

This is the case (using (2.13)) for any $c < 1$ and thus $E(|W_1|^{\alpha b}) < \infty$ for $b \in (-1/\alpha, 0]$. On the other hand (see Feller, 1971, and formula (2.26) below) it is a standard fact about W_1 that if $\alpha \ne 2$ and $g > 0$, $E(|W_1|^g) < \infty$ if and only if $g < \alpha$. Thus $E(|W_1|^{\alpha b}) < \infty$ for $b \in [0, 1)$ which completes the proof.

THEOREM 2.7. *Suppose* $\{A(1)_{ij}: i, j = 1, \cdots, n\}$ *are* i.i.d. *with* $A(1)_{ij}/s$ *being standard symmetric stable with exponent* $\alpha \in (0, 2]$, *while* $s > 0$. *Then the limit in* (1.2) *exists, is nonrandom, is independent of* $x(0)$, *and is given by*

$$\log \lambda = \alpha^{-1}[\log(s^\alpha) + E(\log(|W_1|^\alpha + \cdots + |W_n|^\alpha))]. \tag{2.16}$$

Moreover, for any $x(0) \ne 0$, $t^{-1/2} \log(\lambda^{-t} \| x(t) \|)$ *converges in distribution to* $N(0, \sigma^2)$ *with*

$$\sigma^2 = \alpha^{-2} \operatorname{Var}(\log(|W_1|^\alpha + \cdots + |W_n|^\alpha)).$$

Thus $\{A(t)\}$ *is strongly stable* (*resp. unstable*) *if*

$$s < (\text{resp. } >) e^{-Q_n(\alpha)} \tag{2.17}$$

where

$$Q_n(\alpha) = \alpha^{-1} E[\log(|W_1|^\alpha + \cdots + |W_n|^\alpha)]. \tag{2.18}$$

A sequence of such $n \times n$ *systems with* $s^2 = s_n^2$ *and* $\alpha \ne 2$ *is asymptotically strongly stable if*

$$\limsup_{n\uparrow\infty} (n \log n)^{1/\alpha} s_n < J(\alpha) \tag{2.19}$$

where

$$J(\alpha) = \left[\alpha\pi \Big/ \left\{2\Gamma(\alpha)\sin\left(\frac{\alpha\pi}{2}\right)\right\}\right]^{1/\alpha} \tag{2.20}$$

and asymptotically strongly unstable if

$$\liminf_{n\uparrow\infty} (n \log n)^{1/\alpha} s_n > J(\alpha). \tag{2.21}$$

Figure 1 graphs $J(\alpha)$ for α in (0, 2).

PROOF. Each of the independent components of $A(1)x$ is of the form $\sum_{j=1}^n y_j W_j$ with $y_j = sx_j$ and $W_j = (A(1)_{ij}/s)$. By the first part of Lemma 2.6, each component has the same distribution as $s \| x \| W_0$. Thus $\| A(1)x \| = (\sum_{i=1}^n |(A(1)x)_i|^\alpha)^{1/\alpha}$ has the same distribution as $s \| x \| (\sum_{i=1}^n |W_i|^\alpha)^{1/\alpha}$. The

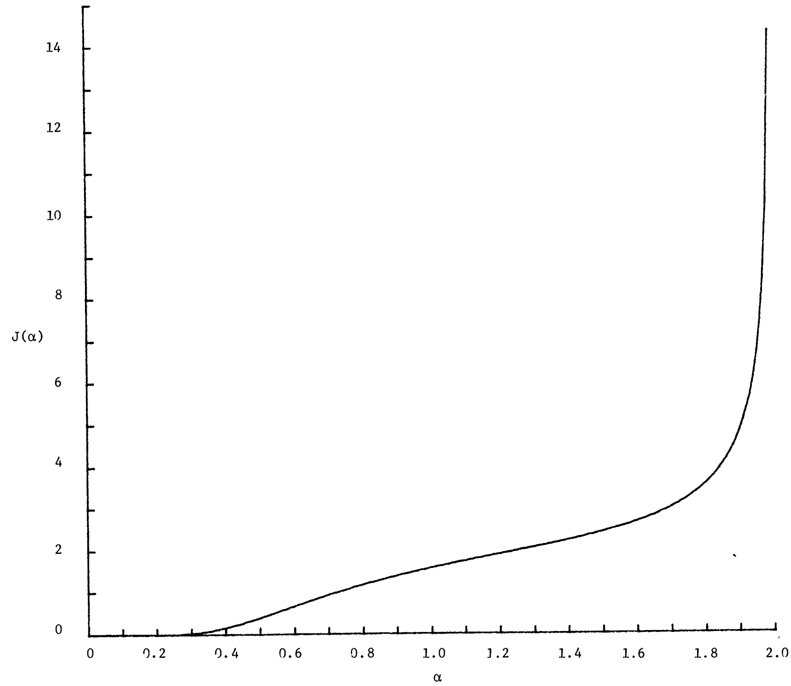

Fig. 1. $J(\alpha) = [\alpha\pi/\{2\Gamma(\alpha)\sin(\alpha\pi/2)\}]^{1/\alpha}$ as a function of α, for α in (0, 2).

theorem now follows from Proposition 2.1 except for (2.19)–(2.21). To prove these we need to show that for α in (0, 2),

$$\lim_{n\uparrow\infty}(n \log n)^{1/\alpha}e^{-Q_n(\alpha)} = J(\alpha), \tag{2.22}$$

or equivalently that

$$\lim_{n\uparrow\infty}E \log[(|W_1|^\alpha + \cdots + |W_n|^\alpha)/(n \log n)] = -\alpha \log(J(\alpha)). \tag{2.23}$$

This is an immediate consequence of the next proposition.

PROPOSITION 2.8. *Let* $W_1, W_2, \cdots$ *be* i.i.d. *standard symmetric stable random variables with exponent* α *in* (0, 2) *and let* $S_n = |W_1|^\alpha + \cdots + |W_n|^\alpha$. *Then*

$$S_n/(n \log n) \to K(\alpha) \ (\textit{in probability}) \ \textit{as} \ n \uparrow \infty \tag{2.23a}$$

and

$$\lim_{n\uparrow\infty}E(\log(S_n/[n \log n])) = \log[K(\alpha)], \tag{2.24}$$

where

$$K(\alpha) = 2\Gamma(\alpha)\sin(\alpha\pi/2)/[\alpha\pi]. \tag{2.25}$$

PROOF. Let $p(v)$ denote the density of $|W_1|^\alpha$. We first note that

$$p(v) = K(\alpha)v^{-2} + O(v^{-3}) \text{ as } v \uparrow \infty. \tag{2.26}$$

This is elementary for $\alpha = 1$ and follows from equation (6.8) (resp. equations (6.8) and (6.10)) of Feller (1971, page 583) for $0 < \alpha < 1$ (resp. $1 < \alpha < 2$). We next define, for $x > 0$,

$$\bar{S}_n = U_1 + \cdots + U_n \text{ where } U_i = |W_i|^\alpha \mathbf{1}_{\{|W_i|^\alpha \le nx\}}, \tag{2.27}$$

and observe that (2.26) implies that, as $n \uparrow \infty$,

$$P(S_n \ne \bar{S}_n) \le nP(U_1 \ne |W_1|^\alpha) = K(\alpha)/x + O(1/nx^2), \tag{2.28}$$

$$E[\bar{S}_n/(n \log n)] = K(\alpha)\log(nx)/\log n + O(1/\log n) \to K(\alpha), \tag{2.29}$$

$$\operatorname{Var}[\bar{S}_n/(n \log n)] \to 0. \tag{2.30}$$

It follows from (2.29), (2.30) and Tchebychev's inequality that $\bar{S}_n/(n \log n) \to K(\alpha)$ (in probability) and (2.23a) then follows since (2.28) is valid for arbitrarily large x. To obtain (2.24) from (2.23a) it suffices to bound the second moment of $\log[S_n/(n \log n)]$ uniformly in n, and to accomplish this, it suffices, as in the proof of Lemma 2.6, to show

$$\limsup_{n\uparrow\infty} E\{[S_n/(n \log n)]^b\} < \infty \tag{2.31}$$

both for some positive b and some negative b. It follows from (2.28), (2.29) and Tchebychev's inequality that for n, x sufficiently large,

$$\begin{aligned} P(S_n/[n \log n] > x) &\le E(\bar{S}_n/[n \log n])x^{-1} + P(S_n \ne \bar{S}_n) \\ &\le C \log(nx)/(x \log n) + C'/x \le (C''/x)\log x, \end{aligned} \tag{2.32}$$

for suitable constants C, C', C'', which implies (2.31) for $b \in (0, 1)$ by a standard argument. The proof is completed by showing (2.31) for $b = -1$. This follows by tedious but straightforward arguments from:

$$E\{[S_n/(n \log n)]^{-1}\} = \int_0^\infty \{g(t/[n \log n])\}^n \, dt, \tag{2.33}$$

$$g(t) \equiv E(\exp[-t|W_1|^\alpha]) \le \begin{cases} 1 - H_1 t|\log t|, & 0 < t < \tfrac{1}{2} \quad (2.34\text{a}) \\ H_2, & \tfrac{1}{2} \le t \le 2 \quad (2.34\text{b}) \\ t^{-\gamma}, & 2 < t < \infty \quad (2.34\text{c}) \end{cases}$$

for some H_1 in $(0, \infty)$, H_2 in $(0, 1)$ and γ in $(0, \infty)$. (2.33) follows from a standard identity for $E(X^{-1})$ for positive X, (2.34a) is a consequence of (2.26), (2.34b) is elementary, and (2.34c) is due to the fact that the density of W_1 is (analytic and hence) bounded near the origin so that $p(v) = O(v^{1/\alpha - 1})$ as $v \downarrow 0$.

REMARK. Generalizations of Theorem 2.7 analogous to Theorem 2.4 (regarded as a generalization of Theorem 2.5), or at least analogous to the situation considered in the remark following Theorem 2.4, are possible. For example, one could take the columns of $A(1)$ to be i.i.d. random n-vectors which are symmetric and stable in the sense that

$$E(\exp(i \sum x_k A(1)_{kj})) = \exp(-D(x))$$

where $D(sx) = |s|^\alpha D(x)$ for real s and $x = (x_1, \cdots, x_n) \in R^n$. The Lévy-Khintchine formula gives a complete characterization of such random vectors (see Feller, 1971). There may also exist generalizations of the criteria (2.19)–(2.21) to distributions of the elements of $A(1)$ that are in the domain of attraction of the symmetric stable laws. We will not pursue such generalizations here.

Other generalizations of Theorems 2.4 and 2.7 are possible. Suppose $\{A(t)\}$ and $\{B(t)\}$ are two mutually independent sequences of i.i.d. $n \times n$ real matrices, each of which satisfies the assumptions of Proposition 2.1, using the same norm for A and B. Let $\log \lambda_A$ and $\log \lambda_B$ denote the corresponding quantities in (2.2). Pick q in $[0, 1]$. Define the sequence $\{C(t)\}$ by $C(t) = A(t)$ w.p. q, $C(t) = B(t)$ w.p. $1 - q$, i.i.d. for all t. If $\log \lambda_C$ is the quantity in (2.2) with A replaced by C, then $\log \lambda_C = q \log \lambda_A + (1 - q) \log \lambda_B$. More generally, for a given norm, let S be the class of distributions on real $n \times n$ matrices C such that the distribution of $\| Cx \| / \| x \|$ is independent of the real nonzero n-vector x. Then S is a convex set and $\log \lambda$ is linear on S. Consequently, for example, if $\{A(t)\}$ had i.i.d. normal columns with one covariance matrix and $\{B(t)\}$ had i.i.d. normal columns with another covariance matrix, the Liapunov exponent $\log \lambda_C$ for the mixture could be found using Theorem 2.4. We thank Morris L. Eaton for suggesting that we try to extend our results to scale mixtures of normal matrices.

3. General stability criteria. The purpose of this section is to extend the stability criteria (2.10)–(2.11) to general situations in which $\lambda = \lambda(x(0))$ is not explicitly calculable. This turns out to be quite easy to accomplish, at least for the stability criterion (see Corollary 3.5 below).

Again in this section $A(t)$, $t = 1, 2, \cdots$ will denote a sequence of i.i.d. $n \times n$ real matrices and $x(t)$ will be as in (1.1). $\| \cdot \|$ will be the ℓ_2 norm and we let S^{n-1} denote $\{x \in R^n \colon \| x \| = 1\}$.

We begin with a result which is a consequence of the arguments of Furstenberg (1963) and of Furstenberg and Kifer (1983) [H. Furstenberg, private communication (Dec. 22, 1983)].

THEOREM 3.1. *Suppose that*

$$\sup_{x \in S^{n-1}} E[\log_+(\| A(1)x \|)] < \infty. \tag{3.1}$$

Then for any $x(0) \neq 0$, *the limit* (1.2) *exists, is nonrandom, and is given by*

$$\log \lambda(x(0)) = \int_{S^{n-1}} E[\log \| A(1)x \|] \mu(dx) \tag{3.2}$$

where μ *is some probability measure on* S^{n-1} *which may depend on* $x(0)$.

As an immediate corollary of Theorem 3.1, we have, using the definitions (1.3), a further result.

COROLLARY 3.2. *If* (3.1) *is valid, then*

$$\log \bar{\lambda} \le \sup_{x \in S^{n-1}} E[\log \| A(1)x \|], \tag{3.3}$$

$$\log \underline{\lambda} \ge \inf_{x \in S^{n-1}} E[\log \| A(1)x \|]. \tag{3.4}$$

THEOREM 3.3. *Suppose each $A(1)_{ij}$ has finite mean and variance and define C to be the positive semidefinite matrix $E(A(1)^T A(1))$, i.e.*

$$C_{ij} = \textstyle\sum_{k=1}^{n} E(A(1)_{ki} A(1)_{kj}). \tag{3.5}$$

Then (3.1) *is valid and*

$$\bar{\lambda} \le [r(C)]^{1/2}, \tag{3.6}$$

where $r(C)$ is the spectral radius (or maximum eigenvalue) of C.

In a previous version of this paper, we established a weaker form of Theorem 3.3 in which (3.6) was replaced by $\log \bar{\lambda} \le [r(C) - 1]/2$. The fact that our previous argument could be improved to yield the present (3.6) was noted independently by Eric Key and by us. (3.6) is sharp in that if $A(1)$ is nonrandom and symmetric, then $\bar{\lambda} = r(A(1)) = [r(C)]^{1/2}$. The last equality is not in general true if $A(1)$ is not symmetric.

PROOF. If $r(C) = 0$, then $A(1) = 0$ identically and both sides of (3.6) are 0. Assume $r(C) > 0$. If every element of $A(1)$ is multiplied by a positive constant K then both $\bar{\lambda}$ and $[r(C)]^{1/2}$ are multiplied by the same K. Take $K = [r(C)]^{-1/2}$. Evidently it entails no loss of generality to assume that $r(C) = 1$. We then use the elementary inequalities

$$\log_+(y) = \tfrac{1}{2}\log_+(y^2) \le \tfrac{1}{2}y^2,$$

$$\log y = \tfrac{1}{2}\log y^2 \le (y^2 - 1)/2,$$

the identity

$$E(\| A(1)x \|^2) = (x, Cx)$$

where $(\cdot, \cdot)$ denotes the standard inner product in R^n, and the fact that

$$\sup_{x \in S^{n-1}} (x, Cx) = r(C),$$

to conclude first that (3.1) is valid and second that

$$\sup_{x \in S^{n-1}} E(\log \| A(1)x \|) \le [r(C) - 1]/2 = 0.$$

It follows from Corollary 3.2 that $\bar{\lambda} \le 1 = [r(C)]^{1/2}$, which completes the proof.

COROLLARY 3.4. *Suppose each $A(1)_{ij}$ has zero mean and finite variance s_{ij}^2. Suppose further that* $\mathrm{Cov}(A(1)_{ki}, A(1)_{kj}) = 0$ *for $i \ne j$ and all k. Then*

$$\bar{\lambda} \le [\max\{(\textstyle\sum_{i=1}^{n} s_{ij}^2)^{1/2} : j = 1, \cdots, n\}]. \tag{3.7}$$

PROOF. Under the hypotheses we may apply Theorem 3.3 with C diagonal and $C_{jj} = \sum_{k=1}^{n} s_{kj}^2$.

COROLLARY 3.5. *Suppose each $A(1)_{ij}$ has zero mean, a common variance s^2 and $\mathrm{Cov}(A(1)_{ki}, A(1)_{kj}) = 0$ for $i \neq j$ and all k. Then $\{A(t)\}$ is strongly stable if $ns^2 < 1$. A sequence of such $n \times n$ systems is asymptotically strongly stable if* (2.10) *is valid.*

PROOF. Immediate from Corollary 3.4.

REMARK. Corollaries 3.4 and 3.5 are applicable both to the case where the $A(1)_{ij}$'s are i.i.d. for all i, j and to the case where the $A(1)_{ij}$'s are i.i.d. for $i \leq j$ while $A(1)_{ji} = A(1)_{ij}$ (so that $A(1)$ is a symmetric matrix) since in both of these situations $A(1)_{ki}$ and $A(1)_{kj}$ are pairwise independent for each k and $i \neq j$.

We now proceed to extend the instability criterion (2.11).

LEMMA 3.6. *For any nonnegative random variable Y, any $\varepsilon > 0$, and $0 < \eta < 1 < \eta' < \infty$, there exist positive finite constants K_ε and $K_{\varepsilon,\eta}$ such that*

$$E(\log_- Y) \leq K_\varepsilon E(Y^{-\varepsilon}), \tag{3.8}$$

$$E(\log Y) \geq -K_{\varepsilon,\eta}E(Y^{-\varepsilon}) + \log \eta \; P(Y < \eta') + \log \eta' \; P(Y \geq \eta'), \tag{3.9}$$

$$\lim_{\eta \downarrow 0} K_{\varepsilon,\eta} = 0. \tag{3.10}$$

PROOF. We let

$$K_{\varepsilon,\eta} = \sup_{0<y<\eta} y^\varepsilon |\log y|$$

and $K_\varepsilon = K_{\varepsilon,1}$. The conclusions then follow by elementary arguments.

PROPOSITION 3.7. *A sequence of $n \times n$ matrix systems $\{A^{(n)}(1), A^{(n)}(2), \cdots\}$ is asymptotically strongly unstable if for each n,* (3.1) *is valid with $A(1)$ replaced by $A^{(n)}(1)$, and if for some $\varepsilon > 0$ and $\eta' > 1$ and any sequence $x^{(n)} \in S^{n-1}$,*

$$\sup_n E(\| A^{(n)}(1)x^{(n)} \|^{-2\varepsilon}) < \infty \tag{3.11}$$

and

$$\lim_{n \uparrow \infty} P(\| A^{(n)}(1)x^{(n)} \|^2 < \eta') = 0. \tag{3.12}$$

PROOF. We apply Corollary 3.2 and Lemma 3.6 with $Y = \| A^{(n)}(1)x \|^2$ and note that the validity of (3.11) and (3.12) for *every* sequence $x^{(n)}$ implies the (seemingly stronger) results that

$$\sup_n (\sup_{x \in S^{n-1}} E(\| A^{(n)}(1)x \|^{-2\varepsilon})) < \infty$$

and

$$\lim_{n\uparrow\infty}\{\sup_{x\in S^{n-1}} P(\| A^{(n)}(1)x \|^2 < \eta')\} = 0.$$

We proceed to obtain conditions which imply (3.11). In the following lemma no assumption is made about the independence of the entries of $A^{(n)}(1)$.

LEMMA 3.8. *In the previous notation, if*

$$B_{ij}^{(n)} = n^{1/2} A_{ij}^{(n)}(1), \quad W_i^{(n)} = (\textstyle\sum_{j=1}^n x_j B_{ij}^{(n)})^2,$$

then

$$E(\| A^{(n)}(1)x \|^{-2\varepsilon}) \le n^{-1} \textstyle\sum_{i=1}^n E(W_i^{(n)})^{-\varepsilon} \tag{3.13}$$
$$= n^{-1} \textstyle\sum_{i=1}^n E | \sum_{j=1}^n x_j B_{ij}^{(n)} |^{-2\varepsilon}.$$

PROOF. For any $\varepsilon > 0$, $g(y) = y^{-\varepsilon}$ is convex on $(0, \infty)$ because $g''(y) = (-\varepsilon)(-\varepsilon - 1) y^{-\varepsilon-2} \ge 0$. Therefore, for $w_i \ge 0$,

$$g([w_1 + \cdots + w_n]/n) \le n^{-1} \textstyle\sum_{i=1}^n g(w_i).$$

But since

$$\| A^{(n)} x \|^2 = n^{-1} \textstyle\sum_{i=1}^n W_i^{(n)},$$

the claimed inequality follows.

We need to bound the right side of (3.13). We do so through a series of lemmas.

LEMMA 3.9. *Let $\phi(t) = Ee^{itW}$ be the characteristic function of a random variable W. Then for any $0 < \varepsilon < 1$*

$$E(| W |^{-\varepsilon}) \le K_\varepsilon' \int_{-\infty}^{\infty} | t |^{\varepsilon - 1} | \phi(t) | \, dt$$

where K_ε' is a finite positive constant that depends only on ε.

PROOF. The Fourier transform (in the sense of generalized functions) of $g(w) = | w |^{-\varepsilon}$ is $[F(g)](t) = K_\varepsilon' \, | t |^{\varepsilon-1}$ for some $K_\varepsilon' > 0$ (see, e.g., Gelfand and Shilov, 1958 [1964], page 359, entry 13). Thus, if $\phi(t)$ is a sufficiently smooth and rapidly decreasing function, it follows that

$$E(| W |^{-\varepsilon}) = \int_{-\infty}^{\infty} g(w) f(w) \, dw = \int_{-\infty}^{\infty} [F(g)](t) \phi(t) \, dt$$

where f is the (necessarily smooth and rapidly decreasing) density of W. For a general ϕ, we approximate W by $W_n + Z/m$ where Z is a standard normal variable independent of W and

$$W_n = \begin{cases} W & \text{if} \quad | W | \le n \\ 0 & \text{otherwise.} \end{cases}$$

The characteristic function ϕ_n of W_n is smooth (in fact, analytic) and the characteristic function $\phi_{n,m}$ of $W_n + Z/m$ is $\exp(-t^2/[2m^2])\phi_n(t)$ which is smooth and rapidly decreasing. Thus

$$E\,|\,W_n + Z/m\,|^{-\varepsilon} = K_\varepsilon' \int_{-\infty}^{\infty} |\,t\,|^{\varepsilon-1} e^{-t^2/(2m^2)} \phi_n(t)\,dt$$

$$\leq K_\varepsilon' \int_{-\infty}^{\infty} |\,t\,|^{\varepsilon-1} e^{-t^2/(2m^2)}\,|\,\phi_n(t)\,|\,dt.$$

Now, $\phi_n(t) \to \phi(t)$ pointwise and $|\,\phi_n(t)\,| \leq 1$, $|\,\phi(t)\,| \leq 1$ for all t. By two applications of the dominated convergence theorem, we have

$$\limsup_{m\uparrow\infty} \limsup_{n\uparrow\infty} E\,|\,W_n + Z/m\,|^{-\varepsilon} \leq K_\varepsilon' \int_{-\infty}^{\infty} |\,t\,|^{\varepsilon-1}\,|\,\phi(t)\,|\,dt.$$

Finally, $|\,W_n + Z/m\,|^{-\varepsilon}$ converges in distribution to $|\,W + Z/m\,|^{-\varepsilon}$ as $n \uparrow \infty$ and if $P(W = 0) = 0$, $|\,W + Z/m\,|^{-\varepsilon}$ converges in distribution to $|\,W\,|^{-\varepsilon}$ so that by standard arguments

$$E\,|\,W\,|^{-\varepsilon} \leq \limsup_{m\uparrow\infty} \limsup_{n\uparrow\infty} E\,|\,W_n + Z/m\,|^{-\varepsilon}.$$

This inequality is also true if $P(W = 0) \neq 0$ since in that case both sides are infinite. The proof is completed by combining the last two inequalities.

LEMMA 3.10. *If, for some δ in $(0, 1)$, $\phi(t) = Ee^{itW}$ satisfies*

$$|\,\phi(t)\,| \leq (1 + t^2)^{-\delta/2} \text{ for all } t \text{ in } (-\infty, \infty),$$

then for any ε in $(0, \delta)$,

$$E\,|\,W\,|^{-\varepsilon} \leq K_{\varepsilon,\delta} < \infty$$

where $K_{\varepsilon,\delta}$ is a finite constant that depends only on ε and δ, and may be taken to decrease as δ increases.

PROOF. This lemma follows immediately from Lemma 3.9.

REMARK. If W has finite variance σ^2, then the hypothesis of Lemma 3.10 is implied by the (seemingly weaker) condition,

$$|\,\phi(t)\,| = O(|\,t\,|^{-\delta_1}) \text{ as } |\,t\,| \uparrow \infty \text{ for some } \delta_1 \text{ in } (0, 1).$$

To see this, first note that the finite variance condition implies, with μ denoting EW,

$$|\,\phi(t)\,| = |\,1 + i\mu t - \sigma^2 t^2/2 + o(t^2)\,| = |\exp[i\mu t - \sigma^2 t^2/2 + o(t^2)]\,|$$

$$= \exp[-\sigma^2 t^2/2 + o(t^2)] = (1 + t^2 + o(t^2))^{-\sigma^2/2} \text{ as } |\,t\,| \downarrow 0.$$

On the other hand, the large $|\,t\,|$ condition implies, by Lemma 3.9, that $E\,|\,W + a\,|^{-\varepsilon} < \infty$ for $0 < \varepsilon < \delta_1$ and any a so that W must be a continuous random variable (this also implies $\sigma^2 > 0$). It follows that $|\,\phi(t)\,| < 1$ for $t \neq 0$ and thus that for any $0 < \varepsilon_1 < K < \infty$, $\sup(|\phi(t)|: \varepsilon_1 \le |t| \le K) < 1$. The large $|t|$, small $|t|$, and intermediate $|t|$ bounds together imply the hypothesis of Lemma 3.10 with δ smaller (if necessary) than either σ^2 or δ_1. The finite variance assumption made at the beginning of the remark can actually be eliminated by a simple truncation argument. Finally, if W has a density $p(x)$ which has a continuous derivative $p'(x)$ with $\int_{-\infty}^{\infty} |p'(x)|\, dx < \infty$, then by standard Fourier transform arguments $|t\phi(t)| \downarrow 0$ as $|t| \uparrow \infty$ so that $|\phi(t)| = o(|t|^{-1})$ as $|t| \uparrow \infty$. Thus the hypothesis of Lemma 3.10 is valid.

LEMMA 3.11. *For $\delta > 0$, let*

$$h_\delta(t) = (1 + t^2)^{-\delta/2}.$$

Then for any $|y| \le 1$,

$$h_\delta(yt) \le [h_\delta(t)]^{y^2}.$$

PROOF. Define $D(u) = (\delta/2)\ln(1 + u)$ on $[0, \infty)$. Then $h_\delta(t) = \exp[-D(t^2)]$. We want to show that for $0 \le c \le 1$, $D(cu) \ge cD(u)$ or equivalently that $D(cu)/cu \ge D(u)/u$ or equivalently that $D(u)/u$ decreases on $(0, \infty)$. This is true because $D(0) = 0$ and $D(u)$ is concave on $[0, \infty)$.

DEFINITION. For a random variable W with characteristic function $\phi(t)$, define

$$d(W) = \sup\{\delta \in [0, 1]: |\phi(t)| \le (1 + t^2)^{-\delta/2}\}.$$

Note that $|\phi(t)| \le (1 + t^2)^{-\delta/2}$ with $\delta = d(W)$.

LEMMA 3.12. *If $W_1, \cdots, W_n$ are independent random variables and $\sum_{j=1}^n x_j^2 = 1$, and $d' = \min\{d(W_j): j = 1, \cdots, n\}$, then*

$$d\left(\textstyle\sum_{j=1}^n x_j W_j\right) \ge d'.$$

PROOF. Denoting the characteristic function of W_j by ϕ_j and of $\sum_j x_j W_j$ by ϕ, we have

$$|\phi(t)| = \Pi\, |\phi_j(x_j t)| \le \Pi\, h_{d(W_j)}(x_j t) \le \Pi\, h_{d'}(x_j t) \le [h_{d'}(t)]^{\Sigma x_j^2},$$

where we use Lemma 3.11 to obtain the last inequality.

PROPOSITION 3.13. *If for each n and each i, $B_{i1}^{(n)}, \cdots, B_{in}^{(n)}$ are independent and*

$$d' = \inf_{i,j,n} d(B_{ij}^{(n)}) > 0,$$

then (3.11) *holds for any ε in* $(0, d'/2)$.

PROOF. This is an immediate consequence of Lemmas 3.8, 3.10 and 3.12.

To analyze (3.12), we consider two cases. In the first, we assume every $A_{ij}^{(n)}(1)$

has finite fourth moment but different elements of $A^{(n)}(1)$ need not be independent. In the second, all elements of $A^{(n)}(1)$ must be independent, but they need have only finite second moments.

LEMMA 3.14. *If*

$$\eta'' \equiv \lim\inf_{n\uparrow\infty} E(\| A^{(n)}(1)x^{(n)} \|^2) > 1,$$

and

$$\lim_{n\uparrow\infty} \operatorname{Var}(\| A^{(n)}(1)x^{(n)} \|^2)/E(\| A^{(n)}(1)x^{(n)} \|^2)^2 = 0,$$

then (3.12) *is valid for any* $\eta' < \eta''$.

PROOF. Let $Y_n = \| A^{(n)}(1)x^{(n)} \|^2$. Then for sufficiently large n, $\eta' < EY_n$. So by Tchebychev's inequality,

$$\begin{aligned} P(Y_n < \eta') &= P(1 - Y_n/EY_n > 1 - \eta'/EY_n) \\ &\le \operatorname{Var}(Y_n/EY_n)/[1 - \eta'/EY_n]. \end{aligned}$$

The right side of this inequality tends to zero since its numerator tends to zero while the lim inf of its denominator is $1 - \eta'/\eta'' > 0$.

LEMMA 3.15. *The* η'' *of Lemma* 3.14 *satisfies*

$$\eta'' \ge \lim\inf_{n\uparrow\infty} \beta(C^{(n)})$$

where $C^{(n)} = C$ *is defined as in* (3.5) *and* β *denotes the minimum eigenvalue of* $C^{(n)}$. *If each* $A_{ij}^{(n)}(1)$ *has zero mean, finite variance* $s_{ij}^{(n)2}$ *and if* $\operatorname{Cov}(A_{ki}^{(n)}(1), A_{kj}^{(n)}(1)) = 0$ *for all n, all k, and all* $i \ne j$, *then* $\eta'' > 1$ *if*

$$\lim\inf_{n\uparrow\infty}[\min\{\textstyle\sum_{i=1}^{n} s_{ij}^{(n)2}: j = 1, \cdots, n\}] > 1.$$

PROOF. The first part of the lemma follows from the identities $\beta(C) = \inf_{x\in S^{n-1}}(x, Cx)$ and $(x, Cx) = E(\| A(1)x \|^2)$. The second part follows as in the proof of Corollary 3.4.

Rather than analyze the variance condition of Lemma 3.14 we content ourselves with the following special case.

PROPOSITION 3.16. *Suppose that for each n, i, j, the element* $A_{ij}^{(n)}(1)$ *of* $A^{(n)}(1)$ *has zero mean, finite variance* $s_{ij}^{(n)2}$ *and finite fourth moment* $q_{ij}^{(n)4}$. *Suppose in addition, for each n, either that the* $A_{ij}^{(n)}(1)$*'s are independent for all i, j or else that they are independent only for* $i \le j$ *with* $A_{ji}^{(n)}(1) = A_{ij}^{(n)}(1)$. *Then* (3.12) *is valid for some* $\eta' > 1$ *if the last inequality in Lemma* 3.15 *is valid and*

$$\lim_{n\uparrow\infty} \frac{\sum_{i=1}^{n} [\max\{s_{ij}^{(n)2}: j = 1, \cdots, n\}]^2}{(\sum_{i=1}^{n} [\min\{s_{ij}^{(n)2}: j = 1, \cdots, n\}])^2} = 0$$

and

$$\lim_{n\uparrow\infty} \frac{\sum_{i=1}^{n} \max\{q_{ij}^{(n)4}: j = 1, \cdots, n\}}{(\sum_{i=1}^{n} \min\{s_{ij}^{(n)2}: j = 1, \cdots, n\})^2} = 0.$$

PROOF. We must show that the second limit of Lemma 3.14 is valid. We have with x and A replacing $x^{(n)}$ and $A^{(n)}(1)$,

$$\mathrm{Var}(\| Ax \|^2) = \mathrm{Var}(\textstyle\sum_i (\sum_j A_{ij}x_j)^2) = \sum_i \mathrm{Var}(Y_i) + \sum_{i\neq i'} \mathrm{Cov}(Y_i, Y_{i'}),$$

where

$$Y_i = (\textstyle\sum_j A_{ij}x_j)^2 = \sum_j x_j^2 A_{ij}^2 + 2 \sum_{j<k} x_j x_k A_{ij} A_{ik}.$$

Under the assumptions of the proposition, for each i, the A_{ij}'s are independent with mean zero ($j = 1, \cdots, n$). It follows that

$$\begin{aligned}\mathrm{Var}(Y_i) &= \textstyle\sum_j x_j^4 \mathrm{Var}(A_{ij}^2) + 2 \sum_{j<k} x_j^2 x_k^2 \mathrm{Var}(A_{ij}A_{ik}) \\ &= \textstyle\sum_j x_j^4 (q_{ij}^4 - s_{ij}^4) + 2 \sum_{j<k} x_j^2 x_k^2 s_{ij}^2 s_{ik}^2.\end{aligned}$$

When the A_{ij}'s are independent for all i, j, $\mathrm{Cov}(Y_i, Y_{i'}) = 0$ for $i \neq i'$. When $A_{ij} = A_{ji}$, one has

$$\mathrm{Cov}(Y_i, Y_{i'}) = \mathrm{Cov}(x_{i'}^2 A_{ii'}^2, x_i^2 A_{i'i}^2) = x_i^2 x_{i'}^2 (q_{ii'}^4 - s_{ii'}^4).$$

Thus, in either case

$$\begin{aligned}\mathrm{Var}(\| Ax \|^2) &\le \textstyle\sum_{i,j} x_j^4 (q_{ij}^4 - s_{ij}^4) + \sum_{i,j,k} x_j^2 x_k^2 s_{ij}^2 s_{ik}^2 \\ &\quad + \textstyle\sum_{i,i'} x_i^2 x_{i'}^2 (q_{ii'}^4 - s_{ii'}^4) \\ &\le \textstyle\sum_{i,j} x_j^2 q_{ij}^4 + \sum_{i,i'} x_i^2 q_{ii'}^4 + \sum_{i,j,k} x_j^2 x_k^2 s_{ij}^2 s_{ik}^2 \\ &\le 2 \textstyle\sum_i (\max_j q_{ij}^4) + \sum_i (\max_j s_{ij}^2)^2.\end{aligned}$$

We have used above that $x_j^2 \le \sum x_k^2 \le 1$ for each j. The proposition now follows from Lemma 3.14, Lemma 3.15, and the proof of Lemma 3.15.

COROLLARY 3.17. *If for each n the $A_{ij}^{(n)}(1)$'s are either* i.i.d. *for all i, j or else* i.i.d. *for $i \le j$ with $A_{ij}^{(n)}(1) = A_{ji}^{(n)}(1)$ and in addition have zero mean, finite variance $s^{(n)2}$ and finite fourth moment $q^{(n)4}$, then* (3.12) *is valid for some $\eta' > 1$ if*

$$\liminf n^{1/2} s^{(n)} > 1$$

and

$$\lim_{n\uparrow\infty} n^{-1}[n^{1/2}q^{(n)}/(n^{1/2}s^{(n)})]^4 = 0.$$

PROOF. This follows from a simple computation, using the expressions of Lemma 3.15 and Proposition 3.16.

Combining Theorem 3.3, Propositions 3.7, 3.13, and Corollary 3.17, we have the following.

THEOREM 3.18. *Suppose for each n that the* $A_{ij}^{(n)}$'s *are either* i.i.d. *for all i, j or else are* i.i.d. *for* $i \le j$ *with* $A_{ij}^{(n)}(1) = A_{ji}^{(n)}(1)$. *Let* $B^{(n)}$ *be a random variable distributed as* $n^{1/2}A_{11}^{(n)}(1)$ *and assume* $B^{(n)}$ *has mean zero and finite variance. Then asymptotic strong instability applies if the following three conditions hold:*

$$\liminf_{n\uparrow\infty}\mathrm{Var}(B^{(n)}) > 1$$

$$\lim_{n\uparrow\infty} n^{-1}E[(B^{(n)}/[\mathrm{Var}\ B^{(n)}]^{1/2})^4] = 0.$$

$$\liminf_{n\uparrow\infty} d(B^{(n)}) > 0$$

where d is defined above. In particular, this is the case if the distribution of $B^{(n)}$ *is independent of n with variance* > 1, *finite fourth moment and positive d.*

Proposition 3.16, Corollary 3.17, and Theorem 3.18 all require the existence of finite fourth moments to insure the validity of (3.12). We proceed to eliminate this requirement in the case where, for each n, the $A_{ij}^{(n)}(1)$'s are independent for all i, j. For the sake of simplicity we henceforth assume that for each n, the $A_{ij}^{(n)}(1)$'s are identically distributed as well as independent for all i, j with zero mean and finite variance $s^{(n)2}$ (but possibly infinite fourth moment). Using the notation of Lemma 3.8, we note that in this case the $W_i^{(n)}$'s are, for each n, i.i.d. nonnegative variables with mean $ns^{(n)2}$.

LEMMA 3.19. *In the* i.i.d. *case under consideration,* (3.12) *is valid if* $\liminf_{n\uparrow\infty} ns^{(n)2} > 1$ *and* $X_i^{(n)} \equiv W_i^{(n)}/[ns^{(n)2}]$ *satisfies the uniform integrability condition*

$$\lim_{K\uparrow\infty}\sup_{n\ge 1} E(X_1^{(n)} 1_{\{|X_1^{(n)}|\ge K\}}) = 0. \tag{3.14}$$

PROOF. It suffices to show that as $n \uparrow \infty$,

$$n^{-1}\sum_{i=1}^n X_i^{(n)} \to 1 \text{ (in probability).}$$

The $X_i^{(n)}$'s are, for each n, i.i.d. nonnegative random variables of mean 1 and the $X_1^{(n)}$'s are uniformly integrable. We will apply Theorem 1 of Feller (1971, page 316). We define

$$X_{k,n} = n^{-1}X_k^{(n)}, \quad X'_{k,n} = \begin{cases} n^{-1}X_k^{(n)}, & \text{if } X_k^{(n)} < sn \\ s, & \text{if } X_k^{(n)} \ge sn, \end{cases}$$

$$b_n = nE(X'_{1,n}).$$

Then, since $E(X_1^{(n)}) = 1$,

$$\begin{aligned} 1 - b_n &= E(X_1^{(n)}) - E(X_1^{(n)} 1_{\{X_1^{(n)} < sn\}}) - nsP(X_1^{(n)} \ge sn) \\ &= E(X_1^{(n)} 1_{\{X_1^{(n)} \ge sn\}}) - snP(X_1^{(n)} \ge sn) \end{aligned}$$

so that by uniform integrability, for any $s > 0$,

$$0 \le 1 - b_n \le E(X_1^{(n)} 1_{\{X_1^{(n)} \ge sn\}}) \to 0 \text{ as } n \uparrow \infty.$$

It follows that to obtain the desired result, it suffices to verify conditions (9.1) and (9.2) of Feller. Since (9.1) is weaker than the first condition of (9.2), it suffices to show

$$\text{(3.15)} \qquad \text{for all } \eta > 0, \quad nP(X_1^{(n)} > n\eta) \to 0 \text{ as } n \uparrow \infty$$

and

$$\text{(3.16)} \qquad \text{for all } s > 0, \quad n \operatorname{Var}(X'_{1,n}) \to 0 \text{ as } n \uparrow \infty.$$

The limit (3.15) follows immediately from

$$nP(X_1^{(n)} > n\eta) \le \eta^{-1} E(X_1^{(n)} 1_{\{X_1^{(n)} > n\eta\}}),$$

and uniform integrability. To obtain (3.16), we have

$$\begin{aligned} n \operatorname{Var}(X'_{1,n}) &\le nE(X'^2_{1,n}) = n^{-1}E(X_1^{(n)2} 1_{\{X_1^{(n)} < sn\}}) + ns^2 P(X_1^{(n)} \ge sn) \\ &\le \varepsilon E(X_1^{(n)} 1_{\{X_1^{(n)} < \varepsilon n\}}) + sE(X_1^{(n)} 1_{\{\varepsilon n \le X_1^{(n)} < sn\}}) + sE(X_1^{(n)} 1_{\{X_1^{(n)} \ge sn\}}) \\ &\le \varepsilon E(X_1^{(n)}) + sE(X_1^{(n)} 1_{\{X_1^{(n)} \ge \varepsilon n\}}) \to \varepsilon \quad \text{as } n \uparrow \infty \end{aligned}$$

by uniform integrability. Since this last estimate is true for arbitrarily small ε, we obtain (3.16) which completes the proof.

Now $X_1^{(n)}$ is given by

$$X_1^{(n)} = (\textstyle\sum_{j=1}^n x_j^{(n)} B_{1j}^{*(n)})^2, \quad B_{1j}^{*(n)} = B_{1j}^{(n)}/(\operatorname{Var}(B_{1j}^{(n)}))^{1/2}$$

with the $B_{1j}^{*(n)}$'s, for each n, being i.i.d. and having mean zero, in the case under consideration.

LEMMA 3.20. *Suppose for each n, $Y^{(n)} = \sum_{j=1}^n x_j^{(n)} Y_j^{(n)}$, where the $Y_j^{(n)}$'s are* i.i.d. *with mean zero and variance one, and $\sum_j (x_j^{(n)})^2 = 1$. Then the $(Y^{(n)2})$'s are uniformly integrable if the $(Y_1^{(n)})^2$'s are uniformly integrable.*

PROOF. Define by continuity in some (a priori n-dependent) neighborhood of $r = 0$, the functions

$$D^{(n)}(r) = \log E(e^{irY^{(n)}}), \; D_1^{(n)}(r) = \log E(e^{irY_1^{(n)}}).$$

Then $D^{(n)}(r) = \sum_j D_1^{(n)}(x_j^{(n)} r)$. So

$$[D^{(n)}(r) - r^2/2]/r^2 = \textstyle\sum_{j=1}^n x_j^{(n)2}[D_1^{(n)}(x_j^{(n)} r) - (x_j^{(n)} r)^2/2]/(x_j^{(n)} r)^2.$$

The proof may be completed by applying the next lemma.

LEMMA 3.21. *Let $Y_1, Y_2, \cdots$ be random variables with mean zero, variance one. Then the Y_n^2's are uniformly integrable if and only if for any $\varepsilon > 0$, there exists $r_0 > 0$ such that if $D_n(r) = \log \phi_n(r)$ and $\phi_n(r) = E(\exp[irY_n])$, then*

$$\text{(3.17)} \qquad D_n^*(r) \equiv |[D_n(r) - r^2/2]/r^2| \le \varepsilon \text{ for } |r| \le r_0, \quad \text{for all } n,$$

or equivalently there exists $r_1 > 0$ such that

$$\text{(3.18)} \qquad \phi_n^*(r) \equiv |[\phi_n(r) - (1 - r^2/2)]/r^2| \le \varepsilon \quad \text{for} \quad |r| \le r_1, \quad \text{for all } n.$$

Proof. We first show that (3.17) and (3.18) are equivalent. Denote $\phi_n(r) - 1$ by $\psi_n(r)$. First, by the triangle inequality, $|D_n/r^2| \le \frac{1}{2} + D_n^*$ and $|\psi_n/r^2| \le \frac{1}{2} + \phi_n^*$. Secondly, also by the triangle inequality,

$$|D_n^* - \phi_n^*| \le \begin{cases} |D_n/\psi_n - 1| \cdot |\psi_n/r^2| \le |\log(1+\psi_n)/\psi_n - 1| \cdot (\frac{1}{2} + \phi_n^*). \\ |\psi_n/D_n - 1|\,|D_n/r^2| \le |(e^{D_n} - 1)/D_n - 1| \cdot (\frac{1}{2} + D_n^*). \end{cases}$$

If we define

$$K_\varepsilon(r) = \sup\{|\log(1+\psi)/\psi - 1| : |\psi| \le (\tfrac{1}{2} + \varepsilon)r^2\},$$

$$H_\varepsilon(r) = \sup\{|(e^D - 1)/D - 1| : |D| \le (\tfrac{1}{2} + \varepsilon)r^2\},$$

then

$$|D_n^* - \phi_n^*| \le \begin{cases} K_\varepsilon(r_1)(\frac{1}{2} + \varepsilon), & \text{for } |r| \le r_1 \text{ if (3.18) is valid.} \\ H_\varepsilon(r_0)(\frac{1}{2} + \varepsilon), & \text{for } |r| \le r_0 \text{ if (3.17) is valid.} \end{cases}$$

A further application of the triangle inequality and the fact that $K_\varepsilon(r) \downarrow 0$, $H_\varepsilon(r) \downarrow 0$ as $r \to 0$ imply the equivalence of (3.17) and (3.18). We proceed to prove the equivalence of (3.18) to the uniform integrability of the Y_n^2's.

First we assume uniform integrability. We use the standard inequalities (see e.g. Feller [1971, page 512]) for real u,

$$|e^{iu} - 1 - iu - (iu)^2/2| \le \begin{cases} |u^3|/3! \text{ (for all } u) \le \varepsilon\, u^2/6, \text{ for } |u| < \varepsilon. \\ |e^{iu} - 1 - iu| + |(iu)^2/2| \le u^2, \text{ for all } u. \end{cases}$$

Then

$$\begin{aligned} r^2\phi_n^*(r) &= |E(e^{irY_n} - 1 - irY_n - (irY_n)^2/2)| \\ &\le (\varepsilon r^2/6)E(Y_n^2 1_{\{|Y_n| < \varepsilon/|r|\}}) + r^2 E(Y_n^2 1_{\{|Y_n| \ge \varepsilon/|r|\}}) \qquad (3.19) \\ &\le r^2[(\varepsilon/6)E(Y_n^2) + E(Y_n^2 1_{\{Y_n^2 \ge \varepsilon^2/r^2\}})]. \end{aligned}$$

Given $\varepsilon > 0$, we may choose r_1 sufficiently small so that, by uniform integrability,

$$\sup_n E(Y_n^2 1_{\{Y_n^2 \ge \varepsilon^2/r_1^2\}}) \le 5\varepsilon/6.$$

It then follows from (3.19) and $E(Y_n^2) = 1$ that (3.18) is valid as desired.

Next we assume that (3.18) is valid. We wish to prove uniform integrability of the Y_n^2's. If uniform integrability is not valid, then there is some subsequence Y_{n_k} which converges in distribution, to some Y, such that

$$E(Y^2) < 1 = \lim_{k\uparrow\infty} E(Y_{n_k}^2). \qquad (3.20)$$

We wish to show that (3.20) is impossible. Defining $\phi^*(r)$ as in (3.18) but with ϕ_n replaced by $\phi(r) = E(\exp[irY])$, we have from the convergence in distribution that $\phi_{n_k} \to \phi$ and so $\phi_{n_k}^* \to \phi^*$. Thus ϕ^* satisfies (3.18) or equivalently

$$\phi(r) = 1 - r^2/2 + o(r^2) \text{ as } r \to 0,$$

which implies [see Feller, 1971, pages 512–513] that $EY = -i\phi'(0) = 0$ and $EY^2 = -\phi''(0) = 1$ thus contradicting (3.20) and completing the proof.

Combining Theorem 3.3, Propositions 3.7, 3.13 and Lemmas 3.19, 3.20 gives the following.

THEOREM 3.22. *Suppose for each n that the* $A_{ij}^{(n)}(1)$*'s are* i.i.d. *for all i, j with mean zero and finite variance* $s^{(n)2}$. *Let* $B^{(n)}$ *be a random variable distributed as* $n^{1/2}A_{11}^{(n)}(1)$ *and let* $B^{*(n)}$ *be distributed as* $A_{11}^{(n)}(1)/s^{(n)}$. *Then asymptotic strong instability applies if the following three conditions all hold*:

$$\liminf_{n\uparrow\infty} ns^{(n)2} > 1$$

$$\textit{uniform integrability of the } [(B^{*(n)})^2]\text{'s}$$

$$\liminf_{n\uparrow\infty} d(B^{(n)}) > 0$$

where d is defined above (*following Lemma* 3.11). *In particular, this is the case if the distribution of* $B^{(n)}$ *is independent of n with finite variance* > 1 *and positive d.*

4. Claims, counterexamples and questions. May (1972) considered the stability of a system of linear ordinary differential equations

$$dx/dt = Ax \tag{4.1}$$

where x is an n-vector and the $n \times n$ real matrix A satisfies $A = B - I$. Here I is the $n \times n$ identity matrix, and the meaning of B here differs from that in the preceding section. The off-diagonal ($i \neq j$) elements B_{ij} of the $n \times n$ matrix B are independent random variables that are equal to 0 with probability $1 - C(0 \leq C \leq 1)$ and drawn from an arbitrary distribution, say F, with probability C. Throughout this section, C is a scalar. The diagonal elements of B are 0. The only restriction May imposes on F is that it have mean 0 and variance a^2.

May (1972) defined the system (4.1) to be stable if $R(A)$, the largest real part of the eigenvalues of A, is negative. Let $P(n, a, C)$ be the probability that, for the given values of n, a, and C, the system (4.1) is stable. May asserted, without proof, that

$$\lim_{n\uparrow\infty} P(n, a, C) = 1 \tag{4.2}$$

if

$$\lim_{n\uparrow\infty} (nC)^{1/2} a < 1, \tag{4.3}$$

while

$$\lim_{n\uparrow\infty} P(n, a, C) = 0 \tag{4.4}$$

if

$$\lim_{n\uparrow\infty} (nC)^{1/2} a > 1. \tag{4.5}$$

Recently, Hastings (1982a) considered the system of difference equations

$$x_{t+1} = Bx_t \tag{4.6}$$

where x_t is an n-vector and the random $n \times n$ matrix B has *all* elements B_{ij} i.i.d. B_{ij} is 0 with probability $1 - C$, and with probability C is drawn from a distribution with mean and all odd moments 0 and variance a^2. The system (4.6) is defined to be stable if $r(B)$, the spectral radius of B, is strictly less than 1. Hastings assumed (page 158) that a is fixed and $C = k/n$, where k is a fixed positive constant. Under these assumptions, he claimed that (4.3) (alternatively (4.5)) implies that (4.6) is stable (alternatively not stable) with probability approaching 1 as $n \uparrow \infty$.

With B as in Hastings (1982a) and $A = B - I$, Hastings (1982b) considered the system (4.1) under the additional hypothesis that the (possibly n-dependent) C satisfies, for some $\varepsilon > 0$ and all large n,

$$C \geq (1 + \varepsilon)n^{-1}\log n. \tag{4.7}$$

He claimed that asymptotic stability in the sense of (4.2) is valid if (4.3) holds, and that asymptotic instability in the sense of (4.4) is valid if (4.5) holds. The announced proof, given in more detail in Hastings (1983), apparently rests on claiming that under (4.3), as $n \uparrow \infty$, the norm of A is, with probability approaching 1, less than a quantity that is less than 1, while under (4.5), as $n \uparrow \infty$, the norm of A is, with probability approaching 1, greater than a quantity greater than 1. The error in these claims will be discussed in detail below.

The difference between the matrix A of Hastings (1982b) and the matrix A of May (1972) is that the diagonal elements of Hastings' A are random variables that have the distribution of the off-diagonal elements shifted by -1, while the diagonal elements of May's A are fixed at -1.

We now show, by explicit counterexamples, that the assertions of May (1972, repeated in 1973) and Hastings (1982a, 1982b) are false in the generality with which they are stated. We also describe some more restricted situations where the criteria of May may turn out to be valid.

EXAMPLE 1. Choose any positive constant k and any i.i.d. real random variables X_{ij}, $i, j = 1, 2, \cdots$ such that $P[X_{ij} = 0] = 0$, $EX_{ij} = 0$ and $P(X_{ij} > 1) > 0$. (E.g. let $X_{ij} \sim N(0, s^2)$ with $s^2 > 0$.) Let $C_n = k/n$ with $0 < k < \infty$ and let $B(n)$ be an $n \times n$ matrix with independent elements

$$\begin{aligned} B_{ij}(n) &= 0 \text{ w.p. } 1 - C_n \\ &= X_{ij} \text{ w.p. } C_n. \end{aligned}$$

PROPOSITION 4.1. *If* $p = P(X_{ij} > 1) > 0$, *then*

$$\liminf_{n\uparrow\infty} P(B(n) \textit{ has a real eigenvalue } > 1) \geq 1 - \exp(-pke^{-k}) > 0.$$

PROOF. Let e_i be the n-vector that has 1 in the ith position, 0 elsewhere. Let B_j be the jth column of B (we drop n here). Then $Be_i = B_i$. Thus if $B_i = ae_i$, a is an eigenvalue of B. We have for a given i

$$\begin{aligned} &P(B_i = ae_i \text{ for some } a > 1) \\ &\quad = P(B_{ii} = a \text{ for some } a > 1 \text{ and } B_{ji} = 0 \text{ for all } j \neq i) \\ &\quad = pC_n(1 - C_n)^{n-1}. \end{aligned}$$

Hence

$$P(B_i = ae_i \text{ for some } a > 1 \text{ and some } i)$$
$$= 1 - [1 - (pk/n)(1 - k/n)^{n-1}]^n.$$

Now we use the standard fact that if b_n is a sequence of real numbers with $\lim_{n\uparrow\infty} b_n = b$, then

$$\lim_{n\uparrow\infty}(1 - b_n/n)^n = e^{-b}.$$

Letting $b_n = pk(1 - k/n)^{n-1}$ gives $b = \lim b_n = pke^{-k}$ and

$$P(B_i = ae_i \text{ for some } a > 1 \text{ and some } i)$$
$$\to 1 - \exp(-pke^{-k}) \text{ as } n \uparrow \infty.$$

COROLLARY 4.2. *For any* k, $0 < k < \infty$, *if* $P(X_{ij} > 1) > 0$, *then*

$$P(r(B(n)) < 1) \not\to 1 \text{ as } n \uparrow \infty, \quad P(R(B(n)) < 1) \not\to 1 \text{ as } n \uparrow \infty,$$

even though when $s^2 = \mathrm{Var}(X_{ij}) < 1/k$, *we have*

$$\lim_{n\uparrow\infty} n\mathrm{Var}\, B_{ij}(n) = \lim_{n\uparrow\infty} nC_n s^2 = ks^2 < 1.$$

In particular, X_{ij} *may be chosen to be normally distributed with mean* 0 *and any positive variance* s^2 *such that* $ks^2 < 1$.

This contradicts Hastings (1982a, pages 156–157) since $k = nC_n$ so that $s^2nC_n < 1$. In fact, ks^2 may be chosen arbitrarily close to 0, provided $p > 0$.

EXAMPLE 2. Choose k, X_{ij} and C_n as in Example 1. Let $A(n)$ be an $n \times n$ matrix with independent elements $A_{ij}(n)$ such that $A_{ii}(n) = 0$, $i = 1, \cdots, n$, and for $i \neq j$, $A_{ij}(n)$ is distributed according to

$$A_{ij}(n) = 0 \text{ w.p. } 1 - C_n$$
$$= X_{ij} \text{ w.p. } C_n.$$

PROPOSITION 4.3.

$$\lim\inf_{n\uparrow\infty} P(R(A(n)) > 1) \geq [1 - \exp(-p^2k^2e^{-2k}/2)]/2.$$

Before presenting the proof of Proposition 4.3, we need the following lemma.

LEMMA 4.4. *For* $2 \leq m \leq n$, *with* n *given, let* U_m *be the event that for some* $i \neq j$, $1 \leq i, j \leq m$, *both* $A_i = ae_j$ *and* $A_j = be_i$ *for some* $a > 1$ *and some* $b > 1$. (*The event* U_m *implies that for some* $i, j \leq m$, $e_i + (a/b)^{1/2}e_j$ *is an eigenvector of* A *with eigenvalue* $(ab)^{1/2} > 1$.) *Let* $P_m = P(U_m)$ *and* $\mathrm{int}(x)$ *be the greatest integer* $\leq x$. *Then*

$$P_n \geq \tfrac{1}{2}[1 - (1 - n[(pk/n)(1 - k/n)^{n-2}]^2)^{\mathrm{int}(n/2)}].$$

PROOF. If $2 \leq m \leq m' \leq n$, then $P_m \leq P_{m'}$. Let V_m be the event that for some $j = 2, \cdots, m$, and some $a > 1$, $b > 1$, we have $A_1 = ae_j$ and $A_j = be_1$. Then

$$P_m = P(U_m \mid V_m)P(V_m) + P(U_m \mid \text{not } V_m)P(\text{not } V_m).$$

Now, by independence of matrix elements, with $p = P[X_{ij} > 1] > 0$,

$$\begin{aligned} P(V_m) &= \textstyle\sum_{j=2}^{m} P(A_1 = ae_j \text{ for some } a > 1) \\ &\qquad \cdot P(A_j = be_1 \text{ for some } b > 1 \mid A_1 = ae_j \text{ for some } a > 1) \\ &= (m-1)P(A_1 = ae_2 \text{ for some } a > 1) \\ &\quad \cdot P(A_2 = be_1 \text{ for some } b > 1) \\ &= (m-1)[(pk/n)(1-k/n)^{n-2}]^2. \end{aligned}$$

Also $P(U_m \mid V_m) = 1$ and it can be seen that

$$P(U_m \mid \text{not } V_m) \geq P(U_{m-2}).$$

Thus

$$\begin{aligned} P_m \geq{} & (m-1)[(pk/n)(1-k/n)^{n-2}]^2 \\ & + [1-(m-1)[(pk/n)(1-k/n)^{n-2}]^2]P_{m-2}. \end{aligned}$$

Letting

$$r = 1 - n[(pk/n)(1-k/n)^{n-2}]^2,$$

we have, for $m - 1 \geq n/2$,

$$P_m \geq (n/2)[(pk/n)(1-k/n)^{n-2}]^2 + rP_{m-2}.$$

Therefore, providing $n - 2j \geq 1 + n/2$, i.e. $j \leq \text{int}(n/2) - 1$, we have

$$P_n \geq (n/2)[(pk/n)(1-k/n)^{n-2}]^2(1 + r + r^2 + \cdots r^j).$$

Taking $j = \text{int}(n/2) - 1$ yields the desired inequality.

PROOF OF PROPOSITION 4.3. Let

$$b_n = [pk(1-k/n)^{n-2}]^2.$$

Then

$$\lim_{n\uparrow\infty} b_n = p^2k^2e^{-2k}.$$

Using Lemma 4.4 gives

$$P_n \geq \tfrac{1}{2}[1-(1-b_n/n)^{\text{int}(n/2)}].$$

Therefore

$$\begin{aligned} &\liminf_{n\uparrow\infty} P([ab]^{1/2} \text{ is an eigenvalue of } A(n) \text{ for some } a > 1 \text{ and } b > 1) \\ &\quad \geq \liminf_{n\uparrow\infty} P_n \geq \tfrac{1}{2} \lim_{n\uparrow\infty}[1-(1-b_n/n)^{\text{int}(n/2)}] \\ &\quad = \lim_{n\uparrow\infty}[1-(1-b_n/n)^{n/2}]/2 \\ &\quad = [1-\exp(-p^2k^2e^{-2k}/2)]/2. \end{aligned}$$

COROLLARY 4.5. *If* $0 < k < \infty$, $p = P[X_{ij} > 1] > 0$, *and* $s^2 = \operatorname{Var} X_{ij} < 1/k$, *then* $n \operatorname{Var} A_{ij} = ks^2 < 1$ *but*

$$P(R(A(n)) > 1) \not\to 0 \text{ as } n \uparrow \infty.$$

This contradicts the claim of May (1972) that (4.3) implies (4.2). As in example 1, k may be chosen arbitrarily large and ks^2 arbitrarily small (but positive), and X_{ij} may be normal with mean 0 and variance s^2.

EXAMPLE 3. Choose positive real constants $d > 1$, $0 < C \leq 1$, and $k < d^{-2}$ and a positive real sequence b_n such that $\lim nb_n^2 = 0$. Let $g_n = k/n$. Let $B(n)$ be an $n \times n$ matrix with elements $B_{ij}(n)$ i.i.d. according to

$$\begin{aligned} B_{ij}(n) &= 0 \text{ w.p. } 1 - C \quad \text{(independent of } n) \\ &= \pm b_n \text{ w.p. } C(1 - g_n)/2 \text{ each} \\ &= \pm d \text{ w.p. } Cg_n/2 \text{ each.} \end{aligned}$$

PROPOSITION 4.6. *There exists a positive sequence* b_n (*with* $\lim nb_n^2 = 0$) *approaching zero sufficiently rapidly so that*

$$\liminf_{n\uparrow\infty} P(B(n) \textit{ has an eigenvalue with modulus and real part} > 1) \geq [1 - \exp(-Cke^{-Ck}/2)].$$

PROOF. Consider the $n \times n$ matrix $D(n)$ with elements i.i.d. according to

$$\begin{aligned} D_{ij}(n) &= 0 \text{ w.p. } 1 - Cg_n \\ &= \pm d \text{ w.p. } Cg_n/2 \text{ each.} \end{aligned}$$

$D(n)$ is constructed from $B(n)$ of Example 3 by replacing the elements equal to $\pm b_n$ with 0. So, for a given n, as b_n approaches 0, the eigenvalues of $B(n)$ approach the eigenvalues of $D(n)$, by the continuity of the roots of a polynomial as functions of its coefficients. However, $D(n)$ is a special case of Example 1 with $X_{ij} = \pm d$, each with probability ½ and $C_n = Cg_n$. By Proposition 4.1, $\liminf_{n\uparrow\infty} P(D(n)$ has an eigenvalue $+d) \geq 1 - \exp(-Cke^{-Ck}/2)$. For each n, b_n may be chosen small enough so that $B(n)$ has some eigenvalue within a distance $(d - 1)/2$ in the complex plane of any given eigenvalue of $D(n)$ (w.p.1). In that case $P(B(n)$ has an eigenvalue with modulus and real part $> 1) \geq P(D(n)$ has an eigenvalue $+d)$.

COROLLARY 4.7. *Choosing* b_n *as in Proposition* 4.6, *we have* $\lim n \operatorname{Var} B_{ij}(n) = Ckd^2 < 1$ *but*

$$P(r(B(n)) < 1) \not\to 1 \quad \textit{as} \quad n \uparrow \infty.$$

PROOF. We have only to compute

$$n \operatorname{Var} B_{ij} = nC(1 - g_n)b_n^2 + nCg_nd^2 \to Ckd^2 < 1.$$

This contradicts Hastings' (1982b) assertion that (4.3) implies (4.2) for the system under consideration, even if $Ckd^2 > 0$ is chosen arbitrarily close to 0, provided $d > 1$.

This example could be perturbed to allow not purely discrete nonzero matrix elements, while keeping C constant.

The examples given so far have shown that the stability, with probability approaching 1, of (4.1) or of (4.6) does not follow from the conditions asserted to be sufficient. The next example deals with instability.

EXAMPLE 4. Let $A(n)$ be a sequence of random real matrices with i.i.d. elements such that $EA_{ij} = 0$, $n \operatorname{Var} A_{ij} \to K < 1$, and

$$\lim_{n\uparrow\infty} P(r(A(n)) < 1) = 1.$$

This guarantees that also

$$\lim_{n\uparrow\infty} P(R(A) < 1) = 1.$$

(For example, let $A(n) = 0$ with probability 1, or, by the results of Geman [1980], let $A_{ij}(n)$ be i.i.d. $N(0, s^2/n)$ with $s^2 < 1/2$.) Let g_n be a positive real sequence such that $\lim n^2 g_n = 0$ and let b_n be a positive real sequence such that $\lim ng_n b_n^2 = \infty$ (e.g., $g_n = n^{-3}$ and $b_n = n^{1+a}$, $a > 0$). Let $D(n)$ be a sequence of random real matrices with i.i.d. elements such that

$$\begin{aligned} D_{ij} &= A_{ij} \text{ with probability } 1 - g_n \\ &= +b_n \text{ w.p. } g_n/2 \\ &= -b_n \text{ w.p. } g_n/2. \end{aligned}$$

PROPOSITION 4.8.

$$\lim_{n\uparrow\infty} P(r(D(n)) < 1) = 1.$$

PROOF. For given i, j, $P(D_{ij} \neq A_{ij}) = g_n$. So considering all n^2 elements simultaneously, $P(D(n) \neq A(n)) \leq n^2 g_n$. Therefore, since $\lim n^2 g_n = 0$, $\lim P(D(n) \neq A(n)) = 0$. Hence the assertion.

COROLLARY 4.9. *Even though* $\lim n \operatorname{Var} D_{ij} = +\infty$,

$$\lim_{n\uparrow\infty} P(r(D(n)) < 1) = 1.$$

PROOF. We compute

$$n \operatorname{Var} D_{ij} = n(1 - g_n)\operatorname{Var} A_{ij} + ng_n b_n^2,$$

$$\lim n \operatorname{Var} D_{ij} = K + \lim ng_n b_n^2 = \infty.$$

This concludes the demonstration that the claims of May (1972, 1973) and Hastings (1982a, 1982b, 1983) concerning stability and instability, with probability approaching 1, are false, at least in their full generality.

Where does the "proof" of Hastings (1983) go wrong? Apparently the error consists in assuming mistakenly that

(4.8) $$\lim_{n\uparrow\infty}\sup_{\|v\|=1}P(\|Av\| \geq 1) = 0$$

implies

(4.9) $$\lim_{n\uparrow\infty}P(\|A\| \equiv \sup_{\|v\|=1}\|Av\| \geq 1) = 0,$$

where A is a random $n \times n$ matrix and v is an n-vector.

Let $\|A\| = \sup\{\|Av\| : \|v\| = 1\}$, where the vector norm is the Euclidean norm $\|v\|^2 = (v, v)$ and $(u, v) = \sum_{k=1}^{n} u_k v_k$ for any n-vectors u and v. Suppose the elements $A_{ij}(n)$ of the $n \times n$ matrix $A(n)$ are i.i.d. $N(0, s^2/n)$. As pointed out by Geman (1980, page 253), it follows from results of Wachter (1974; see also Wachter, 1978) that

$$\lim\inf_{n\uparrow\infty}\|A(n)\|^2 \geq 4s^2.$$

Geman shows that in fact

(4.10) $$\lim_{n\uparrow\infty}\|A(n)\| = 2s.$$

Fix s such that $\frac{1}{2} < s < 1$. Then $n \operatorname{Var} A_{ij}(n) = s^2 < 1$. For any v with $\|v\| = 1$, $\|A(n)v\|^2 = \sum_i (\sum_j A_{ij}(n)v_j)^2 = \sum_i (sn^{-1/2}Z_i)^2 = (s^2/n)\chi_n^2$, where, as before, Z_i are i.i.d. $N(0, 1)$ and χ_n^2 is chi-squared with n degrees of freedom. As $n \uparrow \infty$, χ_n^2/n is asymptotically $N(1, 2/n)$. So $\|A(n)v\|^2$ is asymptotically normal with mean $s^2 < 1$ and variance $2s^4/n \downarrow 0$, i.e. (4.8) holds. However, Geman's result (4.10) contradicts (4.9).

Notwithstanding our counterexamples, May's conjectures may be true under appropriate restrictions. For example, let $D(n)$ be an $n \times n$ matrix with i.i.d. elements having distribution function F (*independent of* n) with mean 0 and finite variance s^2. Let $A(n) = n^{-1/2}D(n)$. We say that $A(n)$ varies with n only by scaling. Let $r_n = r(A(n))$. As $n \uparrow \infty$, three possibilities are:

(i) for all F, $r_n \to s^2$.
(ii) for all F, $r_n \to$ a constant that is a functional of F other than s^2.
(iii) for some F, $r_n \to s^2$ or another functional of F, but for other F, r_n does not converge to a constant or does not converge at all.

Case (i) is May's conjecture when $A(n)$ varies only by scaling. Case (ii) asserts that a universal criterion exists, but it is not May's in general. Case (iii) says that May's criteria are valid only for certain distributions under scaling, including perhaps those studied numerically (McMurtrie, 1975).

Acknowledgments. The authors thank H. Furstenberg for several very useful discussions concerning the topics of this paper in general and Theorem 3.1 in particular, E. S. Key for reading previous drafts carefully, correcting numerous errors, and suggesting several improvements (which had already been obtained independently by the authors), Harry Kesten for suggesting improvements, and R. M. May and H. M. Hastings for their comments on an earlier version of the manuscript. They also thank S. Friedland and E. Litt for transferring drafts between the authors in Jerusalem and New York during the preparation of this paper, and J. Hernandez for typing the manuscript through many drafts.

J. E. C. thanks the National Science Foundation (under grant DEB 80-11026), the John D. and Catherine T. MacArthur Foundation Prize Fellows Program, and Mr. and Mrs. William T. Golden of Olive Bridge, New York, for their support and hospitality during this work.

C. M. N. thanks the University of Arizona, the National Science Foundation (under grant MCS 80-19384), the Courant Institute of Mathematical Sciences of New York University, the Lady Davis Fellowship Trust, and the Institute of Mathematics and Computer Science of the Hebrew University of Jerusalem for their support and hospitality during his sabbatical.

REFERENCES

ABRAMOWITZ, M. and STEGUN, I. A. (1964). *Handbook of Mathematical Functions*, 9th printing. Government Printing Office, Washington, D.C. Dover, New York, reprint, 1970.

FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications,* vol. 2, 2nd ed. Wiley, New York.

FURSTENBERG, H. and KESTEN, H. (1960). Products of random matrices. *Ann. Math. Statist.* **31** 457–469.

FURSTENBERG, H. (1963). Noncommuting random products. *Trans. Amer. Math. Soc.* **108** 377–428.

FURSTENBERG, H. and KIFER, Y. (1983). Random matrix products and measures on projective spaces. Manuscript.

GELFAND, I. M. and SHILOV, G. E. (1958). *Generalized Functions*, vol. 1. Fizmatgiz, Moscow. English translation, Academic, New York, 1964.

GEMAN, S. (1980). A limit theorem for the norm of random matrices. *Ann. Probab.* **8** 252–261.

GIRKO, V. L. (1976). A limit theorem for products of random matrices. *Theor. Probab. Appl.* **21** 197–199.

HASTINGS, H. (1982a). The May-Wigner stability theorem. *J. Theor. Biol.* **97** 155–166.

HASTINGS, H. (1982b). The May-Wigner stability theorem for connected matrices. *Bull. Amer. Math. Soc.* **7** 387–388.

HASTINGS, H. (1983). Stability of community interaction matrices. *Population Biology*, ed. Freedman, H. I. and Strobeck, C. *Lect. Notes in Biomath.* **52** 355–358. Springer, New York.

MAY, R. M. (1972). Will a large complex system be stable? *Nature* **238** 413–414.

MAY, R. M. (1973). *Stability and Complexity in Model Ecosystems.* Princeton University Press, New Jersey.

MCMURTRIE, R. E. (1975). Determinants of stability of large randomly connected systems. *J. Theor. Biol.* **50** 1–11.

WACHTER, K. (1974). Foundations of the asymptotic theory of random matrix spectra. Ph.D. dissertation, Trinity College, Cambridge and St. Catherine's College, Oxford.

WACHTER, K. (1978). The strong limits of random spectra for sample matrices of independent elements. *Ann. Probab.* **6** 1–18.

LABORATORY OF POPULATIONS
ROCKEFELLER UNIVERSITY
1230 YORK AVE., BOX 20
NEW YORK, NEW YORK 10021

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ARIZONA
TUCSON, ARIZONA 85721