

The Distribution of Lyapunov Exponents: Exact Results for Random Matrices

Charles M. Newman*

Department of Mathematics, University of Arizona, Tucson, AZ 85721, USA

Abstract. Simple exact expressions are derived for all the Lyapunov exponents of certain N -dimensional stochastic linear dynamical systems. In the case of the product of independent random matrices, each of which has independent Gaussian entries with mean zero and variance $1/N$, the exponents have an exponential distribution as $N \rightarrow \infty$. In the case of the time-ordered product integral of $\exp[N^{-1/2}dW]$, where the entries of the $N \times N$ matrix $W(t)$ are independent standard Wiener processes, the exponents are equally spaced for fixed N and thus have a uniform distribution as $N \rightarrow \infty$.

Lyapunov exponents arise in the study of dynamical systems, including both deterministic nonlinear and stochastic linear models. Positive exponents indicate sensitive dependence on initial conditions due to instabilities or the occurrence of strange attractors. In the nonlinear context, positive exponents have been observed experimentally in the onset of fluid turbulence [1]; in the linear context, they occur in the proof of localization for one-dimensional disordered systems [2]. The large-dimensional limiting behavior of the exponents is of interest in both contexts: e.g., in the analysis of fully developed turbulence [3] and of localization in more than one dimension (as approximated by one-dimensional strips) [4]. Motivated by Ruelle's proposal [3] that many physical systems should exhibit a limiting density of exponents, we demonstrate explicitly that that is indeed the case for certain highly symmetric linear systems. The work reported here is a direct outgrowth of earlier work on the leading Lyapunov exponent [5].

The linear systems we treat are stochastically time dependent, which distinguishes our results from the classic work of Wigner [6] on the limiting density of states for Hamiltonians which are random but fixed in time. This will lead to the replacement, in one case below, of Wigner's semi-circle law by a triangle law. The methods we use are not immediately applicable either to nonlinear systems or to

* John S. Guggenheim Memorial Fellow. Research supported in part by NSF Grant MCS 80-19384

linear systems with insufficient symmetry. Consequently, one might expect other limiting densities to occur in fluid turbulence and in localization (and in the transfer matrix approach to spin glasses and other disordered Ising models). On the other hand, the universality associated with the high-dimensional limit [see, e.g., the discussion following Eq.(10) below] may make our results directly relevant [7].

The N -dimensional linear systems we consider evolve either in discrete time, according to $x(t) = A(t)x(t-1)$, or in continuous time, according to $dx/dt = (d\tilde{W}/dt)x$, where $A(t)$ and $\tilde{W}(t)$ are matrix valued. In both cases $x(t) = B(t)x(0)$ with $B(t)$ in the discrete case equal to $A(t) \dots A(1)$ and in the continuous case equal to the time ordered product integral of $1 + d\tilde{W}(t)$ or equivalently the time ordered exponential of $\int_0^t dW(t')$ with W related to \tilde{W} in a standard way (see the detailed definition below of Model d). For the stochastic processes $A(t)$ and $W(t)$ we will consider, the Lyapunov exponents are determined by any of three equivalent limits as $t \rightarrow \infty$ [8]. First, they are the set of numbers obtained as $\lim(\log \|x(t)\|)/t$, for arbitrary nonzero $x(0)$, where $\|x\|$ denotes the length of x . Second, they are the N numbers (counting multiplicity), $\mu_1 \geq \mu_2 \geq \dots \geq \mu_N$, obtained as

$$\mu_k = \lim(2t)^{-1} \log(k^{\text{th}} \text{ largest eigenvalue of } B(t)^T B(t)). \quad (1)$$

Third, $\mu_1 + \dots + \mu_k$ is the maximum over all choices of (linearly independent) $x_1(0), \dots, x_k(0)$, of

$$\lim(\log \|x_1(t) \wedge \dots \wedge x_k(t)\|)/t, \quad (2)$$

where $\|x_1 \wedge \dots \wedge x_k\|$ denotes the k -dimensional volume of the parallelogram spanned by x_1, \dots, x_k . Our calculations will be based on (2) which we rewrite in the discrete time case as

$$\lim_{t \rightarrow \infty} t^{-1} \sum_{j=1}^t \log \left(\frac{\|A(j)x_1(j-1) \wedge \dots \wedge A(j)x_k(j-1)\|}{\|x_1(j-1) \wedge \dots \wedge x_k(j-1)\|} \right); \quad (3)$$

in the continuous time case, (3) remains valid with $A(j) = B(j) [B(j-1)]^{-1}$.

We make two important assumptions. The first is that the random $A(j)$'s are statistically independent and identically distributed (with zero probability of being singular). Denote by A a random matrix with the same distribution as any $A(j)$. The second assumption is that of sufficient symmetry so that for any (nonrandom) orthogonal matrix Q , the random matrix $Q^T A^T A Q$ has the same distribution as $A^T A$. Since $\|Ax_1 \wedge \dots \wedge Ax_k\| / \|x_1 \wedge \dots \wedge x_k\|$ is 1) dependent only on $A^T A$ rather than on A itself and 2) is unchanged upon replacement of x_1, \dots, x_k by an orthonormal basis for the subspace spanned by x_1, \dots, x_k , it follows easily from our second assumption that the *distribution of $\|Ax_1 \wedge \dots \wedge Ax_k\| / \|x_1 \wedge \dots \wedge x_k\|$ does not depend on the choice of x_1, \dots, x_k* . Combining this with the first assumption, we conclude that the summands of (3) are statistically independent (and identically distributed) so that by the standard law of large numbers, (3) may be evaluated to yield

$$\mu_1 + \dots + \mu_k = \langle \log \|Ae_1 \wedge \dots \wedge Ae_k\| \rangle, \quad (4)$$

where $\langle \cdot \rangle$ denotes expectation with respect to the distribution of A and e_1, e_2, \dots denote the standard orthonormal basis vectors (i.e., Ae_j is the j^{th} column of A).

In all of the systems we will consider, our second assumption will be a consequence of a stronger symmetry property, namely the existence for any Q of some Q' so that

$$Q'AQ \text{ is equidistributed with } A. \quad (5)$$

In the first class of systems, which we call Model a , (5) is easily seen to be valid with $Q' = 1$.

Model a. The N columns of A are independent random vectors, each with mean zero and a common N -dimensional Gaussian distribution.

Although (4) is valid for Model a , it does not seem possible to obtain a closed form expression for the μ_k 's without further restrictions. We therefore consider a special case of Model a which we call Model b .

Model b. The N^2 entries of A are independent Gaussian variables with mean zero and variance σ^2 .

To evaluate μ_k in Model b , we note that

$$\|Ae_1 \wedge \dots \wedge Ae_k\| = \|Ae_1 \wedge \dots \wedge Ae_{k-1}\| \cdot \|P_k Ae_k\|,$$

where P_k is the (random) projection orthogonal to the subspace spanned by Ae_1, \dots, Ae_{k-1} . Now in Model b , Ae_k is a random vector which has the same distribution as QAe_k for any orthogonal Q which is either nonrandom or random but independent of Ae_k . It follows that if P is the projection onto a j -dimensional subspace which is either nonrandom or random but independent of Ae_k (and with j nonrandom), then $\|PAe_k\|$ has the same distribution as $\|\bar{P}_j Ae_k\|$, where \bar{P}_j is the projection onto the span of e_1, \dots, e_j . In Model 1, Ae_1, \dots, Ae_{k-1} are independent of Ae_k and thus

$$\mu_k = \langle \log \|P_k Ae_k\| \rangle = \frac{1}{2} \langle \log [\|\bar{P}_{N-k+1} Ae_k\|^2] \rangle = \frac{1}{2} \left\langle \log \left[\sum_{i=1}^{N-k+1} (A_{ik})^2 \right] \right\rangle. \quad (6)$$

The Gaussian integral implicit in the rightmost expression of (6) may be evaluated to yield

$$\mu_k = \log \sigma + [\log 2 + \Psi((N-k+1)/2)]/2, \quad (7)$$

where Ψ is the digamma function, $\Psi(u) = \Gamma'(u)/\Gamma(u)$ and Γ is the standard gamma function. The formulas (6) and (7) for the special case, $k=1$, were known previously [5].

To obtain the asymptotic distribution of the exponents in Model b , we let $\sigma = s/\sqrt{N}$, choose $k = k_N$ such that $k_N/N \rightarrow (1-v)$ in $(0, 1)$, and use either (7) and the known asymptotics of Ψ or (6) and the law of large numbers to conclude that $\mu_{k_N} \rightarrow \log s + (\log v)/2$. This means that asymptotically a fraction v of the μ_k 's have values below $\log s + (\log v)/2$ with $(1-v)$ having larger values. It follows that the asymptotic density g of Lyapunov exponents for Model b as $N \rightarrow \infty$ is an exponential,

$$g(u) = \begin{cases} 2 \exp[2(u - \log s)], & \text{for } u < \log s \\ 0, & \text{elsewhere.} \end{cases} \quad (8)$$

It is instructive to consider the exponentials of the μ_k 's since these are analogous to the eigenvalues of a single (symmetric) random matrix, $(A^T A)^{1/2}$, as studied by

Wigner [9]. For the A of Model b with $\sigma = s/\sqrt{N}$, Wigner’s formula for the asymptotic density f of eigenvalues is actually a “quarter-circle” law [9, 10]:

$$f(\lambda) = \begin{cases} (2/\pi s)(1 - (\lambda/2s)^2)^{1/2}, & \text{for } 0 < \lambda < 2s \\ 0, & \text{elsewhere.} \end{cases} \tag{9}$$

The asymptotic density h for the $\exp[\mu_k]$ ’s is immediately obtained from (8), by letting $\lambda = e^\mu$:

$$h(\lambda) = \begin{cases} 2\lambda/s^2, & \text{for } 0 < \lambda < s \\ 0, & \text{elsewhere.} \end{cases} \tag{10}$$

Thus the quarter-circle in $(\lambda/2s)$ is replaced by a triangle in (λ/s) .

What is the status of universality for discrete time models? Wigner’s limiting density is universal, at least to the extent that the matrix A may be taken to have independent entries with a common non-Gaussian distribution of mean zero, scaled by s/\sqrt{N} . It is also known that it is not completely universal, since the scaling of A_{ij} may be chosen to depend on i and N in such a way as to obtain limiting densities different than (9) [10]. The situation for the Lyapunov exponents appears to be precisely analogous. In particular, we note that an i -dependent scaling of the A_{ij} ’s arises as a special case of Model a . It can be shown [11] that there is a rather simple nonlinear mapping between the limiting density for eigenvalues and that for exponents; the relation between (9) and (10) is a special case of that mapping. The mapping is most easily defined in terms of the limiting cumulative distribution function $H(\lambda)$ for the $\exp[\mu_k]$ ’s (the density h , as in (10), is the derivative of H):

$$H(\lambda) = \begin{cases} 0, & \text{for } \lambda \leq [\int t^{-2} f(t) dt]^{-1/2} \\ 1, & \text{for } \lambda \geq [\int t^2 f(t) dt]^{1/2} \\ \text{the unique solution in } (0, 1) \text{ of the equation} \\ \int t^2 [\lambda^2 H + t^2(1 - H)]^{-1} f(t) dt = 1, & \text{for all other } \lambda. \end{cases} \tag{11}$$

For an explicit list of the (probably overly restrictive) hypotheses under which (11) has been proven valid, see Theorem 2.11 of [11].

Preparatory to our discussion of continuous time models, we consider another class of discrete time systems we call Model c . The definition of this model involves an analytic function Φ and a parameter γ . For ease of digestion, the reader may wish to restrict attention to the special case, where $\gamma = 0$ and $\Phi(C) = \exp(\varepsilon C)$ with $\varepsilon > 0$.

Model c. $A = \Phi(C)$ where Φ is an everywhere-convergent power series with $\Phi(0) = 1$ and C is a random matrix whose entries are jointly Gaussian with mean zero and covariance, $\langle C_{ij} C_{i'j'} \rangle = \sigma^2 [\delta_{ii'} \delta_{jj'} + \gamma \delta_{ij'} \delta_{ji'}]$. The parameter γ is necessarily in the interval $[-1, 1]$. We note that (5) is valid with $Q' = Q^{-1} = Q^T$, since

$$\langle \exp[\text{tr}(R^T C)] \rangle = \exp[\text{tr}(\sigma^2 R^T R / 2 + \gamma \sigma^2 R^2 / 2)],$$

which is unchanged when R^T is replaced by $QR^T Q^{-1}$. When $\gamma = 0$, C is just the A of Model b ; when $\gamma = 1$, $C_{ij} = C_{ji}$ for all i and j and when $\gamma = -1$, $C_{ij} = -C_{ji}$ for all i and j .

In order to obtain a system with explicitly calculable Lyapunov exponents, we take a continuous time limit of Model c which is a matrix valued diffusion process we call Model d . This model is related to the theory of stochastic flows [12].

Model d. A is distributed as $B(1)$ from a continuous time model in which $W(t)$ is a matrix valued Wiener process with $W(1)$ distributed as the C of Model c ; i.e., $W' = dW/dt$ is a matrix valued Gaussian white noise with mean zero and covariance

$$\langle W'_{ij}(t_1)W'_{i'j'}(t_2) \rangle = \sigma^2[\delta_{ii'}\delta_{jj'} + \gamma\delta_{ij'}\delta_{ji'}]\delta(t_2 - t_1).$$

$\tilde{W}(t)$ is formally related to $W(t)$ by $d\tilde{W}(t) = dW(t) + [dW(t)]^2/2$ which in this case yields $\tilde{W}(t) = W(t) + \sigma^2(1 + N\gamma)t/2$. The relation between Models c and d is that

$$B(1) = \lim_{m \rightarrow \infty} \exp[C_m(m)] \dots \exp[C_m(1)], \tag{12}$$

where $C_m(j) = W(j/m) - W((j-1)/m)$ is distributed as C/\sqrt{m} .

We now evaluate the μ_k 's for Model d [13]. These have been obtained independently by LeJan [14], Baxendale and Harris [15], and Norris et al. [16], using Ito calculus methods directly suited to continuous time processes. Using Eq. (12) and the fact that Model c satisfies our two hypotheses, we may reexpress (4) for Model d as

$$\mu_1 + \dots + \mu_k = \lim_{m \rightarrow \infty} \frac{m}{2} \langle \log [\| e^{C/\sqrt{m}} e_1 \wedge \dots \wedge e^{C/\sqrt{m}} e_k \|^2] \rangle. \tag{13}$$

Using the fact that $\|x_1 \wedge \dots \wedge x_k\|^2$ is the determinant of the $k \times k$ matrix whose entries are the inner products, $x_i \cdot x_j$, we have

$$\mu_1 + \dots + \mu_k = \lim_{m \rightarrow \infty} \frac{m}{2} \langle \log \det(\bar{P}_k e^{C^T/\sqrt{m}} e^{C/\sqrt{m}} \bar{P}_k) \rangle, \tag{14}$$

where \bar{P}_k is again the projection onto the span of e_1, \dots, e_k . Standard manipulations then yield

$$\begin{aligned} \mu_1 + \dots + \mu_k &= \lim_{m \rightarrow \infty} \frac{m}{2} \langle \text{tr} \{ \bar{P}_k [(C^T + C)/\sqrt{m} + (C^T + C)^2/2m \\ &\quad - (C^T + C)\bar{P}_k(C^T + C)/2m] \bar{P}_k \} + o(1/m) \rangle \\ &= \frac{1}{4} \left\langle \sum_{i=1}^k \sum_{j=k+1}^N (C_{ji} + C_{ij})(C_{ij} + C_{ji}) \right\rangle = \frac{1+\gamma}{2} \sigma^2 \sum_{i=1}^k (N-k), \end{aligned}$$

where we have used in the first equality that $\text{tr}(C^T C + C C^T) = 2\text{tr}(C^T C)$. These expressions for $\mu_1 + \dots + \mu_k$ imply that for each k ,

$$\mu_k = (1 + \gamma)(N - 2k + 1)\sigma^2/2. \tag{15}$$

All Lyapunov exponents vanish when $\gamma = -1$ because $W(t)$ is then skew-symmetric and $B(t)$ is a diffusion in $\text{SO}(N)$ so that $\|x(t)\| = \|x(0)\|$ for all t . For $\gamma \neq -1$, $B(t)$ is a diffusion in the general linear group $\text{GL}(N)$. A diffusion in $\text{SL}(N)$ can be obtained by replacing $W(t)$ by $W(t) - N^{-1}\text{tr}(W(t))$; this leaves (15) unchanged.

The asymptotic behavior of the μ_k 's for Model d is apparent from (15) since the μ_k 's are N evenly spaced points between $-(1 + \gamma)(N - 1)\sigma^2/2$ and $(1 + \gamma)(N - 1)\sigma^2/2$. Choosing $\sigma = s/\sqrt{N}$ for fixed s , the asymptotic density of Lyapunov exponents for Model d as $N \rightarrow \infty$ is uniform between $-(1 + \gamma)s^2/2$ and $(1 + \gamma)s^2/2$.

Acknowledgements. We thank R. Durrett, D. Pickrell, and B. Souillard for useful discussions; P. Baxendale, E. Dynkin, T. Harris, Y. LeJan, H. McKean, Jr., and M. Pinsky for useful correspondence concerning Model d ; D. Ruelle for bringing to our attention the work of Livi, Politi, and Ruffo discussed in [7]; and O. Lanford for his comments on an earlier version of this work.

References

1. Brandstätter, A., Swift, J., Swinney, H.L., Wolf, A.: Low-dimensional chaos in a hydrodynamic system. *Phys. Rev. Lett.* **51**, 1442 (1983)
2. See e.g., Pastur, L.A.: Spectral properties of disordered systems in one-body approximations. *Commun. Math. Phys.* **75**, 179 (1980)
Kunz, H., Souillard, B.: Sur le spectre des opérateurs aux différences finies aléatoires. *Commun. Math. Phys.* **78**, 201 (1980)
3. Ruelle, D.: Large volume limit of the distribution of characteristic exponents in turbulence. *Commun. Math. Phys.* **87**, 287 (1982)
4. Lacroix, J.: Localisation pour l'opérateur de Schrödinger aléatoire dans un ruban. *Ann. Inst. H. Poincaré* **40A**, 97 (1984)
5. Cohen, J.E., Newman, C.M.: *Ann. Probab.* **12**, 283 (1984)
6. Mehta, M.L.: *Random matrices and the statistical theory of energy levels.* New York: Academic Press 1967
7. It was pointed out to us by D. Ruelle that in Livi, R., Politi, A., Ruffo, S.: Distribution of characteristic exponents in the thermodynamic limit, submitted to *J. Phys. A*, there are numerical results concerning the high-dimensional limit for a certain nonlinear deterministic dynamical system (the "Fermi-Pasta-Ulam β -model" with the energy per particle fixed). They suggest, at least for large energy per particle, an asymptotically uniform distribution of Lyapunov exponents. On the other hand, a uniform distribution of exponents is valid exactly for the continuous time linear stochastic dynamical systems presented as "Model d " in this paper [see Eq. (15)]. An asymptotically uniform distribution has also been obtained (numerically) for products of certain random symplectic matrices (G. Paladin and A. Vulpiani, 1985 preprint). The exact nature of the universality implicit in these "coincidences" deserves further investigation
8. Oseledec, V.I.: *Tr. Mosk. O.-Va* **19**, 179 (1968) [*Trans. Mosc. Math. Soc.* **19**, 197 (1968)]
Raghuathan, M.S.: *Isr. J. Math.* **32**, 356 (1979)
9. See p. 7 of Wigner, E.P.: *SIAM Rev.* **9**, 1 (1967)
10. Wachter, K.: *Ann. Probab.* **6**, 1 (1978)
11. Newman, C.M.: Lyapunov exponents for some products of random matrices: exact expressions and asymptotic distributions. In: *Random matrices and their applications.* Cohen, J.E., Kesten, H., Newman, C.M. (eds.). Providence, RI: AMS (to appear)
12. Harris, T.E.: *Ann. Probab.* **9**, 232 (1981)
13. These can also be obtained by first showing that $(B(t)^T B(t))^{1/2}$ is a Brownian motion on the space of positive definite matrices and then applying results of Dynkin, E.B.: *Dokl. Akad. Nauk SSSR* **141**, 288 (1961), *Izv. Akad. Nauk SSSR Ser. Mat.* **30**, 455 (1966), which are based on the theory of Martin boundaries. Our direct derivation is more elementary
14. LeJan, Y.: *C.R. Acad. Sci. Paris, Sér. I* **298**, 361 (1984)
LeJan, Y.: On isotropic Brownian motions, *Z. Wahr.* (to appear)
15. Baxendale, P., Harris, T.E.: Isotropic stochastic flows. *Ann. Probab.* (to appear)
16. Norris, J.R., Rogers, L.C.G., Williams, D.: Brownian motions of ellipsoids. *Trans. A.M.S.* (to appear)

Communicated by O. E. Lanford

Received August 20, 1984; in revised form July 17, 1985