Modified-moment method for the Fokker-Planck equation and some aspects of the thermodynamics of irreversible processes

Byung Chan Eu* and D. K. Bhattacharya
Department of Chemistry, McGill University, 801 Sherbrooke Street West, Montreal, PQ, Canada H3A 2K6
(Received 11 February 1988; revised manuscript received 15 November 1988)

The modified-moment method is applied to solve the Fokker-Planck equation in such a way that the solution is consistent with the requirements of the thermodynamic laws. The formal solution has an attendant mathematical structure for the entropy which is similar to that obtained from the Boltzmann equation. A generalized Gibbs relation is obtained for such a thermodynamic branch of solution. The formalism therefore may be used to study thermodynamic aspects of stochastic processes underlying the Fokker-Planck equation. The method is shown to yield the same analytical results as other methods in the case of linear processes, and in the case of nonlinear processes the moment series converges sufficiently fast to justify a lower-order truncation of the moment series for the distribution function, if the diffusion constant (D) is less than a critical value. In this case the solution of the Fokker-Planck equation tends to the steady-state distribution function. It is shown that the second variation of the nonequilibrium part of the entropy may be used as a Lyapunov function, which provides local criteria of evolution only.

I. INTRODUCTION

Description of macroscopic irreversible processes is often based on the theory of stochastic processes for variables we wish to describe in space-time. Under the Markovian assumption on the processes the description is basically reduced to solving the Fokker-Planck equation[1,2] or a master equation.[3] Especially, the Fokker-Planck equation determines the probability distribution of the stochastic variables as a function of time in the space of the random variables chosen. If the stochastic variables are denoted by $\alpha_1,\alpha_2,\ldots,\alpha_p$, the Fokker-Planck equation for probability distribution or transition probability

$$P(\underline{\alpha}|\underline{\alpha}^0;t)=P(\alpha_1,\alpha_2,\ldots,\alpha_p|\alpha_1^0,\alpha_2^0,\ldots,\alpha_p^0;t)$$

may be written as

$$\frac{\partial}{\partial t}P(\underline{\alpha}|\underline{\alpha}^0;t)=-\sum_{i=1}^{p}\frac{\partial}{\partial\alpha_i}G_i(\underline{\alpha})P+\tfrac{1}{2}\sum_{i=1}^{p}\sum_{j=1}^{p}\frac{\partial}{\partial\alpha_i}D_{ij}(\underline{\alpha})\frac{\partial}{\partial\alpha_j}P(\underline{\alpha}|\underline{\alpha}^0;t)\ , \tag{1.1}$$

where G_i is the "velocity" of process α_i and D_{ij} are diffusion tensors; both G_i and D_{ij} are generally functions of $\alpha_1,\alpha_2,\ldots,\alpha_p$. If the processes are linear, the velocity G_i is a linear combination of the stochastic variables. The Fokker-Planck equation is analytically solvable for arbitrary p if the processes are linear and the diffusion tensors are independent of $\alpha_1,\alpha_2,\ldots,\alpha_p$. For nonlinear processes an analytic solution is generally not possible. There have been a number of formal theories[4,5] proposed to express the distribution function, and the path-integral method[4] and the WKB-type approximation method[6,7] are two typical examples for such theories. They make it possible to express the distribution function in terms of the Onsager-Machlup function,[8] which is the stochastic theory analog of the Lagrangian in classical mechanics. But such methods present formidable computational problems except for linear processes for which there is no need for such theories since the Fokker-Planck equation can be directly and explicitly solved. In any event, if the Fokker-Planck equation is a correct approach to the description of macroscopic irreversible processes it is expected to yield a mathematical structure that would reveal its connection with the thermodynamics of the system under consideration. The reason for such an expectation is that all natural processes must conform to the thermodynamic laws, and the Fokker-Planck equation would not be an exception if it was impeccable as an evolution equation for the distribution function for a realistic physical system. In this paper we present an alternative method of solution for the Fokker-Planck equation which carries with it a concomitant mathematical structure of (irreversible) thermodynamics. We will call such a solution the thermodynamic branch of solution. The method employed in the present investigation is similar in spirit to the modified-moment method[9] used for obtaining the thermodynamic branch of solution for the Boltzmann equation.

In this method the solution of the Fokker-Planck equation is looked for as a functional of the correlation functions or moments of stochastic variables which may be determined by their evolution equations. These evolution equations are macroscopic equations. Since the distribution function is a functional of such macroscopic observables or moments the temporal evolution of the latter determines the evolution of the former, and the thermodynamic laws are brought in to control the evolution of the moments. This approach is advantageous on two sides: for one, inevitable approximations for the distribution function are made such that they would yield macroscopic quantities conforming fully to the thermodynamic laws; and for the other, the resulting mathematical structure becomes easily amenable to interpretation in terms

of the observables measured in the laboratory and it even presents a possibility for formulating a semiempirical theory in which the laboratory information is used for constructing the distribution function that may be employed for some other purposes in connection with the stochastic variables in question. When applied to the Boltzmann equation, the modified-moment method[9] yields the solution (i.e., the distribution function) which reduces to the Chapman-Enskog first-order solution holding near equilibrium. Therefore one is tempted to imagine that the modified-moment method solution would yield a first-order solution for the Fokker-Planck equation which is equivalent to the Chapman-Enskog-type solution. In fact, there is some validity in this expectation since the Fokker-Planck equation may be regarded as being equivalent to a homogeneous Boltzmann equation in which the collision term is replaced by the Fokker-Planck operator $R[P]$: see (2.1) for the definition.

In this paper the actual analysis is made under the assumption of the potential condition being satisfied by the Fokker-Planck operator. This condition is imposed since we wish to obtain the steady-state solution in terms of a potential. Even if this condition were not met, the modified-moment method would be applicable, but the steady-state solution would not be as simple as the case of the potential condition being satisfied.

In Sec. II the H theorem and the entropy are considered. In Sec. III the modified-moment method is developed for the Fokker-Planck equation and the connection with the thermodynamics of the resulting formalism is discussed. The formalism is illustrated with some examples. The examples with linear processes show that the method yields exactly the same distribution functions as by the usual method of solution in the case of Gaussian random processes. In the case of a nonlinear process considered as an example it is numerically shown that the moment series converges sufficiently fast to justify a lower-order truncation of the series. This test indicates that the method can be practical and useful for studying Fokker-Planck equations for nonlinear processes. In Sec. IV a theory of stability is developed based on the entropy calculated from the Fokker-Planck equation. The second variation of the nonequilibrium part of the entropy is found to have the properties required of a Lyapunov function. Section V is for concluding remarks.

II. THE H THEOREM AND THE ENTROPY

Let the distribution function or the transition probability be described by the following Fokker-Planck equation:

$$\frac{\partial}{\partial t}P(\underline{\alpha}|\underline{\alpha}^0;t)=-\sum_{i=1}^{p}\frac{\partial}{\partial\alpha_i}G_i(\underline{\alpha})P+\tfrac{1}{2}\sum_{i=1}^{p}\sum_{j=1}^{p}\frac{\partial}{\partial\alpha_i}D_{ij}(\underline{\alpha})\frac{\partial}{\partial\alpha_j}P(\underline{\alpha}|\underline{\alpha}^0;t)\equiv R[P]\ , \tag{2.1}$$

where $G_i(\underline{\alpha})$ may be a nonlinear function of $\underline{\alpha}$ and D_{ij} may depend on the stochastic variables as well. Whenever convenient, the p-dimensional gradient operator will be abbreviated with ∇. The present study will be confined to the class of processes which satisfies the potential condition:[2(a)] that is, there exists a function $V(\underline{\alpha})$ such that

$$\beta V(\underline{\alpha})=-2\int^{\underline{\alpha}}d\underline{\alpha}\cdot\underline{D}^{-1}\underline{G}(\underline{\alpha})\ . \tag{2.2}$$

Although special and restrictive, this class still covers a sufficiently wide range of processes of interest and also allows a more definitive study of the thermodynamic basis of stochastic processes than those not satisfying the potential condition. In fact, finding the steady-state solution in the case of the potential condition not satisfied is equivalent to solving Pfaff's problem in the theory of differential forms.[10] In this case it is not possible to find $V(\underline{\alpha})$ as easily, and it just might be that the Fokker-Planck equation approach is impractical for all intents and purposes. It must be stressed that the modified-moment method and the thermodynamic theory developed here are equally valid and applicable even if the potential condition is removed, as long as there exists a function $V(\underline{\alpha})$ such that the steady-state solution of the Fokker-Planck equation is represented by the exponential form as in (2.5a) below. If not, it is necessary to develop a theory for V. As mentioned earlier, this is not a trivial problem.

The Fokker-Planck equation is irreversible in the sense that the time-reversal invariance is broken and it admits the H functional as a statistical model for entropy. Before introducing the H functional we define the steady-state solution P_0 of (2.1) by the equation

$$R[P_0]=0\ . \tag{2.3}$$

With the potential condition (2.2) this equation implies

$$P_0=C\exp[-\beta V(\underline{\alpha})]\ ,$$

where β is a parameter as yet undetermined and C is the normalization factor

$$C=\langle P_0\rangle/\langle\exp[-\beta V(\underline{\alpha})]\rangle\ .$$

The angular brackets denote the integration over the stochastic variables $\underline{\alpha}=\{\alpha_i\}$ from $-\infty$ to $+\infty$. If the distribution function is normalized to unity

$$\langle P_0\rangle=\langle P\rangle=1\ , \tag{2.4}$$

then

$$P_0=\exp[-\beta V(\underline{\alpha})]/Q\ , \tag{2.5a}$$

$$Q=\langle\exp[-\beta V(\underline{\alpha})]\rangle\ . \tag{2.5b}$$

The meaning of β will be fixed when the entropy of the system is identified as will be done shortly.

We now introduce the H functional:

$$S(t)=-k_B\langle P(\underline{\alpha}|\underline{\alpha}^0;t)\ln P(\underline{\alpha}|\underline{\alpha}^0;t)\rangle\ , \tag{2.6}$$

where k_B is the Boltzmann constant. This is the statistical formula for the entropy of the system. It is convenient to split $S(t)$ into two parts as follows:

$$S(t)=S_e+S_n \ , \tag{2.7}$$

where

$$S_e=-k_B\langle P\ln P_0(\underline{\alpha})\rangle \ , \tag{2.8}$$

$$S_n=-k_B\langle P\ln(P/P_0)\rangle \ . \tag{2.9}$$

The significance of these two factors will become clear in what follows.

If (2.5a) is inserted into (2.8), there follows a more explicit formula for S_e:

$$S_e=k_B\beta E+k_B\ln Q \ , \tag{2.10}$$

where

$$E=\langle V(\underline{\alpha})P\rangle \ . \tag{2.11}$$

If S, S_e, and S_n are to have the attributes of entropy, it is necessary to choose the parameter β such that

$$\beta=1/k_BT \ , \tag{2.12}$$

where T must be identified with temperature. The reason for this identification will become obvious when the Gibbs relation is derived later. Then by identifying the "work function"

$$\mathcal{A}=-k_BT\ln Q \ , \tag{2.13}$$

we obtain

$$S_e=(E-\mathcal{A})/T \tag{2.14}$$

in analogy to the relation for equilibrium entropy in the canonical-ensemble theory.[11] We caution the reader that despite the symbol, E is not the internal energy in the conventional statistical thermodynamics unless the stochastic variable happens to be the velocity of the particle. Nevertheless, since there must be thermodynamic principles obeyed *on the average* by the stochastic processes under consideration, the analogy taken advantage of above appears inevitable. In this connection it is insightful to note that in the case of a linear process for which $\underline{G}=-\lambda\underline{\alpha}$, there holds

$$\beta V=\underline{D}^{-1}:\lambda\underline{\alpha}\,\underline{\alpha} \ ,$$

where λ is the phenomenological coefficient for the linear process, and the steady-state value E^0 for E is indeed like the energy since

$$E^0=\langle V(\underline{\alpha})P_0\rangle=pk_BT/2 \ ,$$

which is simply the result of the equipartition law. It must be noted that the quadratic form for $V(\underline{\alpha})$ taken above is not the form generally assumed in the subsequent development of the theory presented here. For nonlinear processes the steady-state energy is expected to be a complicated function of T as is the average potential energy in the conventional canonical-ensemble theory of statistical thermodynamics. To make the parameter T appearing in this theory fully determined, we now define T in the case of nonequilibrium by the relation

$$\frac{1}{2}pk_BT\equiv\langle\underline{D}^{-1}:\lambda\underline{\alpha}\,\underline{\alpha}P\rangle \ . \tag{2.12'}$$

Just like the temperature in the Boltzmann kinetic theory this definition together with the attendant thermodynamics makes it possible to determine thermodynamic variables in the units of k_BT, and there now is no undefined parameter in this theory.

We next consider S_n. Differentiation of S_n with t and use of (2.1) and (2.2) yield the well-known inequality

$$\frac{dS_n}{dt}=\tfrac{1}{2}k_B\langle P\nabla\ln(P/P_0)\cdot\underline{D}\cdot\nabla\ln(P/P_0)\rangle\geq 0 \ , \tag{2.15}$$

which embodies the content of the H theorem for the Fokker-Planck equation. We may therefore identify (dS_n/dt) with the entropy production due to the stochastic process represented by the Fokker-Planck equation. We regard $S(t)$ as the entropy of the system. Then it is possible to look upon

$$dS/dt=dS_e/dt+dS_n/dt \tag{2.16}$$

as the entropy balance equation for the system especially in view of the inequality (2.15). The first term on the right in (2.16) does not have a definite sign that can be attached to it. It is related to the rate of change in "energy" as will be shown later; see (3.4) below. Since there are no notions of configuration space and volume in the present Fokker-Planck equation approach, the dS_e/dt term cannot be cast in divergence form as is the case with the entropy balance equation appearing in the Boltzmann kinetic theory where the dS_e/dt term should be replaced by the divergence of entropy flux: $-\nabla_r\cdot\mathbf{J}_s$ where ∇_r is the spatial gradient operator and $\mathbf{J}_s$ is the entropy flux. As will be seen later, this absence of configuration space and the concept of volume in the Fokker-Planck equation approach results in the absence of the pressure-volume work term in the nonequilibrium Gibbs equation (3.10b) below.

III. MODIFIED-MOMENT METHOD FOR THE FOKKER-PLANCK EQUATION

We now look for $P(\underline{\alpha}|\underline{\alpha}^0;t)$ in terms of moments by following the spirit of the modified-moment method.[9] For this purpose it is convenient to introduce the following abbreviations for moments:

$$\begin{aligned}
&h^{(1)}=\underline{\alpha} \ ,\\
&h^{(2)}=(\underline{\alpha}-\bar{\underline{\alpha}})(\underline{\alpha}-\bar{\underline{\alpha}}) \ ,\\
&\vdots\\
&h^{(m)}=(\underline{\alpha}-\bar{\underline{\alpha}})(\underline{\alpha}-\bar{\underline{\alpha}})\cdots(\underline{\alpha}-\bar{\underline{\alpha}}) \ ,\\
&\vdots
\end{aligned} \tag{3.1}$$

where

$$\bar{\underline{\alpha}}=\langle\underline{\alpha}P\rangle \ .$$

Therefore $h^{(1)}$ is a p-dimensional vector and $h^{(2)}$, $h^{(3)}$, etc. are tensors of rank 2, 3, etc. It is possible to use an orthogonal set of tensor polynomials such as tensor Hermite polynomials instead of (3.1), but the advantage of such an orthogonal set is minor in the present case of

nonlinear processes since the steady-state distribution function is no longer Gaussian. The distribution function is sought in the form

$$P=P_0(\underline{\alpha})\exp\left[-\beta\left[\sum_m X^{(m)}\odot h^{(m)}-\mu_n\right]\right], \tag{3.2}$$

where $X^{(m)}$ are the unknowns to be determined such that P satisfies the Fokker-Planck equation and μ_n is defined by

$$\exp(-\beta\mu_n)=\left\langle\exp\left[-\beta\sum_m X^{(m)}\odot h^{(m)}\right]P_0(\underline{\alpha})\right\rangle . \tag{3.3}$$

The symbol $\odot$ stands for the scalar product of the tensors involved. The steady-state distribution function P_0 is given by (2.5a). We will shortly present a method for determining $X^{(m)}$.

Before doing so, it is useful to perform the following formal calculation in order to see the significance of Eq. (2.16) in a better light. The following relation is useful for the purpose:

$$\begin{aligned}\frac{dE}{dt}&=\left\langle V(\underline{\alpha})\frac{\partial P}{\partial t}\right\rangle\\&=\beta^{-1}\left\langle[\ln(1/Q)-\ln P_0]\frac{\partial P}{\partial t}\right\rangle\\&=-\beta^{-1}\left\langle\ln P_0\frac{\partial P}{\partial t}\right\rangle\\&=T(dS_e/dt) .\end{aligned} \tag{3.4}$$

This relation is rather suggestive for the meaning of the parameter T appearing in this theory; see (2.12′). With the definitions

$$A^{(m)}=\langle h^{(m)}P\rangle , \tag{3.5a}$$

$$\Lambda^{(m)}=\langle h^{(m)}R[P]\rangle , \tag{3.5b}$$

$$B^{(m)}=\langle(\partial h^{(m)}/\partial\bar{\alpha})P\rangle \quad (B^{(1)}=B^{(2)}=0) , \tag{3.5c}$$

we find the evolution equations for moments $A^{(m)}$,

$$\frac{dA^{(m)}}{dt}=B^{(m)t}\cdot\Lambda^{(1)}+\Lambda^{(m)} \quad \text{for all } m . \tag{3.6}$$

Here the superscript t means the transpose. This set represents an open hierarchy of coupled ordinary differential equations for $A^{(m)}$ which is equivalent to the Fokker-Planck equation (2.1).

If (3.2) is substituted into the equation

$$\frac{dS_n}{dt}=-k_B\langle\ln(P/P_0)R[P]\rangle , \tag{3.7}$$

we find

$$\frac{dS_n}{dt}=T^{-1}\sum_m X^{(m)}\odot\Lambda^{(m)} , \tag{3.8}$$

and by eliminating the dissipative term $\Lambda^{(m)}$ with the evolution equation for $A^{(m)}$, we obtain

$$\frac{dS_n}{dt}=T^{-1}\sum_m X^{(m)}\odot\left[\frac{dA^{(m)}}{dt}-B^{(m)t}\cdot\frac{dA^{(1)}}{dt}\right] . \tag{3.9}$$

When this result is combined with (3.4), there follows the relation for entropy:

$$\frac{dS_n}{dt}=T^{-1}\left[\frac{dE}{dt}+\sum_m\bar{X}^{(m)}\odot\frac{dA^{(m)}}{dt}\right] , \tag{3.10a}$$

where

$$\bar{X}^{(1)}=X^{(1)}-\sum_{m\geq 3}X^{(m)}\odot B^{(m)t} ,$$

$$\bar{X}^{(m)}=X^{(m)} \quad \text{for } m\geq 2 .$$

Cast in differential form, the relation (3.10a) reads

$$T\,dS=dE+\sum_{m\geq 1}\bar{X}^{(m)}\odot dA^{(m)} . \tag{3.10b}$$

This is a Gibbs relation and an important formal result of this paper which puts the stochastic process in contact with the thermodynamic laws. With the help of this equation we obtain a thermodynamic interpretation of T:

$$T^{-1}=(\partial S/\partial E)_A , \tag{3.11}$$

which is the thermodynamic temperature if S is interpreted as the entropy. This is the basis of the assertion made previously regarding T. We remark that the concept of absolute temperature T is inseparable from the concept of entropy. In other approaches[3(b),12] to the Fokker-Planck equations the temperature is introduced through the fluctuation-dissipation theorem, but the latter is some steps removed from the primordial concept of entropy.

Similarly to (3.11), there holds the relation

$$T^{-1}\bar{X}^{(m)}=(\partial S/\partial A^{(m)})_E , \tag{3.12}$$

which is a natural consequence of (3.10b). As in thermodynamics, if the Gibbs relation is used as a phenomenological equation which determines the entropy, then $\bar{X}^{(m)}$ must be empirically supplied as a function of E and $A^{(m)}$. These empirical relations are constitutive equations characterizing the system in hand. In the present theory $\bar{X}^{(m)}$ are determined from the consistency condition (3.14) below.

With the new interpretation of T and $\bar{X}^{(m)}$ as given in (3.11) and (3.12) we can express the distribution function in a more insightful form in terms of the entropy derivatives:

$$P=\exp\left[-k_B^{-1}(\partial S/\partial E)[V(\underline{\alpha})-\mu]-k_B^{-1}\left[\sum_{m\geq 1}(\partial S/\partial A^{(m)})\odot h^{(m)}-\sum_{m\geq 3}(\partial S/\partial A^{(m)})\odot B^{(m)}\cdot h^{(1)}\right]\right], \tag{3.2'}$$

where

$$\mu=\mu_n-k_BT\ln Q .$$

To find the equation determining $X^{(m)}$ explicitly in terms of moments $A^{(m)}$, we now return to the entropy production (2.15) which may be calculated with (3.2) in the form

$$(dS_n/dt)=(k_B\beta^2/2)\sum_m\sum_l X^{(m)}\odot\langle P\nabla h^{(m)}\cdot\underline{D}\cdot\nabla h^{(l)}\rangle\odot X^{(l)} . \tag{3.13}$$

Since the entropy production can also be given by the form (3.8), comparison of (3.13) with (3.8) yields a set of algebraic equations for $X^{(m)}$:

$$2^{-1}\beta\langle P\nabla h^{(m)}\cdot\underline{D}\cdot\nabla h^{(l)}\rangle\odot X^{(l)}=\Lambda^{(m)}(\{A^{(m)}\}) \tag{3.14}$$

for all m. This is a linear set for tensors $X^{(l)}$ which, when solved, yields $X^{(l)}$ so far undetermined. This set may be solved by the following device. First define a tensor of rank $(m+l)$ by the formula

$$C^{(m+l)}=\langle P\nabla h^{(m)}\cdot\underline{D}\cdot\nabla h^{(l)}\rangle . \tag{3.15a}$$

If the sequence of moments is truncated at q, a vector X can be constructed with the independent elements of vector $X^{(1)}$ and tensors $X^{(2)},\ldots,X^{(q)}$:

$$X=\{X_1^{(1)},X_2^{(1)},\ldots,X_p^{(1)},X_{11}^{(2)},\ldots,X_{111}^{(3)},\ldots,X_{11\cdots1}^{(q)}\ldots\} . \tag{3.15b}$$

The dimension of this vector is

$$d=\sum_{m=1}^{q} d_m ,$$

where d_m is the number of independent elements of $X^{(m)}$. We remark here that, for example, the second-rank tensor $X^{(2)}$ is symmetric and therefore there are only six independent elements. A similar remark applies to other tensors. Then a matrix $\underline{C}$ of dimension d may be constructed with the set of tensors $\{C^{(m+l)}\}$,

$$\{C^{(m+l)}\}\rightarrow\underline{C} . \tag{3.15c}$$

Similarly, the tensors $\Lambda^{(m)}$ induce a corresponding vector $\underline{\Lambda}$ of dimension d. Therefore the set (3.14) may be cast into the following linear set:

$$\underline{C}\cdot\underline{X}=2k_BT\underline{\Lambda} . \tag{3.16}$$

If the matrix $\underline{C}$ is nonsingular, this set can be easily solved:

$$\underline{X}=2k_BT\underline{C}^{-1}\underline{\Lambda} , \tag{3.17}$$

which expresses $X^{(m)}$ in terms of $\{A^{(m)}\}$ since $\underline{\Lambda}$ and $\underline{C}$ are generally dependent on $\{A^{(m)}\}$. This represents the qth moment approximant for $X^{(m)}$ and thus for the distribution function P. The truncation of the sequence of moments must be made such that the following two conditions are first of all satisfied: (1) the boundary condition on P is satisfied: (2) the entropy production calculated with $X^{(m)}$ determined with (3.17) must remain positive. There is another criterion that must be met. It is the convergence of the moment expansion. In this manner the solution procedure for the Fokker-Planck equation is reduced to the solution of algebraic equations (3.14), or equivalently (3.16), and the solution of ordinary differential equations for $A^{(m)}$, i.e., the evolution equations for moments, subject to suitable initial conditions. Therefore, as $q\rightarrow\infty$, (3.17) provides $X^{(m)}$ for all the moments (correlation functions) and the solution of the Fokker-Planck equation is obtained in principle. In practice, however, it will be necessary to resort to further approximations to (3.17) or (3.14) in order to make the equation more manageable.

To summarize the solution procedure, the distribution function is first written in an exponential form as in (3.2) and then the algebraic set of equations (3.14) or (3.16) is solved for $X^{(m)}$ appearing in the distribution function. The terms $X^{(m)}$ will be functions of moments (correlation functions). To find the moments as a function of time, their evolution equations are derived from the Fokker-Planck equation and then solved subject to suitable initial conditions on the moments. When the solutions of the moment evolution equations are substituted into $X^{(m)}(\{A^{(l)}\})$, the distribution function is found as a function of $\underline{\alpha}$ and t for the Fokker-Planck equation under consideration. Note that $A^{(l)}$ are functions of time.

The solution (3.17) provides an alternative interpretation of P. Since $\underline{\Lambda}=0$ if $dA^{(m)}/dt=0$ for all m which defines the steady state of moment evolution, we see that the steady state of the Fokker-Planck equation coincides with the steady state of macroscopic processes defined by

$$(dA^{(m)}/dt)=0$$

or

$$B^{(m)t}\cdot\Lambda^{(1)}+\Lambda^{(m)}(\{A^{(l)}\})=0 .$$

This proposition is true, provided the steady state of the moment evolution equations is stable, and will find an application when we consider the question of stability of macroscopic processes in Sec. IV.

As an illustration of the method presented, we consider

a few simple cases. In the calculations presented below, we assume $\underline{D}$ is constant.

A. Linear one-dimensional process

We consider the case of a linear process characterized by

$$G=-\lambda\alpha\ , \tag{3.18}$$

where λ is a positive number. The evolution equations for $A^{(m)}$ can be easily constructed with (3.5b), (3.5c), and (3.6). For this purpose we first calculate $\Lambda^{(m)}$ and $B^{(m)}$ from (3.5b) and (3.5c) by using integration by parts and the definition of $A^{(m)}$. For $m\geq 3$ the evolution equations are linear,

$$\frac{dA^{(m)}}{dt}=-m\lambda A^{(m)}\quad (m\geq 3)\ .$$

When this is integrated over t and the initial condition $A^{(m)}(0)=0$ is taken for all $m\geq 3$, then we find

$$A^{(m)}(t)=0$$

for all $m\geq 3$. The initial conditions taken here mean that the random variable is not initially correlated at all orders of $m\geq 3$. In fact, we take also $A^{(2)}(0)=0$ as we will see shortly. This means that there are only two moments $A^{(1)}(t)$ and $A^{(2)}(t)$ necessary for the solution and the random process is Gaussian. In this case (3.14) consists of two algebraic equations,

$$\begin{aligned}&\beta D X^{(1)}=-2\lambda A^{(1)}\ ,\\ &2\beta D A^{(2)}X^{(2)}=D-2\lambda A^{(2)}\ .\end{aligned} \tag{3.19}$$

We find $X^{(1)}$ and $X^{(2)}$ trivially:

$$\begin{aligned}&X^{(1)}=-2k_BT\lambda A^{(1)}/D\ ,\\ &X^{(2)}=k_BT(D-2\lambda A^{(2)})/2DA^{(1)}\ .\end{aligned} \tag{3.20}$$

The evolution equations for $A^{(1)}$ and $A^{(2)}$ are calculated similarly to those for $A^{(m)}$, $m\geq 3$:

$$\begin{aligned}&\frac{dA^{(1)}}{dt}=-\lambda A^{(1)}\ ,\\ &\frac{dA^{(2)}}{dt}=D-2\lambda A^{(2)}\ ,\end{aligned} \tag{3.21}$$

which may be solved easily:

$$\begin{aligned}&A^{(1)}(t)=\exp(-\lambda t)A_0^{(1)}\equiv\bar{\alpha}_0\exp(-\lambda t)\ ,\\ &A^{(2)}(t)=(D/2\lambda)[1-\exp(-2\lambda t)]\ ,\end{aligned} \tag{3.22}$$

if $A^{(2)}(0)=0$. Substituting (3.22) into (3.20) yields $X^{(1)}$ and $X^{(2)}$ in terms of time:

$$\begin{aligned}&X^{(1)}(t)=-2k_BT\lambda\bar{\alpha}_0\exp(-\lambda t)/D\ ,\\ &X^{(2)}(t)=\lambda k_BT/D[\exp(-2\lambda t)-1]\ .\end{aligned} \tag{3.23}$$

By substituting these results into (3.2) and calculating P_0 with (2.5) and (3.18), we find the distribution function in the form

$$\begin{aligned}P(t)=&[\lambda/\pi D(1-e^{-2\lambda t})]^{1/2}\\ &\times\exp[-2(\alpha-\bar{\alpha}_0e^{-\lambda t})^2/D(1-e^{-2\lambda t})]\ ,\end{aligned} \tag{3.24}$$

which is exactly the distribution function for a one-dimensional Gaussian random process obtained by other methods.[13]

B. Linear multidimensional process

We next consider a linear process involving more than one random variable. Thus we now have

$$\underline{G}=-\underline{M}\cdot\underline{\alpha}\ , \tag{3.25}$$

where $\underline{M}$ is a positive matrix. For the brevity of presentation we assume that dimension $p=2$. The third and higher moments (triple- and higher-order correlation functions) vanish for the same reason as for the one-dimensional case already discussed. Therefore the entropy production will be positive and P satisfies the boundary condition. Since $X_{21}^{(2)}=X_{12}^{(2)}$ owing to the symmetry of the correlation matrix $A^{(2)}$, there are three independent elements for the 2×2 second-rank tensor $X^{(2)}$. We therefore construct five-dimensional vectors $\underline{X}$ and $\underline{\Lambda}$:

$$\underline{X}=\{X_1^{(1)},X_2^{(1)},X_{11}^{(2)},X_{12}^{(2)},X_{22}^{(2)}\}\ , \tag{3.26a}$$

$$\underline{\Lambda}=\{\Lambda_1^{(1)},\Lambda_2^{(1)},\Lambda_{11}^{(2)},\Lambda_{12}^{(2)},\Lambda_{22}^{(2)}\}\ , \tag{3.26b}$$

and a 5×5 matrix

$$\underline{C}=\begin{bmatrix}\underline{D} & 0\\ 0 & \underline{R}\end{bmatrix}\ , \tag{3.26c}$$

$$\underline{R}=\begin{bmatrix}\{\underline{D}A^{(2)}\}_{1111} & 2\{\underline{D}A^{(2)}\}_{1112} & \{\underline{D}A^{(2)}\}_{1122}\\ \{\underline{D}A^{(2)}\}_{1211} & 2\{\underline{D}A^{(2)}\}_{1212} & \{\underline{D}A^{(2)}\}_{1222}\\ \{\underline{D}A^{(2)}\}_{2211} & 2\{\underline{D}A^{(2)}\}_{2212} & \{\underline{D}A^{(2)}\}_{2222}\end{bmatrix}\ , \tag{3.26d}$$

where

$$\begin{aligned}&\{\underline{D}A^{(2)}\}_{klmn}=D_{km}A_{ln}^{(2)}+D_{kn}A_{lm}^{(2)}+D_{lm}A_{kn}^{(2)}+D_{ln}A_{km}^{(2)}\ ,\\ &\Lambda^{(1)}=-\underline{M}\cdot A^{(1)}\ ,\\ &\Lambda^{(2)}=\underline{D}-(\underline{M}\cdot A^{(2)}+A^{(2)}\cdot\underline{M})\ .\end{aligned} \tag{3.27}$$

Therefore we find

$$\begin{aligned}&X^{(1)}=-2k_BT\underline{D}^{-1}\underline{M}\cdot A^{(1)}\ ,\\ &X_{11}^{(2)}=2k_BT\sum_{k=1}^{3}(\underline{R}^{-1})_{1k}\Pi_k\equiv k_BTK_{11}\ ,\\ &X_{12}^{(2)}=2k_BT\sum_{k=1}^{3}(\underline{R}^{-1})_{2k}\Pi_k\equiv k_BTK_{12}=k_BTK_{21}\ ,\\ &X_{22}^{(2)}=2k_BT\sum_{k=1}^{3}(\underline{R}^{-1})_{3k}\Pi_k\equiv k_BTK_{22}\ ,\\ &\underline{\Pi}=\{\Lambda_{11}^{(2)},\Lambda_{12}^{(2)},\Lambda_{22}^{(2)}\}\ ,\end{aligned} \tag{3.28}$$

where $\Lambda_{kl}^{(2)}$ are the elements of $\Lambda^{(2)}$ given in (3.27). In this particular case the evolution equations for moments are

$$\frac{dA^{(1)}}{dt}=-\underline{M}\cdot A^{(1)}\ ,$$
$$\frac{dA^{(2)}}{dt}=\underline{D}-(\underline{M}\cdot A^{(2)}+A^{(2)}\cdot\underline{M})\ , \tag{3.29}$$

which can be easily solved:

$$A^{(1)}(t)=\exp(-\underline{M}t)A_0^{(1)}\ ,$$
$$A^{(2)}(t)=\underline{H}^{-1}\underline{B}(t)\underline{H}\ . \tag{3.30}$$

Here $\underline{H}$ is a matrix such that

$$\underline{\lambda}=\underline{H}\,\underline{M}\,\underline{H}^{-1}=\begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix}$$

and B is a matrix whose elements are given by

$$B_{ij}(t)=\exp[-(\lambda_i+\lambda_j)t]B_{ij}^0+(\underline{H}\,\underline{D}\,\underline{H}^{-1})_{ij}(\lambda_i+\lambda_j)^{-1}\times\{1-\exp[-(\lambda_i+\lambda_j)t]\}\ ,$$

with B_{ij}^0 denoting the initial value of B_{ij}, i.e.,

$$B_{ij}^0=[\underline{H}A^{(2)}(t=0)\underline{H}^{-1}]_{ij}\ .$$

If $A_{ij}^{(2)}(0)=0$ for all i and j, then

$$B_{ij}(t)=(\underline{H}\,\underline{D}\,\underline{H}^{-1})_{ij}(\lambda_i+\lambda_j)^{-1}\{1-\exp[-(\lambda_i+\lambda_j)t]\} \tag{3.31}$$

and

$$\Lambda_{ij}^{(2)}=\sum_k\sum_l\underline{H}^{-1}_{ik}(\underline{H}\,\underline{D}\,\underline{H}^{-1})_{kl}\exp[-(\lambda_k+\lambda_l)t]\underline{H}_{lj}\ . \tag{3.32}$$

With these results the distribution function can be obtained in terms of $\underline{\alpha}$ and t. The α dependence of the distribution function for an arbitrary p is as follows:

$$P(t)=(\pi^p\det|\underline{L}+\underline{K}|)^{-1/2}\exp[-(\underline{L}+\underline{K}):(\underline{\alpha}-\bar{\underline{\alpha}})(\underline{\alpha}-\bar{\underline{\alpha}})]\ , \tag{3.24'}$$

where

$$\underline{L}=\underline{D}^{-1}\cdot\underline{M}$$

and

$$\underline{K}=\{2\underline{R}^{-1}\cdot\Lambda^{(2)}\}=\begin{bmatrix}K_{11} & K_{12}\\ K_{21} & K_{22}\end{bmatrix}\ ,$$

where K_{ij} are defined in (3.28). The matrix $\underline{K}$ depends on t, $\bar{\underline{\alpha}}$, and $A^{(2)}$. Therefore, the time dependence of P is described by (3.29) or (3.31) and (3.32). The result above agrees exactly with the solution of the Fokker-Planck equation for the linear process in question which can be obtained by other methods.[13]

C. Schloegl model

Schloegl[14] proposed to consider the model

$$A+2X\rightarrow 3X\ ,$$
$$X\rightarrow B\ ,$$

where the mole number of A and the total mole number of all components are held fixed by a suitable device of feeding A into, and drawing B out of, the system. Then the mole number x of component X is the only independent variable whose evolution is given by the equation

$$\frac{dx}{dt}=-(x-x_1)(x-x_2)(x-x_3)\ , \tag{3.33}$$

where x_1, x_2, x_3 may be three different values of x at the steady state. If x is regarded as a stochastic variable and a noise term is added to (3.33), there follows a Langevin equation from which we obtain a one-dimensional Fokker-Planck equation with

$$G=-(\alpha-\alpha_1)(\alpha-\alpha_2)(\alpha-\alpha_3)\qquad(\alpha_i=x_i,\quad i=1,2,3)\ . \tag{3.34}$$

It is important to note here that in obtaining the Fokker-Poanck equation we have not taken the master-equation approach[3(b)] in which a master equation is constructed from (3.33), but the Langevin equation approach in which a Langevin equation with a white-noise source is constructed. The master-equation approach has a different interpretation for diffusion tensor $\underline{D}$ which we do not adopt here. In this particular calculation we assume D is constant for the sake of simplicity of calculation. This assumption is not mandatory and easy to remove. We stress that the aim in this section is not in studying the Schloegl model itself, but in illustrating and examining the modified-moment method for solving the Fokker-Planck equation with the nonlinear flow term provided by the Schloegl model.

Defining

$$\Delta\alpha_i=\bar{\alpha}-\alpha_i=A^{(1)}-\alpha_i\quad(i=1,2,3)\ , \tag{3.35}$$

we obtain

$$\langle GP\rangle=-A^{(3)}-\sum_{i=1}^{3}\Delta\alpha_i A^{(2)}-\Delta\alpha_1\Delta\alpha_2\Delta\alpha_3\ . \tag{3.36}$$

This gives the right-hand side of the evolution equation for $A^{(1)}$. The rest of the evolution equations may be calculated from (3.5b) and (3.5c). We list them below:

$$\frac{dA^{(1)}}{dt}=-\left[A^{(3)}+\sum_{i=1}^{3}\Delta\alpha_i A^{(2)}+\Delta\alpha_1\Delta\alpha_2\Delta\alpha_3\right], \tag{3.37a}$$

$$\frac{dA^{(2)}}{dt}=-2\left[A^{(4)}+\sum_{i=1}^{3}\Delta\alpha_i A^{(3)}+\sum_{i<j}^{3}\Delta\alpha_i\Delta\alpha_j A^{(2)}\right]+D, \tag{3.37b}$$

$$\frac{dA^{(3)}}{dt}=-3\left[(A^{(5)}-A^{(2)}A^{(3)})+\sum_{i=1}^{3}\Delta\alpha_i(A^{(4)}-A^{(2)2})+\sum_{i<j}^{3}\Delta\alpha_i\Delta\alpha_j A^{(3)}\right], \tag{3.37c}$$

and for $m\geq 4$,

$$\frac{dA^{(m)}}{dt}=-m\left[(A^{(m+2)}-A^{(3)}A^{(m-1)})+\sum_{i=1}^{3}\Delta\alpha_i(A^{(m+1)}-A^{(2)}A^{(m-1)})+\sum_{i<j}^{3}\Delta\alpha_i\Delta\alpha_j A^{(m)}\right]$$
$$+2^{-1}m(m-1)DA^{(m-2)}. \tag{3.37d}$$

In the following calculation we choose $x_1=1$, $x_2=2$, and $x_3=3$. Then the steady states $x_1=1$ and $x_3=3$ of (3.33) are stable, whereas the steady state $x_2=2$ is unstable. With this choice for x_i we find

$$\sum_{i=1}^{3}\Delta\alpha_i=3A^{(1)}-6,$$

$$\Delta\alpha_1\Delta\alpha_2\Delta\alpha_3=(A^{(1)}-1)(A^{(1)}-2)(A^{(1)}-3),$$

$$\sum_{i<j}^{3}\Delta\alpha_i\Delta\alpha_j=(A^{(1)}-1)(A^{(1)}-2)+(A^{(1)}-2)(A^{(1)}-3)+(A^{(1)}-1)(A^{(1)}-3).$$

The stationary solution to the Fokker-Planck equation is then given by the formula

$$P_0=N^{-1}\exp[-2D^{-1}(\alpha^4/4-2\alpha^3+11\alpha^2/2-6\alpha)], \tag{3.38}$$

where

$$N=\int_0^\infty d\alpha\exp[-2D^{-1}(\alpha^4/4-2\alpha^3+11\alpha^2/2-6\alpha)].$$

As is shown in Fig. 1, P_0 has a symmetric bimodal distribution with peaks at $\alpha=1$ and 3.

Since the set of evolution equations in (3.37) is open, it must be suitably closed. The simplest way of closing it is to truncate the set at a suitable point and check the convergence of the moment series. We find, for this particular model, the following closure condition most attractive:

$$A^{(m+2)}=A^{(3)}A^{(m-1)}\quad(m\geq 3). \tag{3.39}$$

In addition to this condition, we truncate the moment series at $m=4$ inclusive. This yields the following four coupled nonlinear equations:

$$\frac{dA^{(1)}}{dt}=-A^{(3)}-\sum_{i=1}^{3}\Delta\alpha_i A^{(2)}-\Delta\alpha_1\Delta\alpha_2\Delta\alpha_3, \tag{3.40a}$$

$$\frac{dA^{(2)}}{dt}=-2\left[A^{(4)}+\sum_{i=1}^{3}\Delta\alpha_i A^{(3)}+\sum_{i<j}^{3}\Delta\alpha_i\Delta\alpha_j A^{(2)}\right]+D, \tag{3.40b}$$

$$\frac{dA^{(3)}}{dt}=-3\left[\sum_{i=1}^{3}\Delta\alpha_i(A^{(4)}-A^{(2)2})+\sum_{\substack{i,j\\ i<j}}^{3}\Delta\alpha_i\Delta\alpha_j A^{(3)}\right], \tag{3.40c}$$

$$\frac{dA^{(4)}}{dt}=6DA^{(2)}-4\sum_{\substack{i,j\\ i<j}}\Delta\alpha_i\Delta\alpha_j A^{(4)}. \tag{3.40d}$$

A more mathematically sensible closure condition which we can have but have not used in what follows is the cluster expansion in which the connected clusters (i.e., correlation functions) of order $(m+2)$ or higher are neglected $(m\geq 3)$. More specifically, we can decompose $A^{(m+2)}$ $(m\geq 3)$ into clusters:

$$A^{(m+2)}=A^{(1)}A^{(m+1)}+A^{(2)}A^{(m)}+\cdots+A^{(1)2}A^{(m)}+\cdots+A^{(m+2)}_{\text{connected}}.$$

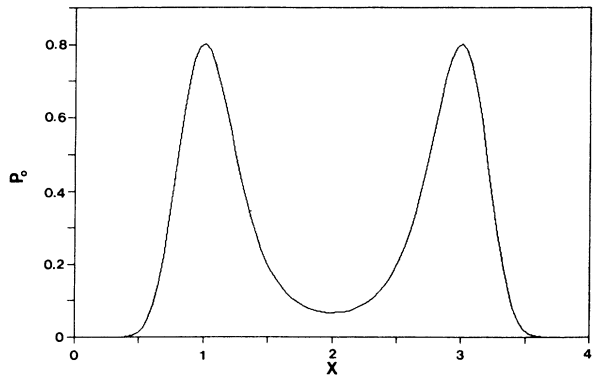

FIG. 1. Stationary distribution function vs x (concentration) for $D=0.16$. It has a bimodal symmetry.

Diagrammatically speaking, if we consider $(m+2)$ points, the last term on the right-hand side represents a connected diagram, while the rest represents disconnected diagrams. Thus, if the connected diagrams of $(m+2)$ points are neglected, the following closure relation holds:

$$A^{(m+2)}=(m+2)!\sum_k{}' \prod_{k=1}^{m+2} A^{(k)l_k}/l_k! ,$$

where the sum over k is subject to the condition $\sum_k kl_k=m+2$ and the prime on the summation sign means the exclusion of the diagrams connecting all $(m+2)$ points where $m\geq 3$. This last condition $m\geq 3$ stems from the fact that the moment series is truncated at $m=4$. If the set is truncated at $m=5$, then the condition is $m\geq 4$ and so on.

To test the accuracy of the truncation scheme (3.39) we will also consider the cases of $m=5$ and 6 later on, in which cases there are, respectively, five and six coupled equations that are structurally similar to those in (3.40).

We solve (3.40) by a sixth-order Runge-Kutta method. In this numerical investigation the diffusion constant is varied. It in fact plays the role of order parameter on which the dynamical behavior of the system is sensitively dependent. Before we start the numerical solution, it is useful to observe that the evolution equations (3.40) remain invariant to the transformation

$$\begin{aligned} &A^{(1)}\rightarrow 4-A^{(1)} ,\\ &A^{(2)}\rightarrow A^{(2)} ,\\ &A^{(3)}\rightarrow -A^{(3)} ,\\ &A^{(4)}\rightarrow A^{(4)} . \end{aligned} \tag{3.41}$$

It is easily verified that the fullest of evolution equations (3.37) also has an analogous symmetry property

$$\begin{aligned} &A^{(1)}\rightarrow 4-A^{(1)} ,\\ &A^{(i)}\rightarrow A^{(i)} \text{ for } i=\text{even}, \quad i>1 ,\\ &A^{(i)}\rightarrow -A^{(i)} \text{ for } i=\text{odd}, \quad i>1 . \end{aligned} \tag{3.42}$$

Next, we calculate the steady states of the set (3.40) and analyze their linear stability. For this purpose it is convenient to introduce

$$u=A^{(1)}-2=x-2, \quad y=A^{(2)} ,$$
$$z=A^{(3)}, \quad w=A^{(4)} .$$

Then the steady states of (3.40) are determined by the algebraic equations

$$\begin{aligned} &z+3uy+u(u^2-1)=0 ,\\ &D-2[w+3uz+(3u^2-1)y]=0 ,\\ &3u(w-y^2)+(3u^2-1)z=0 ,\\ &3Dy-2(3u^2-1)w=0 . \end{aligned} \tag{3.43}$$

These equations lead to the following equations:

$$u[u^{12}-5u^{10}/3+(\tfrac{7}{9}-D/6)u^8-(\tfrac{2}{27}-13D/18)u^6-(\tfrac{5}{81}+7D/27+2D^2/9)u^4+(\tfrac{7}{243}+11D/162+D^2/18)u^2$$
$$-(\tfrac{1}{243}+D/54+D^2/81-D^3/72)]=0 , \tag{3.44a}$$

$$y=[18u^6-24u^4+(3D+6)u^2-D]/[3D+2+6u^2-36u^4] , \tag{3.44b}$$

$$z=-[3uy+u(u^2-1)] , \tag{3.44c}$$

$$w=3Dy/2(3u^2-1) . \tag{3.44d}$$

One obvious solution is

$$u=0 ,$$

which yields

$$x=A^{(1)}=2, \quad y=z=w=0 , \tag{3.45}$$

which is unstable and independent of D. There are an additional 12 solutions to (3.44a). It turns out that for $D\leq 0.17$ there are six real roots satisfying the symmetry (3.41). As D gets larger than 0.17, all but a pair of solutions become complex, but the surviving real roots are unstable. The trajectories of the solutions are indicated in Fig. 2 which we find insightful, since the distribution of roots gives a useful qualitative idea of how the solutions of the evolution equations will behave. The linear stability of the roots is indicated with letter u (unstable) and s (stable) in Fig. 2. In this work we examine the convergence of solutions when $D<0.17$. Since for $D>0.17$ the behavior of the solutions is rather rich in variety and

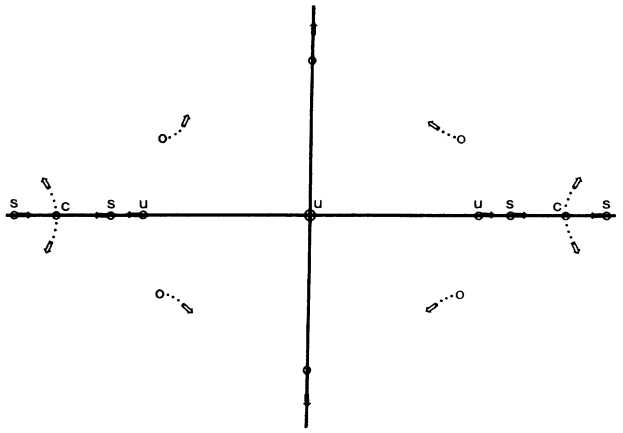

FIG. 2. Trajectories of steady states as the order parameter D is increased. The abscissa is the real axis and the ordinate is the imaginary axis. The letter s stands for a stable steady state, while the letter u is for an unstable state. At point c two steady states merge to a double root. A steady state is located at the origin but it does not move as D is changed. There are 13 roots in all, if the trivial root is included.

appears to warrant a more extensive treatment, we will report the results elsewhere.

For dynamical solutions of (3.40) we thus examine in detail only the subcritical case of $D < D_c \simeq 0.17$. In this regime of D the solutions of (3.40) are sensitive to the initial conditions $\{A^{(m)}\}$ which determine which one of four stable states of (3.40) shown in Fig. 2 is approached as time tends to infinity. Since $\Lambda^{(m)}=0$ at the steady states of the evolution equations and $X^{(m)}$ is directly proportional to $\Lambda^{(m)}$ according to (3.17), the distribution function should be approaching P_0 in long time in this case. Therefore in the subcritical regime of D the dynamics of Brownian motion becomes irrelevant to the long-time distribution of stochastic variables. This is in agreement with the well-known notion of a nonequilibrium system approaching equilibrium. Nevertheless, the nonequilibrium part of the distribution function should play an important role for transient phenomena in the system, and it can be obtained from (3.40), (3.2), (3.15a), and (3.17). Since the main aim in this calculation is to ascertain the validity of truncation in the moment sequence, we check the convergence of the sequence by calculating the moments in three different cases of truncation: $m_{\max}=4$, 5, and 6. There are four evolution equations, i.e., (3.40), in the case of $m_{\max}=4$, five evolution equations in the case of $m_{\max}=5$, and six evolution equations in the case of $m_{\max}=6$. The solutions to these three sets of evolution equations are compared in Table I. The results show that the moment sequence converges fairly rapidly and the truncation at $m=4$ is justifiably sufficient for all practical ends in the case of the model considered. It is verified that the moment sequence also converges in the case of other initial conditions. In this manner the distribution function $P(t)$ is verified to approach the stationary (equilibrium) solution P_0 of the Fokker-Planck equation which has a bimodal symmetry. It is found that there exists an interval of D values in which the solution of the evolution equations becomes a limit cycle. However, it is not known at the present time whether or not this kind of oscillatory solution is consistent with the exact solution of the Fokker-Planck equation for the Schloegl model. We hope to report on this aspect in the near future.

IV. ON THE QUESTION OF STABILITY THEORY

The question of whether or not it is possible to deduce a stability theory for dynamical systems is interesting and, perhaps, even important. Glansdorff and Prigogine[15] proposed a theory in connection with the question, and there has been in the past a controversy[16] surrounding it. Our intention here is not to resolve the controversy, but to examine the behavior of the entropy presented in Sec. III, since the present theory supplies the tool for a mathematically sensible way of discussing the question.

The nonequilibrium part S_n of the entropy S is bounded from above and especially there holds the inequality

$$S_n \leq 0 , \tag{4.1}$$

which can be easily proven by writing

$$S_n = -k_B \langle P \ln(P/P_0) - P + P_0 \rangle ,$$

since $\langle P \rangle = \langle P_0 \rangle$, and by using the inequality[17]

$$x \ln(x/y) - x + y \geq 0 .$$

Since by the H theorem

$$\frac{dS_n}{dt} \geq 0 , \tag{4.2}$$

it is tempting to regard S_n as a Lyapunov function,[18] but since (4.2) holds even for the unstable steady state of processes $\{A^{(m)}\}$ for which dS_n/dt must be negative if it is a Lyapunov function, S_n cannot qualify as a Lyapunov function. But the second variation of S_n can be regarded as a Lyapunov function, subject to a limitation, as will be shown.

TABLE I. Convergence test for the moment sequence. $D=0.1$. A similarly convergent result is obtained for the other steady state in the vicinity of $A^{(1)}=1$. Note that there are four stable steady states as shown in Fig. 2 in the subcritical region.

N	$A^{(1)}$	$A^{(2)}$	$A^{(3)}$	$A^{(4)}$	$A^{(5)}$	$A^{(6)}$
4	2.9518	0.0324	−0.0029	0.0028		
5	2.9514	0.0327	−0.0032	0.0030	−0.0002	
6	2.9448	0.0359	−0.0056	0.0051	−0.0016	0.0007

Let us consider a variation of S_n around the steady state. Recall that the steady state of a macroscopic process coincides with the steady state of the Fokker-Planck equation according to the proposition presented in Sec. III. Taking such a variation, we write

$$S_n = S_n^{\rm st} + \delta S_n + \delta^2 S_n + \cdots \ . \tag{4.3}$$

At the steady state

$$S_n^{\rm st} = 0$$

and the first and the second variations are, respectively, as follows:

$$\delta S_n = \sum_m (\partial S_n / \partial A^{(m)})_{\rm st} \odot \delta A^{(m)} = \sum_m \bar{X}^{(m)}_{\rm st} \odot \delta A^{(m)} \ , \tag{4.4a}$$

$$\delta^2 S_n = \tfrac{1}{2} \sum_{m,l} S^{\rm st}_{ml} \odot \delta A^{(l)} \delta A^{(m)} \ , \tag{4.4b}$$

where the subscript or superscript "st" means the steady-state value, and

$$S^{\rm st}_{ml} = (\partial^2 S_m / \partial A^{(l)} \partial A^{(m)})_{\rm st} \ .$$

Since at the steady state

$$X^{(m)}_{\rm st} = 0 \quad \text{for all } m$$

by the proposition in Sec. III, the first variation δS_n vanishes at the steady state. However, the second variation is not always equal to zero and hence it is the first nonvanishing variation. Since the higher-order variation then may be neglected, we find

$$\delta^2 S_n \le 0 \ , \tag{4.5}$$

owing to the inequality (4.1). Differentiating it with t, we obtain

$$\frac{d}{dt} \delta^2 S_n = \sum_{m,l} S^{\rm st}_{ml} \odot \delta A^{(l)} \delta A^{(m)} \ . \tag{4.6}$$

Since

$$(\partial \delta^2 S_n / \partial \delta A^{(m)}) = \sum_l S^{\rm st}_{ml} \odot \delta A^{(l)} \ ,$$

we may write (4.6) in the form

$$\frac{d}{dt} \delta^2 S_n = \sum_m (\partial \delta^2 S_n / \partial \delta A^{(m)}) \odot \frac{d}{dt} \delta A^{(m)} \ , \tag{4.7}$$

which means that $\delta^2 S_n$ has the Eulerian derivatives[18] as a Lyapunov function should.

If in the vicinity of a steady state

$$\frac{d}{dt} \delta A^{(m)} = \sum_l L_{lm} \odot \delta A^{(l)} \ , \tag{4.8}$$

where L_{lm} are constants, then

$$\frac{d}{dt} \delta^2 S_n = \sum_k \sum_l \left[\sum_m S^{\rm st}_{ml} L_{km} \right] \odot \delta A^{(k)} \delta A^{(l)} \equiv \sum_{k,l} M_{lk} \odot \delta A^{(k)} \delta A^{(l)} \ . \tag{4.9}$$

The matrix $(S^{\rm st}_{ml})$ is negative definite according to the inequality (4.5). If the steady state is stable, then the matrix (L_{km}) must be negative definite and therefore the composite matrix (M_{lk}) is positive definite. In that case

$$\frac{d}{dt} \delta^2 S_n \ge 0 \ . \tag{4.10}$$

On the other hand, if the steady state is unstable, the matrix (L_{km}) is positive and hence the matrix (M_{lk}) is negative. In this case there holds the inequality

$$\frac{d}{dt} \delta^2 S_n < 0 \ . \tag{4.11}$$

Therefore we see that $\delta^2 S_n$ may be regarded as a Lyapunov function and it is possible to tell from $\delta^2 S_n$ whether the macroscopic process in question is stable or not. This conclusion is valid in the case of the Fokker-Planck equation. The result above is in agreement with that obtained by Keizer,[12,19] who used his Σ function. If the evolution equation for P is not the Fokker-Planck equation the same conclusion cannot necessarily be drawn as for the case with the Boltzmann equation for spatially inhomogeneous systems. The conclusion contained in (4.10) and (4.11), however, is about the local stability since the linearized equation (4.8) is valid locally. Such a linearization gives a sensible idea of the evolution of the system if the trajectories are monotonic and confined in the neighborhood of the steady state in question. However, it is often seen in macroscopic dynamics that trajectories can wildly and chaotically oscillate when there are homoclinic points[20] at which stable and unstable lines intersect. Since such circumstances cannot be described by linearized equations such as (4.8), the criteria (4.10) and (4.11) cannot be counted on as even a rough guide for predicting the evolution of a system. In the case of the Schloegl model we have seen some numerical counterexamples against criterion (4.11) which show that the trajectories can return to the vicinity of unstable steady states. (These examples are not presented here in view of their preliminary nature and we will communicate a full account of the study elsewhere.) The conclusion implied by (4.11), however, is that the trajectories cannot return to unstable steady states. From these considerations we can infer that (4.10) and (4.11) are local criteria.

V. CONCLUDING REMARKS

We have presented a method of solving a Fokker-Planck equation by applying the modified-moment method.[9] This method provides an attendant theory for the entropy and entropy production for stochastic irreversible processes. This thermodynamic formalism allows us to calculate the entropy of stochastic processes and its temporal evolution. This method reduces the solution procedure for the Fokker-Planck equation—a partial differential equation—to that of algebraic equations for the entropy derivatives

$$(\partial S / \partial A^{(m)}) = \bar{X}^{(m)} / T$$

and of ordinary differential equations for the moments

chosen, which include the evolution equation for the average values of the stochastic variables themselves. The ordinary differential equations for $A^{(m)}$ are generally nonlinear and therefore are seen to have the characteristic features often seen in nonlinear ordinary differential equations. Since the moments $A^{(m)}$ appear in the distribution function in the present method, their temporal evolutionary characteristics are directly transferred to the distribution function. In this work we have shown that the modified-moment method yields the same results as other methods in the case of linear processes and that in the case of the nonlinear model considered, the moment sequence converges sufficiently fast to justify the truncation of the moment series at a lower order, if $D < D_c$ for the model considered.

If the potential condition is not met, it will be a lot more difficult to find the function $V(\underline{\alpha})$. Nevertheless, the formalism developed in this paper holds in its essence if it is assumed that there exists a function $V(\underline{\alpha})$ satisfying the steady-state Fokker-Planck equation in the manner as in (2.5a). The question of whether there exists such a function must be answered within the context of thermodynamics, but the complete answer is not available as yet.

The structure of the theory of entropy presented is similar to that we see in thermodynamics, and this similarity stems from the common statistical formula for entropy which the present stochastic theory shares with the canonical-ensemble theory in statistical mechanics.[11] The distinctive feature of the modified-moment method lies in the role given to this statistical formula for entropy to play in the solution procedure for the Fokker-Planck equation. We must particularly emphasize the importance of the condition the entropy production imposes on the acceptable approximation for $X^{(m)}$. This condition assures that the approximate solution taken for $X^{(m)}$ will satisfy the positivity of the entropy production, which is generally taken as the second law of thermodynamics. Since the average behavior of stochastic variables must conform to the second law of thermodynamics, the formalism helps us develop inevitable approximate solutions so as to satisfy such a law. The examples for linear processes considered show that the method leads to correct results.

Finally, it seems worthwhile to point out that the quantity used in the stability theory presented is the second variation of S_n, but not the total entropy itself. Therefore $d\delta^2 S_n/dt$ is a variational entropy production in reference to the steady state. This quantity is different from $d\delta^2 S/dt$, which does not quality as a Lyapunov function. It must be stressed that the evolution criteria (4.10) and (4.11) are local criteria, not global ones.

ACKNOWLEDGMENTS

This work was initiated when one of us (B.C.E.) spent a part of his sabbatical leave at Stanford University. He would like to thank Professor John Ross for his hospitality and many valuable discussions and comments. We would also like to thank Sheeba Khan for assistance in computing the numerical solutions. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada.

*Also at the Department of Physics, McGill University.

[1]See, for example, (a) S. Chandrasekhar, Rev. Mod. Phys. **15**, 1 (1943); (b) M. Chen Wang and G. E. Uhlenbeck, *ibid.* **17**, 323 (1945). There are numerous monographs on stochastic theories and Fokker-Planck equations. Some of them are listed in Ref. 2 below.

[2](a) R. L. Stratonovich, *Topics in the Theory of Random Noise* (Gordon and Breach, New York, 1963); (b) J. L. Doob, *Stochastic Processes* (Wiley, New York, 1960).

[3](a) I. Oppenheim, K. E. Shuler, and G. H. Weiss, Adv. Mol. Relaxation Processes **1**, 13 (1967); (b) C. W. Gardiner, *Handbook of Stochastic Methods* (Springer, Heidelberg, 1983).

[4]R. Graham, in *Quantum Statistics in Optics and Solid-State Physics*, Vol. 66 of *Springer Tracts in Modern Physics*, edited by G. Höhler (Springer, Heidelberg, 1973), p. 1.

[5]H. Grabert and M. S. Green, Phys. Rev. A **19**, 1747 (1979).

[6]B. C. Eu, Physica A **90**, 283 (1978).

[7]K. C. Hunt and J. Ross, J. Chem. Phys. **75**, 976 (1981).

[8]L. Onsager and S. Machlup, Phys. Rev. **91**, 1505 (1953); S. Machlup and L. Onsager, *ibid.* **91**, 1512 (1953).

[9]B. C. Eu, J. Chem. Phys. **73**, 2958 (1980); (Erratum) **79**, 3607 (1983); **87**, 1220 (1987). See also, B. C. Eu, J. Nonequil. Thermodyn. **11**, 211 (1986); Acc. Chem. Res. **19**, 153 (1986).

[10]W. Haack and W. Wendland, *Partial and Pfaffian Differential Equations* (Pergamon, Oxford, 1972).

[11]T. L. Hill, *Statistical Mechanics* (McGraw-Hill, New York, 1960).

[12]J. Keizer, *Statistical Thermodynamics of Nonequilibrium Processes* (Springer, Berlin, 1987).

[13]See, for example, Ref. 1(a).

[14]F. Schloegl, Z. Phys. **248**, 446 (1971).

[15]P. Glansdorff and I. Prigogine, *Thermodynamics of Structure, Stability and Fluctuations* (Wiley, New York, 1971).

[16]J. E. Keizer and R. F. Fox, Proc. Natl. Acad. Sci. U.S.A. **71**, 192 (1974); P. Glansdorff, G. Nicolis, and I. Prigogine, *ibid.* **76**, 214 (1979); R. F. Fox, *ibid.* **76**, 2114 (1979); G. Nicolis and I. Prigogine, *ibid.* **76**, 6060 (1979); B. H. Lavenda and E. Santamato, Lett. Nuovo Cimento **33**, 559 (1982).

[17]If we use a new variable u defined by $y = ux$, then $x\ln(x/y) - x + y = x(u - 1 - \ln u)$. For $x > 0$, $u - 1 \geq \ln u$, which is easy to show by setting $f(u) = u - 1 - \ln u$ and providing that $f(u)$ has a minimum at $u = 1$. Then since $f(u) = 0$ at $u = 1$, it follows that $f(u) \geq 0$.

[18]N. Minorsky, *Nonlinear Oscillations* (Van Nostrand, Princeton, NJ, 1962).

[19]J. Keizer, J. Chem. Phys. **65**, 4431 (1976); **69**, 2609 (1978).

[20]H. G. Schuster, *Deterministic Chaos* (Physik-Verlag, Weinheim, 1984).