Nonequilibrium statistical-mechanical approach to discrete-time dynamics

T. Munakata
Department of Applied Mathematics and Physics, Kyoto University, Kyoto 606, Japan
(Received 6 May 1994)

A nonequilibrium statistical-mechanical approach to a general discrete-time dynamical system is presented. A generalized Langevin equation (GLE) with an alternative fluctuation-dissipation theorem is derived for a system with a stationary (invariant) distribution function. A linear response theory is also formulated within a similar framework. The results obtained here are not limited to map dynamics. This is illustrated by applying the GLE to a Hopfield neural network which is synchronously updated following Glauber dynamics.

PACS number(s): 05.20.−y, 02.50.Ey, 05.40.+j

Discrete-time dynamics, in which the state of a system changes only at some prescribed instant of time, has been playing an important role in studies of dynamical behavior in many branches of natural science. Two types of discrete-time dynamics are conceivable. When time evolution of (physical) variables is considered, one usually employs a map to describe system dynamics [1]. The logistic map, modeling the yearly variation of an insect species, is the simplest example. On the other hand, if time evolution of the distribution function is of interest, a discrete-time master equation is used as in the case of neural networks which are updated synchronously [2].

For a continuous-time dynamical system, whose dynamics is governed by, e.g., Hamilton's equation of motion (or a set of Langevin equation), statistical dynamical approaches to irreversible processes, such as a general theory of Brownian motion (GTBM) [3] and a linear response theory (LRT) [4], are formulated with the aid of the stationary distribution function of a Liouville (or a Fokker-Planck) equation. The purpose of this paper is to develop a statistical dynamical theory for a general discrete-time system, which is assumed to have an invariant (stationary) distribution.

Let us first introduce a deterministic (non-noisy) map

$$\mathbf{X}_{n+1}=\mathbf{G}(\mathbf{X}_n)\ , \tag{1}$$

with $\mathbf{X}_n$ denoting the state vector at time $t=n$ ($n=0,1,..,$). Time evolution of the distribution function $P(\mathbf{X};n)$ is governed by [5]

$$P(\mathbf{X};n+1)=\int d\mathbf{X}'P(\mathbf{X}';n)\delta(\mathbf{X}-\mathbf{G}(\mathbf{X}'))=L^{\dagger}P(\mathbf{X};n)\ . \tag{2}$$

We define an innerproduct (F,H) of two arbitrary variables F and H by $(F,H)\equiv\int d\mathbf{X}F(\mathbf{X})H(\mathbf{X})$ and the operator L, adjoint to $L^{\dagger}$, by

$$(L^{\dagger}F,H)=(F,LH)\ . \tag{3}$$

We readily see from Eq. (2) that

$$LF(\mathbf{X})=F(\mathbf{G}(\mathbf{X}))\ . \tag{4}$$

By calculating the average at time $t=n$, that is,

$$\int d\mathbf{X}\,P(\mathbf{X};n)F(\mathbf{X})=\int d\mathbf{X}\,F(\mathbf{X})(L^{\dagger})^nP(\mathbf{X};0)=\int d\mathbf{X}\,P(\mathbf{X};0)(L)^nF(\mathbf{X})\ , \tag{5}$$

we notice that the L denotes the evolution operator of a dynamical variable. Thus, in the Heisenberg representation, a variable $F(\mathbf{X})$ at time $t=n$ is expressed as

$$F_n(\mathbf{X})=(L)^nF(\mathbf{X})\equiv(L)^nF_0(\mathbf{X})\ . \tag{6}$$

In the following we assume that the system has a stationary distribution $P_{\rm st}(\mathbf{X})$, which satisfies $L^{\dagger}P_{\rm st}=P_{\rm st}$, and express the average of $F(\mathbf{X})H(\mathbf{X})$ over $P_{\rm st}(\mathbf{X})$ by $\langle F,H\rangle$. Finally, the z transformation [6] and the projection operator P_A onto the subspace $\mathbf{A}$ are defined by

$$F_z(\mathbf{X})\equiv\sum_{n=0}^{\infty}F_n(\mathbf{X})z^n\ , \tag{7}$$

$$P_AB(\mathbf{X})\equiv\langle B,\mathbf{A}\rangle\langle\mathbf{A},\mathbf{A}\rangle^{-1}\mathbf{A}(\mathbf{X})\ . \tag{8}$$

With these preparations we now derive a generalized Langevin equation (GLE) for a set of K dynamical variables $\mathbf{A}(\mathbf{X})=\{A(1;\mathbf{X}),\dots,A(K;\mathbf{X})\}$. From Eqs. (6) and (7) we see that

$$(I-zL)\,\mathbf{A}_z(\mathbf{X})=\mathbf{A}(\mathbf{X})\ , \tag{9}$$

which represents the equation of motion in the z space. Putting $Q_A\equiv I-P_A$ we have

$$\mathbf{A}_z(\mathbf{X})=P_A\,\mathbf{A}_z(\mathbf{X})+Q_A\,\mathbf{A}_z(\mathbf{X})\equiv\Xi_z\,\mathbf{A}(\mathbf{X})+\mathbf{A}'_z(\mathbf{X})\ . \tag{10}$$

First, we express $\mathbf{A}'_z(\mathbf{X})$ in terms of the correlation matrix, Ξ_z. This is effected by applying Q_A on Eq. (9), resulting in

$$\mathbf{A}'_z(\mathbf{X})=z[I-zQ_AL]^{-1}\Xi_z\mathbf{f}\ , \tag{11}$$

where the random force at $t=0$, $\mathbf{f}(\mathbf{X})$, is defined by

$$\mathbf{f}\equiv Q_AL\,\mathbf{A}\ . \tag{12}$$

Second, we apply P_A on Eq. (9) and then $,\mathbf{A}\rangle\langle\mathbf{A},\mathbf{A}\rangle^{-1}$ on the resulting equation to have

$$\Xi_z=[I-z\Omega+z^2\Phi_z]^{-1}\ , \tag{13}$$

1063-651X/94/50(3)/2351(4)/$06.00

with the frequency matrix Ω and the memory kernel Φ_A defined by

$$\Omega \equiv \langle L\,\mathbf{A}, \mathbf{A}\rangle\langle \mathbf{A}, \mathbf{A}\rangle^{-1} , \quad (14)$$

$$\Phi_z \equiv -\langle L[I - zQ_A L]^{-1}\mathbf{f}, \mathbf{A}\rangle\langle \mathbf{A}, \mathbf{A}\rangle^{-1} . \quad (15)$$

Finally, the GLE in the z space, Eq. (16), is obtained after insertion of Eqs. (13) and (11) into Eq. (10).

$$[I - z\Omega + z^2\Phi_z]A_z(\mathbf{X}) = \mathbf{A}(\mathbf{X}) + z[I - zQ_A L]^{-1}\mathbf{f} . \quad (16)$$

The inverse z transformation $\mathbf{A}_{n+1} = \int dz\ \mathbf{A}_z z^{-(n+2)}/(2\pi i)$ of Eqs. (15) and (16) yields the GLE and the fluctuation-dissipation theorem (FDT) in the t space,

$$\mathbf{A}_{n+1} = \Omega\,\mathbf{A}_n - \sum_{m=0}^{n-1} \Phi_m \mathbf{A}_{n-m-1} + \mathbf{A}\delta_{n+1,0} + \mathbf{f}_n , \quad (17)$$

$$\Phi_m = -\langle L\,\mathbf{f}_m, \mathbf{A}\rangle\langle A, A\rangle^{-1} . \quad (18)$$

The $\mathbf{f}_m$, the random force at time $t=m$, on the right hand side of Eq. (18) is defined by

$$\mathbf{f}_m \equiv (Q_A L)^m \mathbf{f} . \quad (19)$$

Equations (17), (14), and (18) correspond to Eqs. (3.10), (2.19), and (3.12) of [3a], respectively [7]. It is to be noted, in passing, that since $\mathbf{A}$ is a K-dimensional vector $\mathbf{A}(\mathbf{X}) = (A(1;X), \ldots, A(K;\mathbf{X}))$, the frequency matrix Ω, Eq. (14), the memory kernel Φ_z, Eq. (15), and the correlation matrix Ξ_z are all $K\times K$ matrices. If it happens, as in the case of Eq. (30) below, that the dynamical variables of interest $A(\mathbf{a};\mathbf{X})(\mathbf{a}=\{a_1,..,a_k\}$, $-\infty < a_i < \infty$, for $i=1,\ldots,K)$ have a continuous index, a, those matrices become all $\infty\times\infty$ with $\mathbf{a},\mathbf{b}$ element expressed as, e.g., $\Omega(\mathbf{a},\mathbf{b})$, see Eq. (31).

Here it is worthwhile to comment on some characteristics of map dynamics. From Eqs. (2) to (4) it is seen that $L^\dagger[P_{\rm st}(\mathbf{X})G^n(\mathbf{X})] = P_{\rm st}(\mathbf{X})\mathbf{X}^n$ and thus

$$L^\dagger[P_{\rm st}(\mathbf{X})(L)^n\mathbf{A}(\mathbf{X})] = P_{\rm st}(\mathbf{X})(L)^{n-1}\mathbf{A}(\mathbf{X}) , \quad (20)$$

leading to $\langle \mathbf{A}_{n+m}, \mathbf{A}_n\rangle = \langle \mathbf{A}_m, \mathbf{A}\rangle$. If $\mathbf{A}(\mathbf{X})$ is chosen to satisfy $\langle \mathbf{A}\rangle = \mathbf{0}$, it is confirmed that $\langle \mathbf{f}_m\rangle = \mathbf{0}$ for $m=0,1,\ldots$. Finally, with the aid of the relation $\langle P_A F, H\rangle = \langle F, P_A H\rangle$, it can be shown that

$$\begin{aligned}\Psi_{n+m,n} &\equiv \langle \mathbf{f}_{n+m}, \mathbf{f}_n\rangle\langle \mathbf{A}, \mathbf{A}\rangle^{-1} \\ &= \Psi_{m,0} - \sum_{i=1}^{n} \Phi_{n+m-i}\Phi'_{n-1} \\ &\equiv \Psi_{m,0} - D_{m,n} , \end{aligned} \quad (21)$$

where $\Phi'_n \equiv -\langle L\,\mathbf{f}_n, \mathbf{A}\rangle^T\langle \mathbf{A}, \mathbf{A}\rangle^{-1}$ with Q^T meaning the transpose of the matrix Q. When n becomes large $D_{m,n}$ is expected to approach an n-independent value, e.g., D_m. If the L is Hermitian, $\langle LF, H\rangle = \langle F, LH\rangle$, it holds that $\Psi_{n+m,n} = \Psi_{m,0}$ and the noise $\mathbf{f}_m$ is a stationary process. When the memory kernel Φ_m is a rapidly decaying function of time m, the difference $D_{m,n}$ above is of no practical significance.

We now turn to a formulation of a LRT for map with a small time (n)-dependent perturbation $\mathbf{F}(\mathbf{X}_n, n)$ [8],

$$\mathbf{X}_{n+1} = \mathbf{G}(\mathbf{X}_n) + \varepsilon\mathbf{F}(\mathbf{X}_n, n) . \quad (22)$$

Time evolution of the distribution function $P(\mathbf{X}; n+1)$ is modified to order ε as

$$P(\mathbf{X}; n+1) = L^\dagger P(\mathbf{X}; n) - \varepsilon\nabla\cdot L^\dagger[P(\mathbf{X}; n)F(\mathbf{X}, n)] , \quad (23)$$

where ∇ denotes the gradient operator in the $\mathbf{X}$ space. At time $t=0$ we assume the system is in a stationary state, $P(\mathbf{X};0) = P_{\rm st}(\mathbf{X})$. By solving Eq. (23) iteratively we find that

$$P(\mathbf{X}; n) = P_{\rm st}(\mathbf{X}) - \varepsilon \sum_{m=0}^{n-1} (L^+)^{n-m-1}\nabla\cdot L^\dagger \times [P_{\rm st}(\mathbf{X})\mathbf{F}(\mathbf{X}, m)] . \quad (24)$$

The linear response of a variable $A(\mathbf{X}), \delta\langle A\rangle_n \equiv \int d\mathbf{X}[P(\mathbf{X}; n) - P_{\rm st}(\mathbf{X})]A(\mathbf{X})$ is thus given by

$$\delta\langle A\rangle_n = -\varepsilon \sum_{m=0}^{n-1} \int d\mathbf{X}\ A_{n-m-1}(\mathbf{X})\nabla\cdot L^\dagger \times [P_{\rm st}(\mathbf{X})\mathbf{F}(\mathbf{X}, m)] . \quad (25)$$

Putting $F(\mathbf{X}, m) = \mathbf{B}(\mathbf{X})s(m)$, where $s(m)$ denotes a time-dependent field conjugate to $\mathbf{B}$, we obtain after integration by parts

$$\delta\langle A\rangle_n = \varepsilon \sum_{m=0}^{n-1} \phi_{n-m-1}(A|\mathbf{B})s(m) , \quad (26)$$

with the response function $\phi_m(A|\mathbf{B})$ given by [9]

$$\phi_m(A|\mathbf{B}) = \langle \mathbf{B}(\mathbf{X}), L[\nabla\cdot A_m(\mathbf{X})]\rangle . \quad (27)$$

Up to now we have been concerned with a deterministic map (1). When the map becomes noisy under the influence of a (stationary) noise $\mathbf{g}_n$, Eq. (1) is modified to

$$\mathbf{X}_{n+1} = \mathbf{G}(\mathbf{X}_n) + \mathbf{g}_n , \quad (28)$$

and the δ function in Eq. (2) should be changed to the distribution function $P_g(\mathbf{g})$ of the noise, thus

$$P(\mathbf{X}; n+1) = \int d\mathbf{X}' P(\mathbf{X}'; n)P_g(\mathbf{X} - \mathbf{G}(\mathbf{X}')) \equiv L^\dagger P(\mathbf{X}; n) . \quad (29)$$

Time evolution operator L of a dynamical variable is defined as in Eq. (3) as the adjoint of $L^\dagger$, Eq. (29), thus $LF(\mathbf{X}) = \int d\mathbf{X}'F(\mathbf{X}')P_g(\mathbf{X}' - \mathbf{G}(\mathbf{X}))$. $F_n(\mathbf{X}) \equiv (L)^n F(\mathbf{X})$ gives the Heisenberg representation of $F(\mathbf{X})$ at time n as before, Eq. (6). Formal manipulations for the derivation of the GLE (17) remain entirely unaffected by these modifications. As to the LRT we note that Eq. (23) remains intact, leading to the conclusion that Eqs. (26) and (27) are valid for the noise map, too.

Here we show how a dynamical process $\{\mathbf{A}_n\}$ can be classified according to its stochastic properties, based on the GLE (17) or (13). For the purpose we take as a set of variables $\mathbf{A}$ of the GLE the following [10]:

$$A(\mathbf{a};\mathbf{X})\equiv\prod_{i=1}^{K}\delta(A^{(i)}(\mathbf{X})-a^{(i)})=\delta(\mathbf{A}(\mathbf{X})-\mathbf{a}) , \quad (30)$$

where $A(\mathbf{a};\mathbf{X})$ represents the probability for $\mathbf{A}(\mathbf{X})$ to realize a value $\mathbf{a}$. From the fact that $\langle A(\mathbf{a};\mathbf{X}), A(\mathbf{b},\mathbf{X})\rangle=P_{\rm st}^{*}(\mathbf{a})\delta(\mathbf{a}-\mathbf{b})$ with $P_{\rm st}^{*}(\mathbf{a})\equiv\int d\mathbf{X}\,\delta(\mathbf{A}(\mathbf{X})-\mathbf{a})P_{\rm st}(\mathbf{X})$ a stationary distribution of $\mathbf{A}(\mathbf{X})$, we can interpret the two-time correlation function,

$$\Xi_n(\mathbf{a}|\mathbf{b})\equiv\int d\mathbf{c}\langle A_n(\mathbf{a};\mathbf{X}), A(\mathbf{c};\mathbf{X})\rangle\langle A(\mathbf{c};\mathbf{X}), A(\mathbf{b};\mathbf{X})\rangle^{-1}=\langle A_n(\mathbf{a};\mathbf{X}), A(\mathbf{b};\mathbf{X})\rangle/P_{\rm st}^{*}(\mathbf{b})$$

as the transition probability to be in a state $\mathbf{A}=\mathbf{a}$ at time n given that the state at $t=0$ is $\mathbf{A}=\mathbf{b}$. Of course, $A_n(\mathbf{a};\mathbf{X})\equiv L^nA(\mathbf{a};\mathbf{X})$ [see Eq. (6)]. The $\mathbf{a},\mathbf{b}$ element of the frequency matrix (14) is given as

$$\Omega(\mathbf{a},\mathbf{b})=\Xi_1(\mathbf{a}|\mathbf{b}) . \quad (31)$$

If the process $\{A_n\}$ is Markovian, we have by definition

$$\Xi_n(\mathbf{a}|\mathbf{b})=\int d\mathbf{c}_1\cdots d\mathbf{c}_{n-1}\Xi_1(\mathbf{a}|\mathbf{c}_1)\times\Xi_1(\mathbf{c}_1|\mathbf{c}_2)\cdots\Xi_1(\mathbf{c}_{n-1}|\mathbf{b})\equiv\Xi_1^n(\mathbf{a}|\mathbf{b}) . \quad (32)$$

From Eqs. (31), (32), and (13) we see that the Markovian property is equivalent to the vanishing memory $\Phi_z=0$. The next simplest, i.e., doubly Markovian, stochastic process is the one in which $\Phi_m(\mathbf{a}|\mathbf{b})=0$ for $m\geq 1$. In this case we readily see that $\Phi_0(\mathbf{a}|\mathbf{b})=\Omega^2(\mathbf{a}|\mathbf{b})-\Xi_2(\mathbf{a}|\mathbf{b})$. Thus, if we could choose a proper set of variables $\mathbf{A}$, which is expected to form a Markovian process [11], Ω defined by Eq. (14) gives valuable insight into dynamics in the system of interest.

To illustrate this point we consider a neural network composed of N Ising spins, $\mathbf{S}=\{S_1,\ldots,S_N\}$, whose interaction obeys the Hebbian rule [12],

$$J_{ij}=(1/N)\sum_{u=1}^{p}\xi_i^{(u)}\xi_j^{(u)}(1-\delta_{ij}) . \quad (33)$$

Denoting by $P(\mathbf{S};n)$ the probability that the system is in the state $\mathbf{S}$ at time n, synchronous dynamics is defined by $P(\mathbf{S};n+1)=\mathrm{Tr}(\mathbf{S}')W(\mathbf{S}|\mathbf{S}')P(\mathbf{S}';n)\equiv L^{\dagger}P(\mathbf{S};n)$ with the microscopic transition probability

$$W(\mathbf{S}|\mathbf{S}')=\exp[\textstyle\sum\beta h_i'S_i]/\prod 2\cosh(\beta h_i') . \quad (34)$$

Here, $\mathrm{Tr}(\mathbf{S}')$ means the summation over $\mathbf{S}'$, β the inverse temperature, $h_i\equiv\sum J_{ij}S_j$ $(h_i'\equiv\sum J_{ij}S_j')$, and i and j run from 1 to N. The stationary distribution function is given by $P_{\rm st}(\mathbf{S})=C\prod 2\cosh(\beta h_i)$ with C a normalization constant. The operator L, adjoint to $L^{\dagger}$, is given from Eq. (3) with the integration over $\mathbf{X}$ replaced by the summation over $\mathbf{S}$ as $LF(\mathbf{S})=\mathrm{Tr}(\mathbf{S}')\exp[\beta\sum h_iS_i']F(\mathbf{S}')/\prod 2\cosh(\beta h_i)$.

We are interested in dynamics of overlap $o^{(u)}(\mathbf{S})$ defined by $o^{(u)}=\sum\xi_i^{(u)}S_i/N$ $(u=1,\ldots,P)$ and put $A(\mathbf{m};\mathbf{S})\equiv\delta(\mathbf{o}(\mathbf{S})-\mathbf{m})$, Eq. (30) [13]. The (macroscopic) transition probability $\Omega(\mathbf{m},\mathbf{m}')$, Eq. (31), for the order parameter to change from $\mathbf{m}'$ to $\mathbf{m}$ is expressed as

$$\Omega(\mathbf{m},\mathbf{m}')=E(\mathbf{m},\mathbf{m}')/D(\mathbf{m}') , \quad (35)$$

where

$$E(\mathbf{m},\mathbf{m}')\equiv\mathrm{Tr}(\mathbf{S},\mathbf{S}')\times\exp[\beta\textstyle\sum h_i'S_i]\delta(\mathbf{o}(\mathbf{S})-\mathbf{m})\delta(\mathbf{o}(\mathbf{S}')-\mathbf{m}')$$

and

$$D(\mathbf{m}')\equiv\mathrm{Tr}(\mathbf{S},\mathbf{S}')\exp[\beta\textstyle\sum h_i'S_i]\delta(\mathbf{o}(\mathbf{S}')-\mathbf{m}') .$$

$E(\mathbf{m},\mathbf{m}')$ and $D(\mathbf{m}')$ are seen to represent the generalized partition functions, thus in the limit $N\to\infty$ we have $E(\mathbf{m},\mathbf{m}')=\exp[-\beta N e(\mathbf{m},\mathbf{m}')]$ and $D(\mathbf{m}')=\exp[-\beta N d(\mathbf{m}')]$. If we confine ourselves to the case of finite p, that is, $\alpha\equiv p/N\to 0$, $e(\mathbf{m},\mathbf{m}')$ and $d(\mathbf{m}')$ are easily calculated without the replica method. Following closely the saddle-point calculation [2,12], we have

$$d(\mathbf{m}')=\mathbf{m}'\cdot\mathbf{t}'-\langle\langle\ln[\cosh(\beta\xi\cdot\mathbf{m}')]\rangle\rangle/\beta-\langle\langle\ln[\cosh(\beta\xi\cdot\mathbf{t}')]\rangle\rangle/\beta , \quad (36)$$

$$e(\mathbf{m},\mathbf{m}')=\mathbf{t}\cdot\mathbf{m}+\mathbf{t}'\cdot\mathbf{m}'-\mathbf{m}\cdot\mathbf{m}'-\langle\langle\ln[\cosh(\beta\xi\cdot\mathbf{t})]\rangle\rangle/\beta-\langle\langle\ln[\cosh(\beta\xi\cdot\mathbf{t}')]\rangle\rangle/\beta , \quad (37)$$

where $\langle\langle\ \rangle\rangle$ denotes the average over the random pattern ξ and $\mathbf{t}$ and $\mathbf{t}'$ are variationally determined from $\mathbf{m}\equiv\langle\langle\xi\tanh(\beta\mathbf{t}\cdot\xi)\rangle\rangle$ and $\mathbf{m}'=\langle\langle\xi\tanh(\beta\mathbf{t}'\cdot\xi)\rangle\rangle$. The situation is greatly simplified if we consider that only one overlap, say, m_1 is of order 1 and $m_i=t_i=0$ for $i\geq 2$. In this case, we see from Eqs. (36), (37), and (35) that $\Omega(m_1,m_1')=\exp[NT(m_1,m_1')]$ with

$$T(m_1,m_1')=-\beta\{tm_1-m_1m_1'+\ln[\cosh(\beta m_1')]/\beta-\ln[\cosh(\beta t)]/\beta\},$$

where $m_1=\tanh(\beta t)$. Based on this expression for the transition probability we can discuss how the mean μ_n and the mean square deviation σ_n^2, defined by $P_n(m_1)\propto\exp[-N(m_1-\mu_n)^2/(2\sigma_n^2)]$, change by one iteration. Starting from $P_{n+1}(m)=\int dm'\exp\{N[T(m,m')-(m_1-\mu_n)^2/(2\sigma_n^2)]\}$, we can determine μ_{n+1} and σ_{n+1}^2 by saddle-point calculations to have

$$\mu_{n+1}=\tanh(\beta\mu_n) ,$$
$$\sigma_{n+1}^2=(1-\mu_{n+1}^2)[1+(\beta\sigma_n)^2(1-\mu_{n+1}^2)] . \quad (38)$$

It is readily confirmed that μ_n and σ_n^2 converge to the values determined by the equilibrium theory.

In this paper, the GTBM and LRT, which have been playing important roles in continuous-time dynamical systems, are formulated for general discrete-time systems. The last comment is on the GLE (17), which expresses $\mathbf{A}_{n+1}$ in terms of the history $\{\mathbf{A}_m\}(m=n,n-1,\ldots,)$ and the noise $\mathbf{f}_n$. This reminds us of the AR (autoregression) [6], model to reproduce stochastic signals produced by complex systems. For an AR modeling, the coefficients $\{\Phi_m\}$ are regarded as fitting parameters, which are independent of the nature of the noise $\mathbf{f}_m$. Equations (17) and (18) show that the coefficients and the noise are closely related for a very general, even for nonthermal, system.

The author expresses his sincere gratitude to Professors A. Igarashi, H. Sakai, and H. Fujisaka for useful discussions.

[1] H. G. Schuster, *Deterministic Chaos* (VCH, Weinheim, 1988).

[2] D. J. Amit, *Modeling Brain Function* (Cambridge University Press, England, 1989), Chaps. 3 and 4.

[3] (a) H. Mori, Prog. Theor. Phys. **33**, 423 (1965); (b) R. Zwanzig, Phys. Rev. **124**, 983 (1961). See also for a general introduction to the projection operator method, R. Zwanzig, Lect. Theor. Phys. (Boulder) **3**, 106 (1961).

[4] R. Kubo, J. Phys. Soc. Jpn. **12**, 570 (1957). See also L. E. Reichl, *A Modern Course in Statistical Physics* (University of Texas Press, Austin, 1980).

[5] A. Lasota and M. Mackey, *Probabilistic Properties of Deterministic Systems* (Cambridge University Press, England, 1985).

[6] K. J. Astroem, *Introduction to Stochastic Control Theory* (Academic Press, New York, 1970).

[7] We note that the GLE (17) was apparently first derived for a one-dimensional deterministic map $X_{n+1}=G(X_n)$, with $A=X$ to study the time correlation function of X_n. H. Fujisaka and T. Yamada, Z. Naturforsch. **33a**, 1455 (1978). Recently, the projection operator formalism is used in a somewhat different context by W. Just, J. Stat. Phys. **67**, 271 (1992). Our derivation of the GLE closely follows the original one by Mori [3] and elucidates some salient features of the random noise as given below.

[8] This problem is considered by some authors in connection with deterministic one-dimensional maps: J. Helstab, H. Thomas, T. Geisel, and G. Radons, Z. Phys. B **50**, 141 (1983); T. Geisel, J. Heldstab, and H. Thomas, Z. Phys. B **55**, 165 (1984); S. Grossmann, Z. Phys. B **57**, 77 (1984); M. Falcioni, S. Isola, and A. Vulpiana, Phys. Lett. A **144**, 341 (1990). We present a brief derivation of the LRT below, partly because our treatment, extended to include a noisy map and a discrete-time master equation, seems to give a slight generalization of the existing result and partly in order to make the presentation in this paper a coherent one.

[9] For a chaotic map it often happens that the $P_{\rm st}(\mathbf{X})$ has a support on a fractal set. In this case the ∇ in Eq. (24) becomes meaningless and more generally, the calculation of the average $\langle\ \rangle$, as in Eq. (27), becomes rather difficult. However, if a map becomes noise, the distribution function is smeared out, thus removing the difficulty.

[10] H. Mori and H. Fujisaka, Prog. Theor. Phys. **49**, 764 (1973).

[11] M. S. Green, J. Chem. Phys. **20**, 1281 (1952); **22**, 398 (1954).

[12] As a general reference to this subject see Ref. [2] and P. Peretto, *An Introduction to the Modeling of Neural Networks* (Cambridge University Press, England, 1992).

[13] For a continuous-time (asynchronous) system, the overlap dynamics is recently discussed by T. Munakata and Y. Nakamura, Phys. Rev. E **47**, 3792 (1993).