Chaos in Random Neural Networks

H. Sompolinsky[(a)] and A. Crisanti
AT&T Bell Laboratories, Murray Hill, New Jersey 07974, and
Racah Institute of Physics, The Hebrew University, 91904 Jerusalem, Israel[(b)]

and

H. J. Sommers[(a)]
Fachbereich Physik, Universität-Gesamthochschule Essen, D-4300 Essen, Federal Republic of Germany
(Received 30 March 1988)

A continuous-time dynamic model of a network of N nonlinear elements interacting via random asymmetric couplings is studied. A self-consistent mean-field theory, exact in the $N\to\infty$ limit, predicts a transition from a stationary phase to a chaotic phase occurring at a critical value of the gain parameter. The autocorrelations of the chaotic flow as well as the maximal Lyapunov exponent are calculated.

PACS numbers: 05.45.+b, 05.20.−y, 47.20.Tg, 87.10.+e

Theoretical investigations of the onset and the nature of chaotic flows in deterministic dynamical systems have focused, in recent years, mainly on systems with few degrees of freedom.[1] Quite often chaos is achieved in these systems by the variation of a parameter through a sequence of bifurcation points, which represent increasing complexity of the motion. It is still an open question whether these scenarios are realized in large systems which cannot be described by a small number of collective modes. In this Letter we study the nature of chaotic flows in a deterministic nonlinear system—a neural network—consisting of many localized degrees of freedom. We show that in the limit of infinite number of degrees of freedom, there is a sharp transition (as a function of the parameter of nonlinearity) from a stationary state to a chaotic flow. This transition as well as the statistical properties of the chaotic flow is described by time-dependent self-consistent mean-field theory. The maximal Lyapunov exponent is derived from the fluctuations about the mean-field solutions. In the context of neurobiology, the study of chaos in neural networks may be relevant to the understanding of the appearance of spontaneous irregular patterns of activity in neural assemblies.[2]

The model consists of N localized continuous variables ("neurons") $\{S_i(t)\}$, $i=1,\ldots,N$, where $-1\le S_i\le 1$. Associated with each neuron, a local field h_i, $-\infty<h_i<+\infty$, is defined through the relationship $S_i(t)=\phi(h_i(t))$ where $\phi(x)$ is a nonlinear gain function which defines the input (h_i) -output (S_i) characteristics of the neurons. In the biological context, h_i may be related to the membrane potential of the nerve cell and S_i to its electrical activity (e.g., its firing rate). The function $\phi(x)$ is assumed to have a sigmoid shape $\phi(\pm\infty)=\pm 1$, $\phi(-x)=-\phi(x)$. For concreteness we choose the function

$$\phi(x)=\tanh(gx), \tag{1}$$

where the *constant* $g>0$ measures the degree of nonlinearity of the neural response. The dynamics of the network is given by N coupled first-order differential equations ("circuit" equations)[3,4]

$$\dot{h}_i=-h_i+\sum_{j=1}^{N}J_{ij}S_j=-h_i+\sum_{j=1}^{N}J_{ij}\phi(h_j). \tag{2}$$

Here J_{ij} is the synaptic efficacy which couples the output of the (presynaptic) jth neuron to the input of the (postsynaptic) ith neuron, and $J_{ii}=0$. In electrical terms Eqs. (2) are Kirchhoff equations in which the left-hand side represents the current leakage due to the membrane capacitance; the first term in the right-hand side represents the current through the membrane resistance and the last term denotes the input current flowing to the cell due to the activity of the other cells. For simplicity the microscopic time constant has been set equal to unity. We consider here networks with *random* synaptic couplings. Each of the J_{ij}'s is an independent random variable which can be assumed for convenience to have a Gaussian distribution. The mean of J_{ij} is zero whereas the variance is $[J_{ij}^2]_J=J^2/N$. With this normalization the intensive parameter in the case of (1) is the dimensionless *gain parameter gJ*.

If the synaptic matrix **J** were symmetric then Eqs. (2) would describe a relaxation of a global energy function,[5] which for random couplings is just a spin-glass Hamiltonian.[6] Here we study synaptic matrices where J_{ij} and J_{ji} are uncorrelated in which case the dynamics in general is nonrelaxational. In this case the long-time behavior may depend on the particular realization of the J_{ij}'s. However, in the limit $N\to\infty$ a well defined typical behavior exists, the properties of which are described below.

The long-time properties of the solutions of Eqs. (2) for large N have been studied by the method of the dynamical mean-field theory (MFT) originally developed for spin-glasses.[7] This theory is *exact* in the limit of $N\to\infty$. The essential result of this MFT is quite simple. The dynamics of the system at long times can be reduced to a self-consistent equation of a single neuron, which reads

$$\dot{h}_i(t) = -h_i(t) + \eta_i(t). \quad (3)$$

The term η_i is a time-dependent Gaussian field which is generated by the random inputs from the other neurons, i.e., the last term in Eqs. (2). Obviously, the mean of η_i is zero. Its second moment is determined self-consistently from Eq. (2), yielding

$$\langle \eta_i(t)\eta_i(t+\tau)\rangle = J^2 C(\tau), \quad (4)$$

where the average autocorrelation function

$$C(\tau) = [N^{-1}\textstyle\sum_i S_i(t)S_i(t+\tau)]_J$$

is evaluated with use of Eqs. (1) and (3),

$$C(\tau) = \langle \phi(h_i(t))\phi(h_i(t+\tau))\rangle$$

and

$$h_i(t) = \int_{-\infty}^{t} dt' e^{t'-t}\eta_i(t').$$

Square brackets with a subscript J denote average over the distribution of **J**. The angular brackets denote averaging with respect to the Gaussian distribution of η_i. Thus Eq. (4) represents a self-consistent equation for the time-dependent autocorrelation $C(t)$. Note that in deriving Eq. (4) we have assumed that the system reached a satisfactory steady state so that correlations depend only on *time separations* and $C(t) = C(-t)$.[8]

Instead of our solving the equation for $C(t)$ it is more convenient to study explicitly the local-field autocorrelation

$$\Delta(\tau) = \langle h_i(t)h_i(t+\tau)\rangle, \quad (5)$$

which by Eqs. (3) and (4) obeys $\Delta - \ddot{\Delta} = J^2 C$. Equations (4) and (5) can be reduced to the following equation for Δ:

$$\ddot{\Delta} = -\partial V/\partial\Delta, \quad (6)$$

$$V(\Delta) = -\tfrac{1}{2}\Delta^2 + \int_{-\infty}^{+\infty} Dz \left(\int_{-\infty}^{+\infty} Dx\, \Phi((\Delta(0)-|\Delta|)^{1/2}x + |\Delta|^{1/2}z)\right)^2, \quad (7)$$

where $Dz = dz\exp(-z^2/2)/(2\pi)^{1/2}$ and, in general, $\Phi(x) = \int_0^x dy\,\phi(y)$. In the particular example of (1) $\Phi(x) = (gJ)^{-1}\ln\cosh(gJx)$. Equation (6) can be viewed as a one-dimensional motion of $\Delta(t)$ under the Newtonian potential $V(\Delta)$. There are important boundary conditions for the solutions of Eq. (6): (i) $\Delta(t)$ is a differentiable even function, i.e., $\Delta(-t) = \Delta(t)$, $\dot{\Delta}(0) = 0$. This implies that the orbit must have zero initial kinetic energy. (ii) $\Delta(t)$ is bounded by $|\Delta(t)| \le \Delta(0)$. Note that the potential V is a *self-consistent* potential. It depends parametrically on $\Delta(0)$, the value of which has to be consistent with Eq. (6). We now study the solutions of Eqs. (6) and (7).

$gJ < 1$, zero fixed point.— For a gain parameter less than unity, $V(\Delta)$ is of the form shown in Fig. 1(a), in the allowed regime $|\Delta(t)| \le \Delta(0)$. The only *bounded* classical orbit with $\dot{\Delta}(0) = 0$ is $\Delta(t) \equiv 0$. The vanishing of the ("steady state") equal-time correlations implies that the system (2) flows to the zero fixed point $\{h_i \equiv 0\}$, for all (or almost all) initial conditions. The stability of the zero fixed point below $gJ = 1$ can be deduced by our linearizing the equations of motion (2), noting that the maximal real part of the eigenvalues of **J** is 1.[9]

$gJ > 1$, chaotic phase.— For $gJ > 1$ the form of $V(\Delta)$ is not unique but depends on the assumed magnitude of $\Delta(0)$. Furthermore, for a given $V(\Delta)$ there is a continuum of self-consistent solutions of Eq. (6). In general there exists a value $\Delta_1(gJ)$ such that for $\Delta(0)$ in the regime $0 < \Delta(0) < \Delta_1$, $V(\Delta)$ has the form shown in Fig. 1(b). For such a potential the solutions of Eq. (6) are periodic orbits, implying that the system converges to *limited cycles.* If $\Delta(0) > \Delta_1$ the potential has the double-well form of Fig. 1(c). In this case, the solution with the lowest (Newtonian) energy [denoted by a in Fig. 1(c)] has $\Delta(t) \equiv \Delta$, where $\Delta > 0$ is independent of time. This static solution corresponds to a nonzero fixed point of the system (2), and would be analogous to a spin-glass freezing which occurs in systems with random symmetric **J**.[6] Other solutions with negative energies (denoted by b) correspond to limited cycles with nonzero averages. Solutions with positive energies (denoted by c) are oscillations around zero values.

Which of the many solutions correspond to stable attractors? In fact, a study of the stability of the MFT to fluctuations, which will be outlined below, shows that *none* of the above solutions (for $gJ > 1$) *are stable.* The only stable solution is the one with zero energy, denoted by d in Fig. 1(c), for which $\Delta(t)$ *decreases monotonically to zero* as $t \to \infty$. The decay of the correlations between two points along the flow implies that the flow is chaotic. For $gJ \simeq 1^+$, the chaotic solution for $\Delta(t)$ is $\Delta(t) \simeq \epsilon\cosh^{-2}(\epsilon t/\sqrt{3})$, where $\epsilon \equiv gJ - 1 \ll 1$. Note that as $gJ \to 1^+$ the amplitude of the flow vanishes as $\Delta(0) \sim \epsilon$. At the same time the relaxation time of the autocorrelation diverges as $\tau \propto 1/\epsilon$. In general, this relaxation time is given by $\tau^{-1} = [-\partial^2 V(0)/\partial\Delta^2]^{1/2}$ as can be deduced from Eq. (6). In the "Ising" limit,[10] $gJ \to \infty$, the relaxation time approaches the value $\tau^{-1} \to (1-2/\pi)^{1/2}$.

Fluctuations and Lyapunov exponent.— In order to study the stability of the flows to fluctuations we add to Eq. (2) an infinitesimal external source, $h_i^0(t)$, and study the linear perturbation of the flow, given by $\chi_{ij}(t,t')$

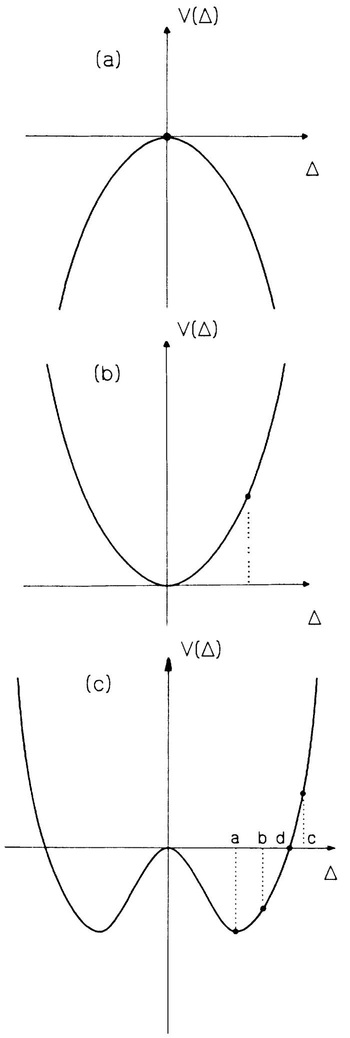

FIG. 1. Qualitative shape of the self-consistent classical potential $V(\Delta)$ in the range $|\Delta| \leq \Delta(0)$. (a) The case of $gJ < 1$. (b) $gJ > 1$ and $0 < \Delta < \Delta_1$. (c) $gJ > 1$ and $\Delta > \Delta_1$ (see text). A solution starting at the point marked a is a static state. The points b and c are examples of initial conditions leading to solutions which oscillate around nonzero and zero values, respectively. The orbit starting at d decays to zero. The dots mark the energies of the various solutions.

$\equiv \delta h_i(t)/\delta h_j^0(t')$. It is more convenient to study the average quantity $\chi^2(t)$ defined by

$$\chi^2(t) \equiv \lim_{\tau\to\infty} \frac{1}{N} \left[\sum_{ij} \chi_{ij}^2(\tau+t,\tau) \right]_J, \quad t>0, \tag{8}$$

which yields information regarding the maximal Lyapunov exponent λ. This exponent, which measures the sensitivity of the flow to a perturbation of the initial conditions in a random direction, is given by the *asymptotic* time dependence $\chi^2(t)$, i.e.,[11]

$$\lambda = \lim_{t\to\infty} \ln(\chi^2(t))/2t. \tag{9}$$

In the case of a stable point, λ is negative. If the attractor is a stable limit cycle or quasiperiodic, λ should be zero. On the other hand, in the case of a chaotic attractor λ is positive indicating the exponential *growth* of a perturbation of the initial condition in random directions.

Calculating the fluctuations about the MFT we find that $\chi^2(t)$ can be written as

$$\chi^2(t) = \sum_{n=0}^{\infty} \chi_n \exp(2\omega_n t),$$

$t > 0$, with $\omega_n = -1 \pm (1-E_n)^{1/2}$, where E_n are the eigenvalues of the one-dimensional Schrödinger equation,

$$(-\partial^2/\partial t^2 - \partial^2 V/\partial \Delta^2)\psi_n(t) = E_n \psi_n(t), \tag{10}$$

and $E_{n+1} \geq E_n$. From Eq. (9) it follows that

$$\lambda = -1 + (1-E_0)^{1/2}, \tag{11}$$

where E_0 is the ground-state energy of Eq. (10). The "fluctuation potential" $W(t) \equiv -\partial^2 V/\partial\Delta^2$ is determined by the classical solution, $\Delta(t)$, about which perturbations are made. When Δ is constant, $W(t)$ is also constant, $W(t) \equiv W$, and the spectrum of E_n is the continuum $E \geq W$ yielding $\lambda = -1 + (1-W)^{1/2}$. In the case of the zero fixed point, $W = 1 - g^2J^2$; hence $\lambda = gJ - 1$ indicating the instability of this fixed point when $gJ > 1$. The static "spin-glass" fixed point also has a negative W, i.e., a positive λ (for all $gJ > 1$), implying that it is not a stable fixed point.

The analysis of Eq. (10) for a time-dependent $\Delta(t)$ is more complicated. However, by differentiation of Eq. (6) it is readily seen that $\psi(t) = \dot{\Delta}(t)$ is an eigenstate of Eq. (10) with $E = 0$. This state represents a fluctuation of the initial condition in the direction of the flow. In the case of oscillatory $\Delta(t)$, the zero-energy eigenstate is part of a band of eigenvalues extending to *negative* energies, leading again to a *positive* Lyapunov exponent.[12]

In the case of the chaotic solution, $W(t)$ is a symmetric potential well with a single minimum at $t = 0$. The zero-energy solution is a bound state with an odd wave function, $\psi(t) = \dot{\Delta}(t)$, which has a single node, $\psi(0) = 0$. Thus there is exactly one bound state, the ground state, with *negative* energy yielding a positive λ as expected for a chaotic flow. We thus conclude that *the chaotic solution is the only stable solution of the MFT for* $gJ > 1$.[12] Solving Eq. (10) for the chaotic solution near $gJ \simeq 1$ one finds that $\lambda \sim \epsilon^2/2$, i.e., it vanishes continuously as $\epsilon \equiv gJ - 1 \to 0$. In the Ising limit, $gJ \to \infty$, λ diverges as $\ln(gJ)$, similar to the result for products of random matrices.[13]

An important question is how many positive Lyapunov exponents appear above $gJ = 1$. The distribution of the Lyapunov exponents cannot be evaluated from $\chi^2(t)$. However, the fact that the exponential divergence of $\chi^2(t)$ with time is observed in the theory even after the limit $N \to \infty$ has been taken indicates that the number of positive Lyapunov exponents for $gJ > 1$ is macroscopic, i.e., proportional to N.

To what extent do the above predictions correspond to the typical behavior of a system with a finite, large value of N? To check this point we have solved Eqs. (2) numerically for systems with sizes up to $N = 1000$. A detailed account of the simulations will be given elsewhere. The most important novel feature of the numerical results is the existence of an intermediate regime of gJ separating the stationary and the chaotic phases. For sizes $N \geq 100$, in almost all cases, a rapid decay to a zero fixed point is observed for $gJ > 1$, whereas for $gJ > 2$ the behavior is chaotic. However, as gJ is increased above unity one first observes the appearance of either nonzero stationary states, or more often, *limited cycles*. These limited cycles become increasingly more complex as gJ is increased until the motion becomes chaotic. The limit cycles vary strongly from one realization of **J** to another and are therefore probably not related directly to the unstable periodic solutions of the MFT. In fact, comparing the behavior of different system sizes we find that the range of gJ where this intermediate behavior is observed shrinks as N increases. It therefore seems that a system with a finite number of degrees of freedom N undergoes a transition from a stationary state to chaos through intermediate stages of limit cycles with increasing complexity. However, the range of gJ where these bifurcations occur shrinks to zero with N so that in the limit $N \rightarrow \infty$ a *sharp* transition to chaos emerges as predicted by the MFT.

Several intriguing questions remain to be answered. First, the above proposed picture for the behavior of large finite systems has yet to be systematically tested. In addition, further characterization of the strange attractor, e.g., evaluation of the whole distribution of Lyapunov exponents, is needed. Finally, it will be very interesting to study extensions of the present model to networks with *short-range* synaptic matrices in finite spatial dimensions. Detailed presentation of the calculations will be given elsewhere.

Special thanks to P. C. Hohenberg for most valuable discussions. We have also benefitted from discussions with D. Ben Simon, B. Derrida, D. S. Fisher, D. Kleinfeld, P. B. Littlewood, I. Procaccia, B. Shraiman, and A. Vulpiani. This research is partially supported by a grant of the U.S.A.-Israel Binational Science Foundation. One of us (H.J.S.) acknowledges support from the Deutsche Forschungsgemeinschaft, Sonderforschungsbereich No. 237.

[a]Permanent address.

[b]Present address: Institute for Advanced Studies, Hebrew University, 91904 Jerusalem, Israel.

[1]For a review on the theory of deterministic chaos see, e.g., J.-P. Eckmann and D. Ruelle, Rev. Mod. Phys. **57**, 617 (1985); G. Paladin and A. Vulpiani, Phys. Rep. **156**, 147 (1987).

[2]Most previous studies of these phenomena employ neural models with few degrees of freedom. See, e.g., M. R. Guevara *et al.*, IEEE Trans. Syst. Man Cybern. **13**, 790 (1983).

[3]This model is similar to that studied by S. I. Amari, IEEE Trans. Syst. Man Cybern. **2**, 643 (1972), who, however, did not investigate self-consistent *time-dependent* solutions.

[4]Collective complex dynamical behavior which is specific to *discrete-time* systems has been studied by M. Y. Choi and B. A. Huberman, Phys. Rev. B **28**, 254 (1983).

[5]For recent study of networks with an energy function see, e.g., J. J. Hopfield and D. W. Tank, Science **233**, 625 (1986), and references therein.

[6]For a review on spin-glass theory see, e.g., K. Binder and A. P. Young, Rev. Mod. Phys. **58**, 801 (1986).

[7]H. Sompolinsky and A. Zippelius, Phys. Rev. B **25**, 6860 (1982); A. Crisanti and H. Sompolinsky, Phys. Rev. A **36**, 4922 (1987).

[8]Usually the autocorrelation is defined by averaging over initial conditions. However, after averaging over J_{ij}'s, the steady-state limit of C is dependent of the initial conditions.

[9]H. J. Sommers, A. Crisanti, H. Sompolinsky, and Y. Stein, Phys. Rev. Lett. **60**, 1895 (1988).

[10]A. Crisanti and H. Sompolinsky, Phys. Rev. A **37**, 4865 (1988).

[11]In principle the average over **J** should be performed on $\ln\chi^2$; however, in the present case χ^2 is self-averaging so that the averaging and the logarithm can be interchanged.

[12]Formally we find that the MFT solutions are stable only if there is a gap in the Spectrum of E_n below $E = 0$. The unstable periodic solutions presumably correspond to the set of unstable limit cycles which reside on the chaotic attractor. See D. Ruelle, *Thermodynamic Formalism* (Addison-Wesley, Reading, MA, 1978); and D. Auerbach *et al.*, Phys. Rev. Lett. **58**, 2387 (1987).

[13]J. E. Cohen and C. M. Newman, Ann. Probab. **12**, 283 (1984); C. M. Newman, Commun. Math. Phys. **103**, 121 (1986); A. Crisanti, G. Paladin, and A. Vulpiani, J. Stat. Phys. (to be published).