Z. Wahrscheinlichkeitstheorie verw. Gebiete
60, 291–314 (1982)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

A Chaos Hypothesis for Some Large Systems of Random Equations*

Stuart Geman[1] and Chii-Ruey Hwang[2]

[1] Division of Applied Mathematics, Brown University, Providence, Rhode Island 02912, USA
[2] Institute of Mathematics, Academia Sinica, Taipei, Taiwan, R.O.C.

Summary. This paper is about the behavior of solutions to large systems of linear algebraic and differential equations when the coefficients are random variables. We will prove a law of large numbers and a central limit theorem for the solutions of certain algebraic systems, and the weak convergence to a Gaussian process for the solution of a system of differential equations. Some of the results were surprisingly difficult to prove, but they are all easily anticipated from a "chaos hypothesis": i.e. an assumption of near independence for the components of the solutions of large systems of weakly coupled equations.

1. Introduction

In a large and homogeneous system of interacting particles, it is natural to think of the states of individual particles as being "nearly independent". For example, in his derivation of the Maxwell velocity distribution for an ideal gas, Boltzmann assumed that velocities are pairwise independent (Boltzmann's so-called chaos hypothesis). Although he was unable to rigorously justify the assumption, its consequence, the Maxwell distribution, has been experimentally verified.

We will demonstrate that a similar assumption of statistical chaos is appropriate in other systems as well: a "chaos hypothesis" leads easily to true conjectures, although many of these are quite difficult to prove. We will examine systems of random equations in which intuition suggests the near pairwise independence of all of the random variables involved. In each case, the *assumption* of independence leads, through a law of large numbers (LLN) or the central limit theorem (CLT), to a conjecture about the distribution of the solutions, and in each case the validity of the conjecture will be established. Although the discussion here is of linear systems (algebraic and differential), we have examples in nonlinear systems as well (see [6]).

* Supported by the National Science Foundation under grant MCS76-80762, by the U.S. Air Force under grant AFOSR 78-3514 and the U.S. Army under grant DAAG 2980-K-0006

0044-3719/82/0060/0291/$04.80

We will begin in Sect. 2 with a discussion of a linear system of algebraic equations with random coefficients. The assumption of independence will lead us to conjecture a law of large numbers for the solution, as the number of equations grows. This conjecture is followed by a formal proof. In Sect. 3 we will study the same system under a different normalization. Here, the natural conjecture is a central limit theorem for the solution, and, here again the conjecture can be shown to be true. In Sect. 4, we will look at systems of differential equations with random constant coefficients. A chaos hypothesis leads us to predict that the solutions converge, as the number of equations increases, to Gaussian processes with specified mean and covariance functions. We will prove that this is indeed the limiting behavior.

This research was motivated in part by our belief that the systems studied here are of some interest in their own right. What, if any, regular behavior can we expect in a large linear system of algebraic or differential equations in which the coefficients are random variables? To us, a natural question, and one whose answer will possibly be of some use in applications. Mostly, though, we had hoped to gain techniques which would apply to more sophisticated versions of the same problem: relating a "local chaos" to a "global regularity". But we have not succeeded in finding a general approach; our proofs are specialized to their particular systems. Nevertheless, we are convinced that there is a common method of proof, more natural and more general. This would be worth finding, for it is certainly true that chaos is operating in a similar fashion in a broad class of systems.

We would like to remark finally on a technique, not yet very common, used in developing our theorems: namely, we have repeatedly used a computer to convince ourselves of the validity of a conjecture before or during our attempts to prove it. We found some of these proofs to be quite challenging. It is unlikely that we would have been able to complete these without first knowing that the conjectures were (in all likelihood) true. This knowledge was obtained through simulation.

2. LLN for an Algebraic System of Equations

Let $\{w_{ij}\}$, $i=1,2,\ldots;\ j=1,2,\ldots$ be independent and identically distributed random variables with zero means. For each $n=1,2,\ldots$ define W_n to be the $n\times n$ matrix whose (i,j) component is w_{ij}. Given a sequence of numbers $\alpha_1,\alpha_2,\ldots$ define, for each n, $v_n=(\alpha_1,\ldots,\alpha_n)^T$. Finally, for each n define a random vector $x_n=(x_{n,1},\ldots,x_{n,n})^T$ as the solution to the equation

$$x_n=v_n+\frac{1}{n}W_n x_n \tag{2.1}$$

i.e.

$$x_{n,i}=\alpha_i+\frac{1}{n}\sum_{j=1}^{n} w_{ij}x_{n,j} \qquad 1\leqq i\leqq n.$$

For a given realization of the matrix W_n, the components of x_n are deterministically related. Still, we would expect that when n is large each component of x_n would be "nearly independent" of each of the remaining components, as well as of each element in the matrix W_n. For n large, then, we might expect that

$$\frac{1}{n}\sum_{j=1}^{n} w_{ij} x_{n,j} \approx E w_{i1} x_{n,1} \approx E[w_{i1}]\, E[x_{n,1}] = 0.$$

In other words, it is natural to conjecture that for each i, $x_{n,i} \to \alpha_i$ in some sense. The logic is of course very loose, and in fact a straightforward argument shows that $E[x_{n,1}]$ does not generally exist. Nonetheless:

Theorem 1. *Define x_n by (2.1) whenever $I - \frac{1}{n} W_n$ is nonsingular. Otherwise, define $x_n = 0$.*

i) *If $E|w_{11}|^8 < \infty$ and $(\alpha_1, \alpha_2, \ldots) \in l_\infty$, then for each $i = 1, 2, \ldots, x_{n,i} \to \alpha_i$ a.s.*

ii) *If $E|w_{11}| < 1$, $E[e^{i\lambda w_{11}}]$ is analytical at $\lambda = 0$, and $(\alpha_1, \alpha_2, \ldots) \in l_\infty$, then $\sup_{1 \leq i \leq n} |x_{n,i} - \alpha_i| \to 0$ a.s.*

iii) *If $E|w_{11}|^8 < \infty$ and $(\alpha_1, \alpha_2, \ldots) \in l_2$, then $(x_{n,1}, \ldots, x_{n,n}, 0, 0, \ldots) \to (\alpha_1, \alpha_2, \ldots)$ in l_2 a.s.*

If $\alpha_1 = \alpha_2 = \ldots = \alpha$, and if $E w_{11} = m$ (instead of 0) then the same heuristics lead to the conjecture $x_{n,i} \to \frac{\alpha}{1-m}$ for each i:

Theorem 2. *Define x_n by (2.1) whenever $I - \frac{1}{n} W_n$ is nonsingular and define $x_n = 0$ otherwise. If $|m| < 1$ and if there exists a β such that $E|w_{11}|^n \leq n^{\beta n}\ \forall n \geq 2$, then for each $i = 1, 2, \ldots, x_{n,i} \to \frac{\alpha}{1-m}$ a.s.*

Certainly other (and probably better) results along these lines are possible. But the point is this: in each case an assumption of "statistical chaos" leads us to conjecture exactly the right asymptotic (i.e. large system) behavior.

Proof of Theorem 1. The proofs of i and iii are based on the following three lemmas. The first we cite without proof (cf. Chung [4], Chap. 5).

Lemma 2.1. *If $x_1, x_2, \ldots$ are uncorrelated random variables, and their second moments have a common bound, then*

$$\frac{1}{m}\left\{\sum_{i=1}^{m} x_i - E\sum_{i=1}^{m} x_i\right\} \to 0 \quad \text{a.s.}$$

Lemma 2.2. *If $E|w_{11}|^8 < \infty$, then $\left\|\left(\frac{W_n}{n}\right)^2\right\| \to 0$ a.s.*[1]

[1] When applied to a matrix, $\| \ \|$ denotes induced operator norm. (Recall that $\| \ \|^2$ is dominated by the "Euclidean norm" of the matrix: the square root of the sum of the squares of its components.) When applied to a vector, $\| \ \|$ denotes Euclidean or l_2 norm

Proof of Lemma 2.2

$$\left\|\left(\frac{W_n}{n}\right)^2\right\|^2 \leqq \frac{1}{n^4}\sum_{ij}\left(\sum_k w_{ik}w_{kj}\right)^2$$
$$=\frac{1}{n^4}\sum_{ijk}(w_{ik}w_{kj})^2+\frac{2}{n^4}\sum_{ij}\sum_{k>h}w_{ik}w_{kj}w_{ih}w_{hj}. \tag{2.2}$$

Let $S_n=\frac{1}{n^4}\sum_{ijk}(w_{ik}w_{kj})^2$. Then, since S_n^2 has n^6 terms, and since

$$E(w_{ik}w_{kj})^2(w_{lr}w_{rs})^2 \leqq E|w_{11}|^8,\; ES_n^2 \leqq cn^6/n^8=c/n^2.$$

Hence $E\sum_{n=1}^{\infty} S_n^2<\infty$, which implies $S_n \to 0$ a.s.

Concerning the indices of the second expression in (2.2), let $I=\{i,j,k,h)$: at least two of these indices are paired$\}$, and let $J=\{(i,j,k,h)\colon k>h\}$. The size of I is of order n^3, and therefore, reasoning as above,

$$\frac{2}{n^4}\sum_{I\cap J} w_{ik}w_{kj}w_{ih}w_{hj} \to 0 \text{ a.s.}$$

For (i,j,k,h) and (i',j',k',h') in $I^c\cap J$, we claim that

$$w_{ik}w_{kj}w_{ih}w_{hj} \tag{2.3}$$

and

$$w_{i'k'}w_{k'j'}w_{i'h'}w_{h'j'} \tag{2.4}$$

are orthogonal, unless $(i,j,k,h)=(i',j',k',h')$. Obviously the expressions in (2.3) and (2.4) have zero expectation. If they are not orthogonal, then each element in (2.3) matches an element in (2.4). If (say) $i \neq i'$, then either ($ik=k'j'$ and $ih=h'j'$) or ($ik=h'j'$ and $ih=k'j'$). Both of these choices leads to the conclusion $k=h$, which is not allowed. Hence $i=i'$. Similar reasoning establishes $j=j'$, $k=k'$, and $h=h'$. An application of Lemma 2.1, to the expression

$$\frac{2}{n^4}\sum_{I^c\cap J} w_{ik}w_{kj}w_{ih}w_{hj}$$

then completes the proof of Lemma 2.2.

Lemma 2.2 implies that for almost every ω, when n is sufficiently large,

$$\sum_{k=0}^{\infty}\left\|\left(\frac{W_n}{n}\right)^k\right\| \leqq \left\|\frac{W_n}{n}\right\|\sum_0^{\infty}\left\|\left(\frac{W_n}{n}\right)^2\right\|^k+\sum_0^{\infty}\left\|\left(\frac{W_n}{n}\right)^2\right\|^k$$

converges. Hence

Lemma 2.3. *If* $E|w_{11}|^8<\infty$, *then, almost surely,* $(I-W_n/n)^{-1}$ *exists and equals* $\sum_0^{\infty}\left(\frac{W_n}{n}\right)^k$ *when n is sufficiently large.*

We may now proceed with the proof of the Theorem:

i) It is, of course, enough to show $x_{n,1} \to \alpha_1$ a.s. Almost surely, for n sufficiently large,

$$x_n - v_n = \frac{1}{n} W_n \left(I - \frac{1}{n} W_n\right)^{-1} v_n = \sum_1^\infty \left(\frac{W_n}{n}\right)^k v_n.$$

Let A_1 denote the first row of a matrix A. For n large

$$x_{n,1} - v_{n,1} = \left(\frac{W_n}{\sqrt{n}}\right)_1 \left\{\sum_0^\infty \left(\frac{W_n}{n}\right)^k\right\} \frac{v_n}{\sqrt{n}} \Rightarrow |x_{n,1} - v_{n,1}|$$

$$\leqq \left|\frac{1}{n} \sum_1^n \alpha_i w_{1i}\right| \tag{2.5}$$

$$+ \left\|\left(\frac{W_n}{\sqrt{n}}\right)_1\right\| \left\|\frac{W_n}{n} \frac{v_n}{\sqrt{n}}\right\| \tag{2.6}$$

$$+ \left\|\left(\frac{W_n}{\sqrt{n}}\right)_1\right\| \left(\sup_{k=1,2,\dots} |\alpha_k|\right) \sum_{k=2}^\infty \left\|\left(\frac{W_n}{n}\right)^k\right\|. \tag{2.7}$$

'2.5'$\to 0$[2], a.s., by an application of Lemma 2.1. $\|W_n/\sqrt{n})_1\|$, in (2.6), converges (a.s.) to $(E w_{11}^2)^{1/2}$ by the usual law of large numbers. For the second expression in (2.6), write

$$\left\|\frac{W_n}{n} \frac{v_n}{\sqrt{n}}\right\|^2 = \frac{1}{n^3} \sum_{ki} \alpha_i^2 w_{ki}^2 + \frac{2}{n^3} \sum_k \sum_{i>j} \alpha_i \alpha_j w_{ki} w_{kj},$$

and apply Lemma 2.1 to each term. Conclude that $\left\|\frac{W_n}{n} \frac{v_n}{\sqrt{n}}\right\| \to 0$ a.s., and therefore '2.6'$\to 0$ a.s. In (2.7), for k even, write $\|(W_n/n)^k\| \leqq \|(W_n/n)^2\|^{k/2}$, and for k odd write $\|(W_n/n)^k\| \leqq \|W_n/n\| \; \|(W_n/n)^2\|^{(k-1)/2}$. Since $\|W_n/n\|^2 \leqq \frac{1}{n^2} \sum_{ij} w_{ij}^2 \to E w_{11}^2$, Lemma 2.2 implies '2.7'$\to 0$ a.s.

iii) First, we claim $\left\|\frac{1}{n} W_n v_n\right\| \to 0$ a.s.

$$\left\|\frac{1}{n} W_n v_n\right\|^2 = \frac{1}{n^2} \sum_{k=1}^n \left(\sum_{i=1}^n \alpha_i w_{ki}\right)^2$$

$$= \frac{1}{n^2} \sum_{ki} \alpha_i^2 w_{ki}^2 \tag{2.8}$$

$$+ \frac{2}{n^2} \sum_k \sum_{i>j} \alpha_i \alpha_j w_{ki} w_{kj}. \tag{2.9}$$

[2] Sometimes, it will be convenient to use line numbers to represent expressions

In (2.8) apply Lemma 2.1:

$$\frac{1}{n^2}\sum_{ki}\alpha_i^2 w_{ki}^2 - E\frac{1}{n^2}\sum_{ki}\alpha_i^2 w_{ki}^2 \to 0 \text{ a.s.}$$

But $E\frac{1}{n^2}\sum_{ki}\alpha_i^2 w_{ki}^2 = \frac{1}{n}Ew_{11}^2\sum_{i=1}^{n}\alpha_i^2 \to 0$. Hence '2.8'$\to 0$ a.s. For (2.9) we can write '2.9'$=\frac{2}{n^2}\sum_{k=1}^{n} t_k^{(n)}$, where $t_1^{(n)}, \ldots, t_n^{(n)}$ are i.i.d. with $Et_1^{(n)}=0$ and

$$\begin{aligned} E(t_1^{(n)})^2 &= E(\sum_{i>j}\alpha_i\alpha_j w_{1i}w_{1j})^2 = \sum_{i>j}\alpha_i^2\alpha_j^2 E w_{1i}^2 w_{1j}^2 \\ &= \sum_{i>j}\alpha_i^2\alpha_j^2 (Ew_{11}^2)^2 \leqq \left(\sum_{i=1}^{n}\alpha_i^2\right)^2 (Ew_{11}^2)^2. \end{aligned}$$

So $E(\text{'2.9'})^2 = O(1/n^3) \Rightarrow$ '2.9'$\to 0$ a.s.

Now, to prove $x_n \to v_n$ in l_2, use Lemma 2.3: when n is large

$$x_n = \left(I - \frac{1}{n}W_n\right)^{-1} v_n = v_n + \sum_{k=1}^{\infty}\left(\frac{W_n}{n}\right)^k v_n.$$

$$\begin{aligned} \text{But } \left\|\sum_{k=1}^{\infty}\left(\frac{W_n}{n}\right)^k v_n\right\| &\leqq \left\|\frac{1}{n}W_n v_n\right\| \sum_{k=0}^{\infty}\left\|\left(\frac{W_n}{n}\right)^2\right\|^k \\ &\quad + \|v_n\|\sum_{k=1}^{\infty}\left\|\left(\frac{W_n}{n}\right)^2\right\|^k \to 0 \text{ a.s.} \end{aligned}$$

by Lemma 2.2, and the fact (demonstrated above) that $\left\|\frac{1}{n}W_n v_n\right\| \to 0$ a.s.

ii) We will make use of the following lemma. The proof is omitted; it is a typical application of techniques used in proving large deviation theorems for sums of independent random variables (cf. Chernoff [3]).

Lemma 2.4. *If $x_1, x_2, \ldots$ are zero mean iid random variables, if $E[e^{i\lambda x_1}]$ is analytic at $\lambda=0$, and if $(\alpha_1, \alpha_2, \ldots) \in l_\infty$, then for every $\varepsilon>0$ there exists an $\varepsilon'>0$ such that*

$$P\left\{\left|\frac{1}{n}\sum_{1}^{n} x_i\alpha_i\right| > \varepsilon\right\} \leqq 2e^{-n\varepsilon'}$$

for all $n=1, 2, \ldots$.

For n sufficiently large, equation (2.1) holds and

$$\begin{aligned} x_{n,i} - \alpha_i &= \frac{1}{n}\sum_{j=1}^{n} w_{ij}(x_{n,j}-\alpha_j) + \frac{1}{n}\sum_{j=1}^{n} w_{ij}\alpha_j \\ &\Rightarrow |x_{n,i}-\alpha_i| \leqq (\sup_{1\leqq j\leqq n}|x_{n,j}-\alpha_j|)\frac{1}{n}\sum_{j=1}^{n}|w_{ij}| + \left|\frac{1}{n}\sum_{j=1}^{n} w_{ij}\alpha_j\right|. \end{aligned}$$

Let $S_n = \sup_{1 \leq i \leq n} |x_{n,i} - \alpha_i|$. Then

$$S_n \leqq S_n \sup_{1 \leqq i \leqq n} \frac{1}{n} \sum_{j=1}^{n} |w_{ij}| + \sup_{1 \leqq i \leqq n} \left| \frac{1}{n} \sum_{j=1}^{n} w_{ij} \alpha_j \right|. \tag{2.10}$$

Lemma 2.4 implies that $\sup_{1 \leqq i \leqq n} \left| \frac{1}{n} \sum_{j=1}^{n} w_{ij} \alpha_j \right| \to 0$, almost surely:

$$P\left(\sup_{1 \leqq i \leqq n} \left| \frac{1}{n} \sum_{j=1}^{n} w_{ij} a_j \right| > \varepsilon \right)$$

$$\leqq \sum_{i=1}^{n} P\left(\left| \frac{1}{n} \sum_{j=1}^{n} w_{ij} \alpha_j \right| > \varepsilon \right) \leqq 2n e^{-n\varepsilon'}$$

and apply the Borel-Cantelli Lemma. We can use a similar argument to show that $\frac{1}{n} \sum_{j=1}^{n} |w_{ij}|$ converges uniformly in i to $E|w_{11}|$ (almost surely). What is needed is the analyticity of

$$E[e^{i\lambda(|w_{11}| - E|w_{11}|)}],$$

or equivalently $E[e^{i\lambda|w_{11}|}]$, at $\lambda = 0$ (then apply Lemma 2.4 to $\frac{1}{n} \sum_{j=1}^{n} \{|w_{ij}| - E|w_{11}|\}$). Let $f(\lambda) = E[e^{i\lambda w_{11}}]$, which is, by assumption, analytic at $\lambda = 0$. Then, it is well known that for some $\lambda_0 > 0$,

$$f(\lambda) = \sum_{k=0}^{\infty} \frac{i^k \lambda^k}{k!} E[w_{11}^k]$$

for all $|\lambda| < \lambda_0$. Define $a_k = E|w_{11}|^k / k!$. Then

$$\overline{\lim_{k \to \infty}} \sqrt[k]{a_k} \leqq \overline{\lim_{k \to \infty}} (E|w_{11}|^{2k})^{\frac{1}{2k}} / (k!)^{1/k}$$

$$\leqq \overline{\lim_{k \to \infty}} \frac{(E|w_{11}|^{2k})^{\frac{1}{2k}}}{\{(2k)!\}^{\frac{1}{2k}}} \overline{\lim_{k \to \infty}} \frac{\{(2k)!\}^{\frac{1}{2k}}}{(k!)^{1/k}}$$

$$\leqq 2 \overline{\lim_{k \to \infty}} \frac{|E[w_{11}^k]|^{1/k}}{(k!)^{1/k}} < \infty,$$

(the latter inequality because $f(\lambda)$ is analytic at $\lambda = 0$). Hence, for all $|\lambda|$ sufficiently small,

$$E[e^{i\lambda|w_{11}|}] = E \sum_{k=0}^{\infty} \frac{i^k \lambda^k}{k!} |w_{11}|^k$$

$$= \sum_{k=0}^{\infty} \frac{i^k \lambda^k}{k!} E|w_{11}|^k.$$

Therefore, $E[e^{i\lambda|w_{11}|}]$ is analytic at $\lambda=0$, and with probability one,

$$\frac{1}{n}\sum_{j=1}^{n}|w_{ij}|\to E|w_{11}|$$

uniformly in i.

Finally, return to (2.10):

$$\begin{aligned}\overline{\lim_{n\to\infty}}\, S_n &\leqq \overline{\lim_{n\to\infty}}\, S_n \overline{\lim_{n\to\infty}} \sup_{1\leqq i\leqq n}\frac{1}{n}\sum_{j=1}^{n}|w_{ij}| \\ &+\overline{\lim_{n\to\infty}} \sup_{1\leqq i\leqq n}\left|\frac{1}{n}\sum_{j=1}^{n} w_{ij}\alpha_j\right| = E|w_{11}|\overline{\lim_{n\to\infty}}\, S_n.\end{aligned}$$

Since $E|w_{11}|<1$, $\overline{\lim_{n\to\infty}}\, S_n=0$ a.s.

Proof of Theorem 2. Let M_n be the $n\times n$ matrix with every component equal to m, let $\hat{W}_n=W_n-M_n$, and let s_n be the n-dimensional column vector with each component equal to $\alpha/(1-m)$. Under the moment condition on w_{11}, the theorem in Geman [5] says that $\|\hat{W}_n/\sqrt{n}\|\to 2\sqrt{E(w_{11}-m)^2}$ almost surely, which implies $\|\hat{W}_n/n\|\to 0$ a.s. From this we can conclude that

$$\|W_n/n\|\to|m| \text{ a.s.: } \|M_n/n\|-\|\hat{W}_n/n\|\leqq\|W_n/n\|\leqq\|M_n/n\|+\|\hat{W}_n/n\|$$

and $\|M_n/n\|=|m|$. Since $|m|<1$, (2.1) holds for all n sufficiently large (a.s.). After some algebra, (2.1) can be written

$$\begin{aligned} x_n-s_n &= \frac{1}{n}\hat{W}_n(x_n-s_n)+\frac{1}{n}M_n(x_n-s_n)+\frac{1}{n}\hat{W}_n s_n \\ \Rightarrow \|x_n-s_n\| &\leqq \left(\left\|\frac{\hat{W}_n}{n}\right\|+\left\|\frac{M_n}{n}\right\|\right)\|x_n-s_n\|+\left\|\frac{\hat{W}_n}{\sqrt{n}}\right\|\left\|\frac{s_n}{\sqrt{n}}\right\|. \end{aligned} \tag{2.11}$$

And, therefore, since $\|s_n/\sqrt{n}\|=|\alpha/(1-m)|$, there is a constant c such that $\overline{\lim}\,\|x_n-s_n\|\leqq c$ a.s.

Return to (2.11):

$$\begin{aligned}\left|x_{n,1}-\frac{\alpha}{1-m}\right| &= |(x_n-s_n)_1| \\ &\leqq \left\|\frac{\hat{W}_n}{n}\right\|\|x_n-s_n\|+\left|\frac{m}{n}\sum_{i=1}^{n}(x_{n,i}-s_{n,i})\right|+\left|\frac{\alpha}{1-m}\frac{1}{n}\sum_{i=1}^{n}(w_{1i}-m)\right| \\ &\leqq \left\|\frac{\hat{W}_n}{n}\right\|\|x_n-s_n\|+\|x_n-s_n\|\frac{|m|}{\sqrt{n}}+\left|\frac{\alpha}{1-m}\right|\left|\frac{1}{n}\sum_{1}^{n}(w_{1i}-m)\right|\to 0. \text{ a.s.} \quad \square\end{aligned}$$

3. CLT for an Algebraic System of Equations

The discussion in Sect. 2 suggests that a renormalization may produce a more interesting limit. For each n define 1_n to be the n dimensional column vector composed only of 1's, and define $x_n=(x_{n,1}, \ldots, x_{n,n})^T$ as the solution to

$$x_n = 1_n + \frac{1}{\sqrt{n}} W_n x_n$$
$$\text{i.e. } x_{n,i} = 1 + \frac{1}{\sqrt{n}} \sum_{j=1}^{n} w_{ij} x_{n,j} \quad 1 \leqq i \leqq n \tag{3.1}$$

where W_n is again composed of independent and identically distributed random variables $\{w_{ij}\}$ with zero means. As before, reason from intuition: when the system is large all variables should be "nearly independent". This leads to the conjecture that for fixed m

$$(x_{n,1}, \ldots, x_{n,m}) \xrightarrow{W} \text{i.i.d. } N(\mu, \gamma^2). \tag{3.2}$$

To determine μ and γ^2, *pretend* that all moments of $x_{n,i}$ exist and that all variables actually are independent:

$$\mu = E x_{n,i} = 1 + \frac{1}{\sqrt{n}} \sum_{j=1}^{n} E w_{ij} E x_{n,j} = 1 \tag{3.3}$$

$$\begin{aligned} E x_{n,i}^2 &= E\left\{1 + \frac{1}{\sqrt{n}} \sum_{j=1}^{n} w_{ij} x_{n,j}\right\}^2 \\ &= 1 + \frac{1}{n} \sum_{j=1}^{n} E w_{ij}^2 E x_{n,j}^2 \\ &= 1 + E w_{11}^2 E x_{n,i}^2 \\ &\Rightarrow \gamma^2 = E x_{n,i}^2 - \{E x_{n,i}\}^2 = \frac{E w_{11}^2}{1 - E w_{11}^2}. \end{aligned} \tag{3.4}$$

Although here it is much harder to prove, a chaos hypothesis has again led us to precisely the right asymptotic (large system) behavior:

Theorem 3. *Define* x_n *by* (3.1) *whenever* $I - \frac{1}{\sqrt{n}} W_n$ *is nonsingular. Otherwise, define* $x_n = 0$. *If* $E w_{11}^2 < \frac{1}{4}$ *and if there exists a constant* α *such that* $E|w_{11}|^n \leqq n^{\alpha n}$ $\forall n \geqq 2$, *then, for every fixed m,* (3.2) *holds with* μ *and* γ^2 *given by* (3.3) *and* (3.4) *respectively.*

Remark. From (3.4), it appears unlikely that (3.2) will hold, for *any* γ^2, when $E w_{11}^2 > 1$. But the condition $E w_{11}^2 < 1/4$ is probably too strong. For example, computer simulations support (3.2) for the particular case w_{11} uniform on $[-1, 1]$, in which $E w_{11}^2 = 1/3$.

Proof. Put $\sigma^2 = Ew_{11}^2$. Fix m and $\lambda_1, \ldots, \lambda_m$. We will show that

$$E\exp\left\{i\sum_1^m \lambda_i x_{n,i}\right\} \to \exp\left\{i\sum_1^m \lambda_i - \frac{1}{2}\frac{\sigma^2}{1-\sigma^2}\sum_1^m \lambda_i^2\right\}. \tag{3.5}$$

For each p and n define

$$x_n^p = 1_n + \sum_{k=1}^{p-1}\left(\frac{W_n}{\sqrt{n}}\right)^k 1_n.$$

An obvious consequence of Lemma A.1. (see Appendix) is that

$$\sum_{i=1}^m \lambda_i x_{n,i}^p \xrightarrow{W} N\left(\sum_{i=1}^m \lambda_i, \frac{\sigma^2-\sigma^{2p}}{1-\sigma^2}\sum_{i=1}^m \lambda_i^2\right). \tag{3.6}$$

Now define $e_n^p = x_n - x_n^p$. Then, whenever $I - \frac{1}{\sqrt{n}} W_n$ is nonsingular,

$$\begin{aligned}
x_n - x_n^p &= 1_n + \frac{1}{\sqrt{n}} W_n x_n - x_n^p \\
&= \frac{1}{\sqrt{n}} W_n(x_n - x_n^p) + \frac{1}{\sqrt{n}} W_n x_n^p - \sum_{k=1}^{p-1}\left(\frac{W_n}{\sqrt{n}}\right)^k 1_n \\
&= \frac{1}{\sqrt{n}} W_n(x_n - x_n^p) + \sum_{k=1}^{p}\left(\frac{W_n}{\sqrt{n}}\right)^k 1_n - \sum_{k=1}^{p-1}\left(\frac{W_n}{\sqrt{n}}\right)^k 1_n \\
&= \frac{1}{\sqrt{n}} W_n(x_n - x_n^p) + \left(\frac{W_n}{\sqrt{n}}\right)^p 1_n
\end{aligned}$$

i.e.

$$e_n^p = \frac{1}{\sqrt{n}} W_n e_n^p + \left(\frac{W_n}{\sqrt{n}}\right)^p 1_n. \tag{3.7}$$

We claim

$$\overline{\lim_{p\to\infty}}\ \overline{\lim_{n\to\infty}}\ P\left(\left|\sum_1^m \lambda_i e_{n,i}^p\right| > \varepsilon\right) = 0 \qquad \forall \varepsilon > 0. \tag{3.8}$$

To prove (3.8), choose $\delta < 1$ such that $\delta > 2\sigma$ (recall that $\sigma^2 < \frac{1}{4}$). Let $A_n = \{\omega : |W_n/\sqrt{n}| < \delta\}$. Then (see Geman [5]) $I_{A_n}(\omega) \to 1$ a.s. and for (3.8) it will suffice to show

$$\overline{\lim_{p\to\infty}}\ \overline{\lim_{n\to\infty}}\ P(|I_{A_n} e_{n,i}^p| > \varepsilon) = 0 \qquad \forall \varepsilon > 0 \tag{3.9}$$

for $i = 1, 2, \ldots, m$. From (3.7):

$$
\begin{aligned}
&|I_{A_n} e_n^p| \leqq \delta |I_{A_n} e_n^p| + \sqrt{n}\,\delta^p \\
&\quad \Rightarrow I_{A_n} \sum_1^n |e_{n,i}^p|^2 \leqq \frac{n\delta^{2p}}{(1-\delta)^2} \\
&\quad \Rightarrow E\{I_{A_n} |e_{n,i}^p|^2\} \leqq \frac{\delta^{2p}}{(1-\delta)^2} \quad \text{(by symmetry).}
\end{aligned}
$$

Hence

$$
\begin{aligned}
&\overline{\lim_{p\to\infty}}\ \overline{\lim_{n\to\infty}}\ P(|I_{A_n} e_{n,i}^p| > \varepsilon) \\
&\quad \leqq \overline{\lim_{p\to\infty}}\ \overline{\lim_{n\to\infty}}\ E\{I_{A_n} |e_{n,i}^p|^2\}/\varepsilon^2 \leqq \overline{\lim_{p\to\infty}} \frac{\delta^{2p}}{(1-\delta)^2\,\varepsilon^2} = 0
\end{aligned}
$$

which is (3.9).

Finally, return to (3.5):

$$
\begin{aligned}
&\overline{\lim_{n\to\infty}} \left| E \exp\left\{ i \sum_1^m \lambda_i x_{n,i} \right\} - \exp\left\{ i \sum_1^m \lambda_i - \frac{1}{2} \frac{\sigma^2}{1-\sigma^2} \sum_1^m \lambda_i^2 \right\} \right| \\
&\quad \leqq \overline{\lim_{n\to\infty}} \left| E \exp\left\{ i \sum_1^m \lambda_i x_{n,i} \right\} - E \exp\left\{ i \sum_1^m \lambda_i x_{n,i}^p \right\} \right| \\
&\qquad + \overline{\lim_{n\to\infty}} \left| E \exp\left\{ i \sum_1^m \lambda_i x_{n,i}^p \right\} - \exp\left\{ i \sum_1^m \lambda_i - \frac{1}{2} \frac{\sigma^2}{1-\sigma^2} \sum_1^m \lambda_i^2 \right\} \right| \\
&\quad = \overline{\lim_{n\to\infty}} \left| E\left[\exp\left\{ i \sum_1^m \lambda_i x_{n,i}^p \right\} \left(\exp\left\{ i \sum_1^m \lambda_i e_{n,i}^p \right\} - 1 \right) \right] \right| \\
&\qquad + \left| \exp\left\{ i \sum_1^m \lambda_i - \frac{1}{2} \frac{\sigma^2 - \sigma^{2p}}{1-\sigma^2} \sum_1^m \lambda_i^2 \right\} \right. \\
&\qquad \left. - \exp\left\{ i \sum_1^m \lambda_i - \frac{1}{2} \frac{\sigma^2}{1-\sigma^2} \sum_1^m \lambda_i^2 \right\} \right| \quad \text{(by 3.6)} \\
&\quad \leqq \overline{\lim_{n\to\infty}}\ E \left| \exp\left\{ i \sum_1^m \lambda_i e_{n,i}^p \right\} - 1 \right| \\
&\qquad + \left| \exp\left\{ -\frac{1}{2} \frac{\sigma^2}{1-\sigma^2} \sum_1^m \lambda_i^2 \right\} \left(\exp\left\{ \frac{1}{2} \frac{\sigma^{2p}}{1-\sigma^2} \sum_1^m \lambda_i^2 \right\} - 1 \right) \right|.
\end{aligned}
$$

Finally, let $p \to \infty$; then both terms above converge to 0, the former by (3.8) and the latter because $\sigma^2 < 1$. This establishes (3.5) and completes the proof. □

4. CLT for a System of Differential Equations

In this article we will be content to limit the discussion to linear systems. As remarked earlier, we have examples of a "chaos principle" operating in some nonlinear systems (specifically, an LLN for a nonlinear system of differential equations), but these will be discussed elsewhere. Still, within the linear context we can considerably extend our treatment by introducing time dependence; here we propose to treat a random system of ordinary differential equations in a manner analogous to the approach of Sect. 2 and 3. In other words, we will first make a chaos hypothesis for the system, then derive its consequences in the form of a conjecture, and then rigorously prove the conjecture.

Follow the previous notation: $\{w_{ij}\}$, $i=1,2,\ldots$, $j=1,2,\ldots$ is an iid collection of zero mean and finite variance random variables, and W_n is the $n\times n$ matrix $\{w_{ij}\}$ $1\leqq i\leqq n$, $1\leqq j\leqq n$. α will indicate an arbitrary constant, and 1_n will represent the column n-vector of 1's. For each n, define an R^n-valued random process $x_n(t)=(x_{n,1}(t),\ldots,x_{n,n}(t))^T$ by

$$\dot{x}_n=\alpha x_n+\frac{1}{\sqrt{n}}W_n x_n \qquad x_n(0)=1_n \tag{4.1}$$

i.e.

$$\dot{x}_{n,i}(t)=\alpha x_{n,i}(t)+\frac{1}{\sqrt{n}}\sum_{j=1}^{n} w_{ij}x_{n,j}(t) \qquad x_{n,i}(0)=1$$
$$1\leqq i\leqq n$$

($\cdot$ means derivative with respect to t). We might as well assume that $Ew_{11}^2=1$, since any other variance can be brought to 1 by scaling t and changing α.

What sort of behavior should we expect from (4.1) when n is large? We might try to infer the asymptotic properties from a direct solution of the equation:

$$x_n(t)=e^{\alpha t}\exp\left\{\frac{1}{\sqrt{n}}W_n t\right\}1_n,$$

but this expression does not suggest any particular limiting behavior. Let us instead reason intuitively: first, since the components $x_{n,1}(\cdot),\ldots,x_{n,n}(\cdot)$ are becoming "weakly coupled", we expect asymptotic (large n) independence - at least for any fixed set, $x_{n,1}(\cdot),\ldots,x_{n,m}(\cdot)$. To guess the limit for a particular component, rewrite (4.1) as follows:

$$x_{n,i}=e^{\alpha t}+\frac{1}{\sqrt{n}}\sum_{j=1}^{n} w_{ij}\int_0^t e^{\alpha(t-u)}x_{n,j}(u)\,du. \tag{4.2}$$

For any collection of times $t_1,\ldots,t_l$, and any constants $\lambda_1,\ldots,\lambda_l$,

$$\sum_{k=1}^{l}\lambda_k x_{n,i}(t_k)=\sum_{k=1}^{l}\lambda_k e^{\alpha t_k}+\frac{1}{\sqrt{n}}\sum_{j=1}^{n} w_{ij}\left\{\sum_{k=1}^{l}\lambda_k\int_0^{t_k} e^{\alpha(t_k-u)}x_{n,j}(u)\,du\right\},$$

and our usual chaos hypothesis suggests that the right hand side is asymptotically Gaussian. We have, then, the conjecture that for fixed m, $(x_{n,1}(\cdot),\ldots,x_{n,m}(\cdot))$ converges weakly to *iid* Gaussian processes.

For the mean and correlation functions return to (4.2) and pretend, whenever necessary, to have complete independence:

$$\mu(t)=Ex_{n,i}(t)=e^{\alpha t}+\frac{1}{\sqrt{n}}\sum_{j=1}^{n} Ew_{ij}\int_0^t e^{\alpha(t-u)}Ex_{n,j}(u)\,du$$
$$=e^{\alpha t}$$

$$\begin{aligned}r(t,s)&=Ex_{n,i}(t)\,x_{n,i}(s)\\&=E\left\{\left(e^{\alpha t}+\frac{1}{\sqrt{n}}\sum_{j=1}^{t} w_{ij}\int_0^t e^{\alpha(t-u)}x_{n,j}(u)\,du\right)\right.\\&\quad\times\left.\left(e^{\alpha s}+\frac{1}{\sqrt{n}}\sum_{j=1}^{n} w_{ij}\int_0^s e^{\alpha(s-v)}x_{n,j}(v)\,dv\right)\right\}\\&=e^{\alpha t}e^{\alpha s}+\frac{1}{n}\sum_{j=1}^{n} Ew_{ij}^2\int_0^t\int_0^s e^{\alpha(t-u)}e^{\alpha(s-v)}r(u,v)\,du\,dv\\&=e^{\alpha t}e^{\alpha s}+\int_0^t\int_0^s e^{\alpha(t-u)}e^{\alpha(s-v)}r(u,v)\,du\,dv.\end{aligned}$$

Differentiating twice:

$$r_{ts}=\alpha r_t+\alpha r_s+(1-\alpha^2)\,r. \tag{4.3}$$

The appropriate boundary conditions are $r(t,0)=Ex_{n,i}(t)\;x_{n,i}(0)=Ex_{n,i}(t)=e^{\alpha t}$ and, similarly, $r(0,s)=e^{\alpha s}$. The solution of (4.3), say by the method of Laplace transforms, is:

$$r(t,s)=e^{\alpha(t+s)}J_0(2\sqrt{-ts})=e^{\alpha(t+s)}\sum_{k=0}^{\infty}\frac{t^k s^k}{(k!)^2},$$

J_0 being the zero'th Bessel function of the first kind. It will be conventient to work with the covariance function, $c(t,s)=Ex_{n,i}(t)\,x_{n,i}(s)-\mu(t)\,\mu(s)$. The conjecture becomes

$$c(t,s)=e^{\alpha(t+s)}\sum_{k=1}^{\infty}\frac{t^k s^k}{(k!)^2}.$$

Once again, we have identified correctly the large n behavior:

Theorem 4. *Suppose there exists a constant β such that $E|w_{11}|^n\leqq n^{\beta n}$ $\forall n\geqq 2$. Then for each fixed m and for each fixed T, $(x_{n,1}(\cdot),\ldots,x_{n,m}(\cdot))$ converges weakly on $[0,T]$ to independent and identically distributed Gaussian processes with mean*

$$\mu(t)=e^{\alpha t}$$

and covariance

$$c(t,s)=e^{\alpha(t+s)}\sum_{k=1}^{\infty}\frac{t^k s^k}{(k!)^2}.$$

Remarks. 1. The asymptotics of J_0 (cf. Olver [7]) provide us with some insight into the nature of the limiting process:

$$c(t,s)=e^{\alpha(t+s)}\{J_0(2\sqrt{-ts})-1\}\sim\tfrac{1}{2}e^{\alpha(t+s)}(\pi^2 ts)^{-1/4}e^{2\sqrt{ts}}$$

for large t or s, with t and s positive. Hence:

a) $\operatorname{Var}(x_{n,1}(t))=c(t,t)\to\begin{cases}0 & \alpha\leqq -1\\ \infty & \alpha>-1.\end{cases}$

b) Correlation coefficient $=c(t,s)/\sqrt{c(t,t)\,c(s,s)}\to 0$ for all α, as $t\to\infty$ or $s\to\infty$ with the other variable remaining fixed.

2. We believe that the analogous theorem holds for the equation

$$\dot{x}_n=\alpha x_n+\frac{1}{\sqrt{n}}W_n x_n+1_n.$$

Although, in this case, the limiting process appears to have a much more complicated covariance function.

Proof. The solution to (4.1) is

$$x_n(t)=e^{\alpha t}\exp\left\{\frac{1}{\sqrt{n}}W_n t\right\}1_n.$$

If we define $y_n(t)=e^{-\alpha t}x_n(t)$, then the theorem becomes: $(y_{n,1}(\cdot),\ldots,y_{n,m}(\cdot))$ converges weakly on $[0,T]$ to i.i.d. Gaussian processes with constant mean $=1$ and covariance

$$\sum_{k=1}^{\infty}\frac{t^k s^k}{(k!)^2}.$$

We will show first that $\{(y_{n,1}(\cdot),\ldots,y_{n,m}(\cdot))\}_{n=1}^{\infty}$ is tight in $C^m[0,T]$, and then that the finite dimensional distribution functions converge appropriately.

In light of the symmetry in this problem, tightness will require only that (cf. Billingsley [1], Chap. 2)

$$\lim_{\delta\to 0}\overline{\lim_{n\to\infty}}P(\sup_{|s-t|<\delta}|y_{n,1}(s)-y_{n,1}(t)|>\varepsilon)=0 \tag{4.4}$$

for arbitrary $\varepsilon>0$ (s and t in $[0,T]$). From $y_n(t)=\exp\left\{\frac{1}{\sqrt{n}}W_n t\right\}1_n$ we get

$$y_{n,1}(s)-y_{n,1}(t)=\sum_{k=1}^{\infty}\left\{\left(\frac{1}{\sqrt{n}}W_n\right)^k 1_n\right\}_1\frac{s^k-t^k}{k!}.$$

Let

$$\alpha_k(n)=\left|\left\{\left(\frac{1}{\sqrt{n}}W_n\right)^k 1_n\right\}_1\right| \quad\text{and}\quad \beta_k(\delta)=(T+\delta)^k-T^k.$$

Then

$$\sup_{|s-t|<\delta} |y_{n,1}(s)-y_{n,1}(t)| \leqq \sum_{k=1}^{\infty} \alpha_k(n) \frac{\beta_k(\delta)}{k!}.$$

In Geman [5] it was shown that $\left|\frac{1}{\sqrt{n}} W_n\right| \to 2$ a.s. It follows that for any $\gamma > 2$, $I_{A_n}(\omega) \to 1$ a.s., where $A_n = \left\{\omega : \left|\frac{1}{\sqrt{n}} W_n\right| < \gamma\right\}$. Return now to (4.4):

$$\begin{aligned}
&\lim_{\delta\to 0} \overline{\lim_{n\to\infty}} P\big(\sup_{|s-t|<\delta} |y_{n,1}(s)-y_{n,1}(t)| > \varepsilon\big) \\
&\quad \leqq \lim_{\delta\to 0} \overline{\lim_{n\to\infty}} P\left(\sum_{k=1}^{\infty} \alpha_k(n) \frac{\beta_k(\delta)}{k!} > \varepsilon\right) \\
&\quad = \lim_{\delta\to 0} \overline{\lim_{n\to\infty}} P\left(\sum_{k=1}^{\infty} I_{A_n} \alpha_k(n) \frac{\beta_k(\delta)}{k!} > \varepsilon\right) \\
&\quad \leqq \lim_{\delta\to 0} \overline{\lim_{n\to\infty}} \frac{1}{\varepsilon} \sum_{k=1}^{\infty} E(I_{A_n} \alpha_k(n)) \frac{\beta_k(\delta)}{k!}.
\end{aligned}$$

Arguing, again, by symmetry:

$$\begin{aligned}
E(I_{A_n} \alpha_k(n))^2 &= \frac{1}{n} E\left\{ I_{A_n} \left| \left(\frac{1}{\sqrt{n}} W_n\right)^k 1_n \right|^2 \right\} \\
&\leqq \frac{1}{n} E\{\gamma^{2k} |1_n|^2\} = \gamma^{2k}.
\end{aligned}$$

Finally then,

$$\begin{aligned}
&\lim_{\delta\to 0} \overline{\lim_{n\to\infty}} P\big(\sup_{|s-t|<\delta} |y_{n,1}(s)-y_{n,1}(t)| > \varepsilon\big) \\
&\quad \leqq \lim_{\delta\to 0} \frac{1}{\varepsilon} \sum_{k=1}^{\infty} \frac{\gamma^k \beta_k(\delta)}{k!} = 0.
\end{aligned}$$

It remains for us to establish appropriate convergence for the finite dimensional distribution functions. Choose, arbitrarily, numbers $\{\lambda_{ij}\}$ where $1 \leqq i \leqq m$ and $1 \leqq j \leqq l$, and times $t_1, \ldots, t_l$ where $t_j \in [0, T]$ for each j. With

$$\hat{c}(t, s) = \sum_{k=1}^{\infty} \frac{t^k s^k}{(k!)^2}$$

we wish to show

$$\begin{aligned}
&E \exp\left\{ i \sum_{i=1}^{m} \sum_{j=1}^{l} \lambda_{ij} y_{n,i}(t_j) \right\} \\
&\quad \to \exp\left\{ i \sum_{i=1}^{m} \sum_{j=1}^{l} \lambda_{ij} - \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{l} \sum_{q=1}^{l} \lambda_{ij} \lambda_{iq} \hat{c}(t_j, t_q) \right\}.
\end{aligned}$$

For this purpose, define

$$y_n^p = \sum_{k=0}^{p} \left(\frac{W_n}{\sqrt{n}}\right)^k 1_n \frac{t_k}{k!}$$

and

$$e_n^p = \sum_{k=p+1}^{\infty} \left(\frac{W_n}{\sqrt{n}}\right)^k 1_n \frac{t^k}{k!}.$$

Observe that $y_n = y_n^p + e_n^p$.

$$\overline{\lim_{n\to\infty}} \left| E \exp\left\{ i \sum_{i=1}^{m} \sum_{j=1}^{l} \lambda_{ij} y_{n,i}(t_j) \right\} - \exp\left\{ i \sum_{i=1}^{m} \sum_{j=1}^{l} \lambda_{ij} - \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{l} \sum_{q=1}^{l} \lambda_{ij} \lambda_{iq} \hat{c}(t_j, t_q) \right\} \right|$$

$$\le \overline{\lim_{p\to\infty}}\ \overline{\lim_{n\to\infty}} \left| E \exp\left\{ i \sum_{i=1}^{m} \sum_{j=1}^{l} \lambda_{ij} y_{n,i}(t_j) \right\} - E \exp\left\{ i \sum_{i=1}^{m} \sum_{j=1}^{l} \lambda_{ij} y_{n,i}^p(t_j) \right\} \right| \tag{4.5}$$

$$+ \overline{\lim_{p\to\infty}}\ \overline{\lim_{n\to\infty}} \left| E \exp\left\{ i \sum_{i=1}^{m} \sum_{j=1}^{l} \lambda_{ij} y_{n,i}^p(t_j) \right\} - \exp\left\{ i \sum_{i=1}^{m} \sum_{j=1}^{l} \lambda_{ij} - \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{l} \sum_{q=1}^{l} \lambda_{ij} \lambda_{iq} \hat{c}(t_j, t_q) \right\} \right|. \tag{4.6}$$

In fact, the expressions in both (4.5) and (4.6) are zero. Start with (4.5) (use '4.5' to refer to the expression in (4.5)):

$$\text{'4.5'} \le E \left| \exp\left\{ i \sum_{i=1}^{m} \sum_{j=1}^{l} \lambda_{ij} e_{n,i}^p(t_j) \right\} - 1 \right|$$

and therefore it will suffice to show

$$\overline{\lim_{p\to\infty}}\ \overline{\lim_{n\to\infty}} P(|e_{n,i}^p(t_j)| > \varepsilon) = 0$$

for every $\varepsilon > 0$, $1 \le i \le m$, and $1 \le j \le l$. Take γ and A_n just as in the first part of this proof. Then

$$\overline{\lim_{p\to\infty}}\ \overline{\lim_{n\to\infty}} P(|e_{n,i}^p(t_j)| > \varepsilon) = \overline{\lim_{p\to\infty}}\ \overline{\lim_{n\to\infty}} P(I_{A_n} |e_{n,i}^p(t_j)| > \varepsilon)$$

$$\le \overline{\lim_{p\to\infty}}\ \overline{\lim_{n\to\infty}} \frac{1}{\varepsilon^2} E |I_{A_n} e_{n,i}^p(t_j)|^2$$

$$= \overline{\lim_{p\to\infty}}\ \overline{\lim_{n\to\infty}} \frac{1}{\varepsilon^2} \frac{1}{n} E |I_{A_n} e_n^p(t_j)|^2 \text{ (once again, by symmetry)}$$

$$\le \overline{\lim_{p\to\infty}}\ \overline{\lim_{n\to\infty}} \frac{1}{\varepsilon^2} \frac{1}{n} E \left(I_{A_n} \sum_{k=p+1}^{\infty} \left| \frac{W_n}{\sqrt{n}} \right|^k |1_n| \left| \frac{t_j^k}{k!} \right| \right)^2$$

$$\le \overline{\lim_{p\to\infty}}\ \overline{\lim_{n\to\infty}} \frac{1}{\varepsilon^2} \frac{1}{\mathrm{n}} \left(\sqrt{\mathrm{n}} \sum_{k=p+1}^{\infty} \frac{\gamma^k t_j^k}{\mathrm{k}!} \right)^2 = 0.$$

For the expression in (4.6), make the definition

$$\alpha(i,k,n)=\left[\left(\frac{W_n}{\sqrt{n}}\right)^k 1_n\right]_i$$

and write

$$\sum_{i=1}^{m}\sum_{j=1}^{l}\lambda_{ij}\,y_{n,i}^{p}(t_j)=\sum_{i=1}^{m}\sum_{j=1}^{l}\sum_{k=0}^{p}\lambda_{ij}\,\alpha(i,k,n)\frac{t_j^k}{k!}.$$

Then

$$\begin{aligned}
\text{`4.6'}\leq \overline{\lim_{p\to\infty}}\ \overline{\lim_{n\to\infty}}\Bigg|&E\exp\left\{i\sum_{i=1}^{m}\sum_{j=1}^{l}\sum_{k=0}^{p}\lambda_{ij}\,\alpha(i,k,n)\frac{t_j^k}{k!}\right\}\\
&-\exp\left\{i\sum_{i=1}^{m}\sum_{j=1}^{l}\lambda_{ij}-\frac{1}{2}\sum_{i=1}^{m}\sum_{k=1}^{p}\sum_{j=1}^{l}\sum_{q=1}^{l}\lambda_{ij}\lambda_{iq}\frac{t_j^k t_q^k}{(k!)^2}\right\}\Bigg| \qquad (4.7)\\
+\overline{\lim_{p\to\infty}}\ \overline{\lim_{n\to\infty}}\Bigg|&\exp\left\{i\sum_{i=1}^{m}\sum_{j=1}^{l}\lambda_{ij}-\frac{1}{2}\sum_{i=1}^{m}\sum_{k=1}^{p}\sum_{j=1}^{l}\sum_{q=1}^{l}\lambda_{ij}\lambda_{iq}\frac{t_j^k t_q^k}{(k!)^2}\right\}\\
&-\exp\left\{i\sum_{i=1}^{m}\sum_{j=1}^{l}\lambda_{ij}-\frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{l}\sum_{q=1}^{l}\lambda_{ij}\lambda_{iq}\,\hat{c}(t_j,t_q)\right\}\Bigg|. \qquad (4.8)
\end{aligned}$$

The expression in (4.7) is zero as a direct consequence of Lemma A.1 (of Appendix); the expression in (4.8) is zero since

$$\lim_{p\to\infty}\sum_{1}^{p}\frac{t_j^k t_q^k}{(k!)^2}=\hat{c}(t_j,t_q). \quad \square$$

Appendix

Let $\{w_{ij}\}$, $i=1,2,\ldots$, $j=1,2,\ldots$ be a collection of i.i.d. zero mean random variables, having all moments finite. Let σ^2 denote the second moment, Ew_{11}^2, and, for each n, let W_n denote the $n\times n$ matrix $\{w_{ij}\}$ $1\leq i,j\leq n$. The result presented here (Lemma A.1) describes some aspects of the asymptotic (large n) distribution of W_n. It has been referenced in the proofs of Theorems 3 and 4, and is the essential part of both of these.

Lemma A.1 is an instance of the central limit theorem for dependent random variables. As there is a large literature addressed to weak limits for dependent random variables, one might expect to discover that the result is a corollary of a more general, known result. To the best of our knowledge, this is not the case. The best results, among those potentially relevant to the problem here, appear to be those by Chen [2]. Unfortunately, the dependencies among the random variables discussed below are too strong to fit either Chen's theorems or, it would seem, any obvious modification of them.

Let 1_n be the column vector of 1's and define, for each i, k and n,

$$\begin{aligned}
\alpha(i,k,n)&=\left[\left(\frac{W_n}{\sqrt{n}}\right)^k 1_n\right]_i = i^{\text{th}} \text{ component of } \left(\frac{W_n}{\sqrt{n}}\right)^k 1_n\\
&=\left(\frac{1}{\sqrt{n}}\right)^k\sum_{l_1,\ldots,l_k} w_{il_1}w_{l_1l_2}\ldots w_{l_{k-1}l_k}.
\end{aligned}$$

Lemma A.1. *For any m, and any set of m distinct ordered pairs* $(i_1, k_1), \ldots, (i_m, k_m)$,

$$(\alpha(i_1, k_1, n), \ldots, \alpha(i_m, k_m, n)) \xrightarrow{W} (Z_1, \ldots, Z_m)$$

where $Z_1, \ldots, Z_m$ *are independent zero mean normal random variables with* $EZ_j^2 = \sigma^{2k_j}$.

Remark. So, for example, the 1st and 2nd row sums of $\frac{1}{n} W_n^2$ are asymptotically independent $N(0, \sigma^4)$ random variables.

Proof of Lemma A.1. By the "method of moments".

Fix m ordered pairs $(i_1, k_1), \ldots, (i_m, k_m)$, and define $Z_1, \ldots, Z_m$ as in the statement of the lemma. Let $n_1, \ldots, n_m$ be an arbitrary collection of m integers. We will show that

$$E \prod_{j=1}^{m} \alpha(i_j, k_j, n)^{n_j} \xrightarrow[n\to\infty]{} E \prod_{j=1}^{m} Z_j^{n_j} = \begin{cases} \prod_{j=1}^{m} \left\{ \sigma^{k_j n_j} \prod_{p=1}^{n_j/2} (2p-1) \right\} & \text{all } n_j\text{'s even} \\ 0 & \text{otherwise.} \end{cases} \tag{a.1}$$

Let C_p^j, $1 \leqq p \leqq n_j$, denote an element of that sum which comprises the p^{th} occurance, in $\prod_{j=1}^{m} \alpha(i_j, k_j, n)^{n_j}$, of $\alpha(i_j, k_j, n)$. C_p^j has the form

$$C_p^j = w_{i_j l_1} w_{l_1 l_2} \ldots w_{l_{k_j-1} l_{k_j}}.$$

Call such a sequence of w's a *chain.* Using this notation we can write

$$E \prod_{j=1}^{m} \alpha(i_j, k_j, n)^{n_j} = \left(\frac{1}{\sqrt{n}}\right)^{\sum_{j=1}^{m} n_j k_j} \sum E\{C_1^1 \ldots C_{n_1}^1\, C_1^2 \ldots C_{n_2}^2 \ldots C_1^m \ldots C_{n_m}^m\} \tag{a.2}$$

where the sum is taken over all free indices, i.e. all indices except the first index of each chain (i_j in the example above). Notice that there are $\sum_{j=1}^{m} n_j k_j$ free indices.

The following definitions and conventions will be used:

1. A particular element of the sum in (a.2), (without the expectation) will be called a *chain sequence* (i.e. what is within the brackets { }).
2. An *element* of a chain sequence is a particular w in that sequence.
3. The elements will be considered to be ordered by their appearance in the chain sequence, the left-most element being the first.

For example, if $m=2$, $i_1=1$, $i_2=3$, $k_1=1$, $k_2=2$, $n_1=1$, and $n_2=2$,[3] then

[3] These same parameters will be assumed in all of our specific examples used below

$$E\prod_{j=1}^{m}\alpha(i_j,k_j,n)^{n_j}$$
$$=E[\alpha(1,1,n)\,\alpha(3,2,n)^2]$$
$$=\left(\frac{1}{\sqrt{n}}\right)^5\sum E\{w_{1l_1^1}\,w_{3l_1^2}\,w_{l_1^2l_2^2}\,w_{3l_1^3}\,w_{l_1^3l_2^3}\}$$

where the latter summation is over l_1^1, l_1^2, l_2^2, l_1^3, and l_2^3. Here,

$$w_{15}\,w_{34}\,w_{42}\,w_{31}\,w_{18}$$

is an example of a chain sequence, consisting of the chains w_{15}, $w_{34}\,w_{42}$, and $w_{31}\,w_{18}$. The first element is w_{15}, the second is w_{34}, etc.

It is clear that a chain sequences will give rise to a nonzero contribution in (a.2) only if every element is paired to (identical to) at least one other element. In these chain sequences, we will distinguish two types of elements:

1. Call an element a *First* if it is not repeated to its left.
2. All other elements are *Seconds*.

Thus, for example,

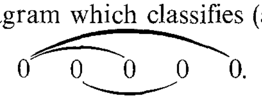
$$\begin{matrix} w_{15} & w_{31} & w_{15} & w_{31} & w_{15} \\ F & F & S & S & S \end{matrix} \tag{a.3}$$

is a chain sequence whose expectation is not necessarily zero, and may therefore contribute to (a.2). The First and Second elements have been indicated by "F's" and "S's" respectively.

Every chain sequence in which each element is paired to at least one other element can be uniquely classified according to its *pairing diagram*:

For every Second, draw an arc which connects the position of that Second to the position of the (unique) First to which it is identical.

For example, the pairing diagram which classifies (a.3) is


(a.4)

Notice that, with fixed $i_1,\ldots,i_m$, $k_1,\ldots,k_m$, and $n_1,\ldots,n_m$, there are only a finite number of pairing diagrams, and that this number does not change with n. *Pairing class* will refer to the set of all chain sequences with a given pairing diagram.

We wish to count the number of chain sequences in a given pairing class (this number *does* depend on n). In particular, we wish to identify those pairing classes which are large enough to contribute asymptotically to the sum in (a.2).

Before proceeding with these combinatorics in the general case, it will be illustrative to look more closely at the case defined by the parameters in our particular example. It is not hard to see that the pairing diagram, (a.4), is the *only* pairing diagram associated with this example; *every* chain sequence with nonzero expectation belongs to the pairing class associated with (a.4). Taking into account the restrictions on indices imposed by this pairing diagram, we have the following generic representation for a member of this pairing class:

$$w_{1l_1^1}\,w_{31}\,w_{1l_1^1}\,w_{31}\,w_{1l_1^1}. \tag{a.5}$$

Since l_1^1 is the only free index in this representation, it is evident that for this example there are exactly n elements in the pairing class. And, since the expectation of every element in a pairing class is identical, there is a constant c such that

$$E \prod_{j=1}^{m} \alpha(i_j, k_j, n)^{n_j} = \left(\frac{1}{\sqrt{n}}\right)^5 c\, n \to 0$$

as $n \to \infty$, which is consistent with (a.1).

Similarly, for the purpose of counting the number of members in a pairing class in the general case, we will derive a generic representation, such as (a.5), for the members of a given class. The numbers of free indices in this generic representation determines the size of the pairing class. A generic representation for members of a pairing class can be derived from the pairing diagram by a relabeling process, as follows:

Begin with a generic representation for a chain sequence, in which all of the indices, except the left most index of each chain, are free (i.e. the analogue of

$$w_{1 l_1^1} w_{3 l_1^2} w_{l_1^2 l_2^2} w_{3 l_1^3} w_{l_1^3 l_2^3}$$

for the general case). Fix a pairing class. When referring to a generic chain sequence, First and Second will mean those elements which are Firsts and Seconds in the members of this pairing class.

Now, beginning with the left-most First (which is also the first element of the chain sequence), relabel the indices of each Second connected to this First – so that the indices of each Second are identical to those of the First. However, one or more of the Seconds may contain a non-free first index (by being the first member of a chain). In this case, the First, and all connected Seconds, must assume this non-free index as first index. The relabeling of indices is next extended to neighbors of elements which share a relabeled index.

Before relabeling, a free index can be assigned an "order", depending on its position in the chain sequence. We will say that one index is of higher order than another if it originally appears to the right of the other index. As relabeling proceeds, "order" will continue to refer to the *original* position of an index.

Now choose the next most left First. Beginning with its left most Second, relabel indices to reflect the fact that the First and the Second are the same element. Given a match of two free indices, always relabel the index of higher order, replacing it by the lower order index in all appearances in the chain sequence. (Note that the Second may have the lower order index, having inherited this from a relabeled neighboring element.) If an index is fixed, then, as before, the matching index (in all its appearances) must be relabeled.

Continue this procedure by choosing successive Firsts.

In summary, the following relabeling procedure is to be carried out for each pairing class:

1. Begin with the left-most First and proceed to the right.
2. For each First, begin with the left-most Second and proceed to the right.

3. Relabel indices to reflect the matching of a Second to its First (as defined by the pairing diagram).

a) Any index matched to a constant is made constant.

b) If neither of the matched indices is constant, then relabel the index of higher order

c) Any time an index is relabeled, relabel all occurrences of that index in the chain sequence.

As an illustration, let us carry out the relabeling procedure for the single pairing class in our specific example:

$$\begin{aligned} & w_{1l_1^1} w_{3l_1^2} w_{l_1^2 l_2^2} w_{3l_1^3} w_{l_1^3 l_2^3} \\ & \rightarrow w_{1l_1^1} w_{3l_1^2} w_{1l_1^1} w_{3l_1^3} w_{l_1^3 l_2^3} \\ & \rightarrow w_{1l_1^1} w_{31} w_{1l_1^1} w_{3l_1^3} w_{l_1^3 l_2^3} \\ & \rightarrow w_{1l_1^1} w_{31} w_{1l_1^1} w_{3l_1^3} w_{1l_1^1} \\ & \rightarrow w_{1l_1^1} w_{31} w_{1l_1^1} w_{31} w_{1l_1^1}. \end{aligned}$$

We claim that all free indices originally belonging to Second elements no longer appear in the chain sequence after the relabeling process. To see this, first observe that the order of the index at a given location is never increased. If a free index of a Second element is unchanged at the time at which that element is matched with its First, then the corresponding index of the First, being to the left, must be of lower order (the index at that location may have been changed, but not to an index of higher order). Hence, the free index of the Second will be lost upon relabeling. On the other hand, if a free index of a Second was changed before matching, then, since all occurances of that index were changed, it is already lost from the chain sequence.

Define a run (or "run of Seconds") to be a maximal sequence of consecutive Seconds *within a chain.* A run with k elements has, before relabeling,

k free indices if it contains the first element of the chain, and

$k+1$ free indices if it does not contain the first element of the chain.

If $\sum_{j=1}^{m} n_j k_j$ is odd, then for each pairing class there are at least $\frac{1}{2} \sum_{j=1}^{m} n_j k_j + \frac{1}{2}$ Seconds, which will "cost" at least as many free indices when paired. Hence there will be no more than $\frac{1}{2} \sum_{j=1}^{m} n_j k_j - \frac{1}{2}$ free indices after pairing and, therefore, no more than

$$n^{\frac{1}{2} \sum_{j=1}^{m} n_j k_j - \frac{1}{2}}$$

terms in each pairing class. Since the summation in (a.2) is multiplied by

$$\left(\frac{1}{\sqrt{n}}\right)^{\sum_{j=1}^{m} n_j k_j}$$

and since the expectation of every chain sequence in a pairing class is identical and finite, the contribution from such pairing classes goes to zero as $n \rightarrow \infty$.

If $\sum_{j=1}^{m} n_j k_j$ is even, then there are at least $\frac{1}{2} \sum_{j=1}^{m} n_j k_j$ Seconds; the same reasoning shows that the only pairing classes which make an asymptotic contribution are those with *exactly* $\frac{1}{2} \sum_{j=1}^{m} n_j k_j$ Seconds. Furthermore, each run of these Seconds must contain the first element of its chain (so as not to lose an additional free index).

In summary, the only pairing classes which have enough elements to contribute to the limit are those for which

1. $\sum_{j=1}^{m} n_j k_j$ is even,
2. there are $\frac{1}{2} \sum_{j=1}^{m} n_j k_j$ Seconds, and
3. each runs of Seconds begins with the first element of a chain.

Next we argue that, among the classes defined by 1, 2, and 3 above, only those for which every run of *Firsts* begins with the first element of a chain are relevant in the limit. To see this, follow the recipe for relabeling indices, but resolve matches by changing the index of *lower* order. Reasoning as before, conclude that relabeling leads to a loss of free indices equal to

k for each run of Firsts containing k elements and the first element of a chain, and

$k+1$ for each run of Firsts containing k elements but not containing the first element of a chain.

Since there are $\frac{1}{2} \sum_{j=1}^{m} n_j k_j$ Firsts in the relevant pairing classes, these contain only runs of Firsts which begin with the first element of a chain.

Putting together the previous 2 paragraphs, we conclude that the relevant pairing classes:

1. contain $\frac{1}{2} \sum_{j=1}^{m} n_j k_j$ Firsts, each matched to exactly one Second, and
2. have the property that every run, of Firsts or Seconds, is a chain.

Now, of these, we need only consider those classes for which, under the original relabeling scheme, no index of a First is changed. In all other classes, the number of free indices will be less than $\frac{1}{2} \sum_{j=1}^{m} n_j k_j$ (this being the number of indices associated with Firsts, in the relevant classes, before relabeling).

In the relevant pairing classes, it must be the case that two consecutive Seconds in the same chain are paired, preserving order, to consecutive Firsts of a common chain. Otherwise, one of the two Firsts would have an altered index reflecting the shared index of the consecutive Seconds. Consider a run of Seconds (which, in the relevant pairing classes, is a full chain). The first element must be paired to the first element of another chain, in order that no index of a First be altered. (This, of course, requires that these first elements have the same fixed first index.) Each successive Second is necessarily paired to each successive First, until there are no more Seconds in the chain of Seconds. If there is another First in the chain of Firsts, then it must be paired to an element of a different chain, which element can not be the first element of that different chain (due to the fixed index of first chain elements). But this requires that this other chain be a chain of Seconds, and this leads to a violation of the above-mentioned fact that consecutive Seconds of a common chain pair to consecutive Firsts of a common chain, preserving order.

In short, in the asymptotically relevant pairing classes, every chain of Firsts must be matched to (exactly) one chain of Seconds of identical length, and identical first (fixed) index. The expectation of any chain sequence of this type is exactly

$$\prod_{j=1}^{m} (\sigma^2)^{k_j \frac{n_j}{2}} \quad \text{(all } n_j\text{'s must be even).}$$

Furthermore, the generic member of such a pairing class has $\frac{1}{2}\sum_{j=1}^{m} n_j k_j$ free indices, on which the only restrictions are those which assure that no further pairings are created, beyond those defined by the pairing diagram. It then follows that there are between

$$\frac{1}{n^2}\sum_{j=1}^{m} n_j k_j \quad \text{and} \quad \frac{n!}{\left(n-\frac{1}{2}\sum_{j=1}^{m} n_j k_j\right)!}$$

members of each of the relevant pairing classes. And therefore, the asymptotic contribution to (a.2) from each of the relevant pairing classes is

$$\lim_{n\to\infty}\left(\frac{1}{\sqrt{n}}\right)^{\sum_{j=1}^{m} n_j k_j} \prod_{j=1}^{m} (\sigma^2)^{k_j \frac{n_j}{2}} \quad \{\#\text{ members}\}$$
$$= \prod_{j=1}^{m} (\sigma^2)^{k_j \frac{n_j}{2}}. \tag{a.6}$$

We have shown that the relevant pairing classes are those in which each chain is matched to exactly one other chain, the latter having identical values for the parameters i_j and k_j (notation as in (a.1) and (a.2)). Since, for each j, there are n_j such chains, there are

$$(n_j-1)(n_j-3)\dots(1)=\prod_{p=1}^{n_j/2}(2p-1)$$

such pairings for each set of parameters i_j and k_j. And, since there are m sets of these parameters ($1 \leqq j \leqq m$), there are

$$\prod_{j=1}^{m}\prod_{p=1}^{n_j/2}(2p-1) \tag{a.7}$$

asymptotically relevant pairing classes. Combining (a.6) and (a.7), and comparing to (a.1), completes the proof. □

Acknowledgement. Before proofs were complete, computer simulation was used to accumulate evidence for the validity of some of these conjectures. Ulf Grenander taught us the utility of this approach. And, it was he who suggested the moments method to prove Lemma A.1.

References

1. Billingsley, P.: Convergence of probability measures. New York: Wiley 1968
2. Chen, L.H.Y.: Two central limit problems for dependent random variables. Z. Wahrscheinlichkeitstheorie verw. Gebiete **43**, 223–243 (1978)
4. Chernoff, H.: A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. Annals Math. Statist. Vol. **23**, 493–507 (1952)
4. Chung, K.L.: A course in probability theory. New York: Academic Press 1974
5. Geman, S.: A limit theorem for the norm of random matrices. Annals Probab. **8**, 252–261 (1980)
6. Geman, S.: Almost sure stable oscillations in a large system of randomly coupled equations. SIAM J. Applied Math. (to appear)
7. Olver, F.W.: Bessel functions of integer order. In: Handbook of Mathematical Functions. Ed. M. Abramowitz and I.A. Stegun. New York: Dover 1965

Received August 4, 1980; in revised form October 26, 1981

Z. Wahrscheinlichkeitstheorie verw. Gebiete
60, 315–334 (1982)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

Asymptotic Theory of Grenander's Mode Estimator

Peter Hall

Australian National University, Dept. of Statistics, P.O. Box 4, Canberra ACT 2600, Australia

Summary. The simplicity and good performance of Grenander's mode estimator have found favour with several authors; see for example [6, 11]. However, mathematical difficulties have precluded a proper examination of its theoretical properties [3], and most of the available studies concentrate on Monte Carlo simulations [2–6, 8, 11]. In this paper we give a rigorous, theoretical account of the estimator's properties. In particular, we derive a central limit theorem which describes the influence of the two adjustable parameters on the behaviour of the estimator.

1. Introduction and Summary

Grenander [12] suggested a one-stage mode estimator which provided a practical alternative to Parzen's [13] two-stage, density-based procedure. The simplicity and excellent performance of Grenander's estimator have found favour with several authors. For example, after a comparative study of mode estimators, Ekblom [11] concluded that "on the whole, Grenander's direct mode estimator is the one to be preferred of the four types examined". Adriano, Gentle and Sposito [6] recommended Grenander's estimator, together with a variant of an estimator proposed by Venter [17] and Sager [16], over the other estimators they studied. Dalenius [8] compared three different estimators, and showed that Grenander's estimator has smaller standard error than the other two.

These favourable conclusions have generated interest in Grenander's method, particularly during recent years. However, the majority of studies (see for example [2–6, 8, 11]) have consisted of Monte Carlo experiments, with very little attention being paid to theoretical results. There does not even exist a description of central limit theory for Grenander's estimator, since "the analytic complexity of the estimator makes a mathematical study ... quite difficult" [3].

0044-3719/82/0060/0315/$04.00