Journal of Statistical Physics, Vol. 11, No. 2, 1974

Probabilistic Kinetics of Macroprocesses in Broken Microscopic Reversibility

Koichiro Matsuno[1]

Received April 9, 1974

Probabilistic kinetics following the Pauli master equation without microscopic reversibility determines an asymptotic structure of macroprocesses in a coarse-grained phase space of many degrees of freedom. The structure, which is asymptotically realized, minimizes its irreversible decay rate among various candidates. This least irreversible decay rate is consistent with the assertion for the minimum K-entropy which has been argued to apply to the nonequilibrium asymptotic state. The irreversible decay rate is a state function characteristic of macrostructure on a coarse-grained time scale. Macrofluctuations, which always appear around the asymptote as fluctuations of the state function, do not obey the central limit theorem, implying that fluctuations whose characteristic times are not less than some finite value are never excluded.

KEY WORDS: Detailed balance; K-entropy; master equation; microscopic reversibility.

1. INTRODUCTION

1.1. Introductory Remarks

Near the thermal equilibrium point the kinetic equation used for describing the probabilistic kinetics of macroprocesses, which result following a certain projection eliminating microscopic motion,[1,2],[2] may sometimes be subject to

[1] Central Research Laboratories, Nippon Electric Company, Ltd., Kawasaki, Japan.

[2] Although a linear equation of collective variables is obtained with the use of a certain projection in Ref. 1, a nonlinear kinetic equation follows as for the distribution of collective variables. In the latter scheme the structure of the distribution is of more

the microscopic reversibility of Onsager.[3] In particular, if the microscopic reversibility, that is, the symmetry between conjugate transition probabilities, is kept in the Pauli master equation, the probability measure of finding a system in any one of the available macrostates will be equal, being consistent with the principle of equal weight (see Ref. 4, especially Appendix I). However, it does not always occur that various kinetic processes off equilibrium are in accord with microscopic reversibility.

Tomita and Tomita[5] have recently pointed out that a nonlinear Markov process does not yield a strict detailed balance[6],[3] based upon microscopic reversibility and they proposed cyclic balance instead. Hence one may ask what sort of macroscopic characteristics *on the whole* would be available if microscopic reversibility does not hold. This is the problem we investigate in this paper.

It is of fundamental importance for the study of kinetics on a macroscale to establish a probabilistic kinetic equation by applying a certain projection to microscopic mechanics, e.g., the Liouville equation.[2],[4] Examples of kinetic equations include the Fokker–Planck equations derived for a single-mode laser oscillation,[8] the Navier–Stokes fluid system,[9–11] and interacting phonons.[12],[5] Although such a kinetic equation must be found for each practical case, there still exists a macroscopic argument based upon the Markovian postulate. In particular, the Kramers–Moyal expansion[13] applied to the master equation reveals various nonlinear phenomena off equilibrium for the case that the relevant macrovariables are extensive.[5,14] However, we shall not argue the derivation of the probabilistic kinetic equation from microscopic mechanics in spite of its importance, and rather shall discuss what macroscopic properties could be expected supposing that microscopic reversibility does not hold for a given kinetic equation such as the Pauli master equation.

We shall concentrate on the asymptotic behavior of macroprocesses in broken microscopic reversibility and on their fluctuations, because of the

interest than the collective variables themselves. If one is allowed to divide various kinetic stages into two sets of slow and fast processes and also to suppose that only the slow process is responsible for the structural change in the distribution function, the master equation will result. A crucial aspect of this argument is seen in the assertion that the fast process may not interfere with the slow one. Also see Chapter 4 of Ref. 27.

[3] The classical Stosszahlansatz giving the probability distribution as proportional to phase space volume results in detailed balance only in the linear approximation.

[4] Some authors, including, for example, Ueyama,[7] have obtained the master equation in the so-called $\lambda^2 t$ limit within the lowest perturbation expansion. It is, however, not clear whether the resulting master equation could deal with any slow process, some contributions of which are critically examined in Ref. 2. Also see Section 8 of the present paper and footnote 10.

[5] A Fokker–Planck equation follows from a nonlinear Langevin equation of many coupled modes with the use of the Gaussian–Markov ansatz. Also see Ref. 27.

complexities of the subject. The kinetics prior to reaching the asymptote is beyond the scope of this paper.

Macroprocesses will be supposed to have a finite number of degrees of freedom. We shall further suppose that the dynamics is classical. Extension to the quantal case will be discussed later.

Since only observable quantities are concerned, it would be impossible for a limited observer to locate the representative point of the finite system in phase space, that is, Γ space, in a pointwise manner. A coarse graining of Γ space necessarily follows. One may thus regard Γ space as a set of partitions in such a manner that as long as the representative point is found inside a given local element of the partitioned phase space, one could assign to the point only the name or the number of the coordinate fixed to the element and could not tell where inside the element the point is really found. The coordinate of the representative point in Γ space becomes discrete instead of being continuous because of the limitations of an observer who is interested only in macroprocesses *as a whole*. At this point one should note that the way of partitioning Γ space is specific to the observer and is by no means prepared a priori.

One may imagine another limitation which inevitably accompanies the observer. Observation always takes place with the aid of an operation which integrates events appearing during a predetermined time interval which never vanishes and which is characteristic of the observer. Hence it would not be possible for the observer to track events occurring during a time interval less than the predetermined one. When the transition dynamics of the representative point is followed in the partitioned Γ space the observer cannot precisely keep up with a fast process in which the point transits from a partitioned element to another within a time less than the predetermined interval given as a coarse-grained time unit.

As a result, two kinds of coarse graining necessarily follow. One is for the partitioning of Γ space. The other is for time measurement. When asymptotic behavior and its fluctuations are investigated for a system with a finite number of degrees of macroscopic freedom, coarse grainings for both phase space and time must be employed because of the finiteness of the observer. We discuss this in Section 2 in a stricter form which will be of use in the following sections. The coarse grainings which inevitably accompany a finite observer will be stated as a definition. A general theorem derived from the definition will play a central role in the remaining discussion.

1.2. Brief Summary

A nonequilibrium asymptotic structure of a system with a finite number of degrees of freedom has been argued to occur at a minimum point of the Krylov–Kolmogorov dynamic entropy or, equivalently, the Kolmogorov–

Sinai entropy, which is usually called the K-entropy.[15,16] Although the original derivation is based upon a somewhat heuristic argument, the minimum K-entropy will be rederived in Section 3 following the theorem of Section 2. The K-entropy is simply a measure indicating the degree of stochastic instability, i.e., mixing, as one meets in the ergodic problem of orbits in a dynamical system.[17,18] An irreversible transition of the representative point in the coarse-grained Γ space is a specific example of general mixing phenomena, that is to say, a process of losing information. Here we use the term information in the sense of the Kolmogorov entropy of partition, which is a generalization of Shannon's original usage.[19] The stated minimum K-entropy implies that the asymptotic structure would seem to a finite observer to be realized at a possible least amount of the associated K-entropy, which is another way of saying the rate of losing information characteristic of the macrostructure through irreversible processes.

The theorem of Section 2 enables us to find an asymptotic solution to the Pauli master equation of a representative point moving in the coarse-grained Γ space. The solution is presented in Section 4. One observes that the asymptotic structure comprising macroprocesses without microscopic reversibility is realized in such a way that the representative point remains *mostly* inside a specific element of the coarse-grained Γ space. The specific element is found to minimize the net transition probability per unit time for the representative point to transfer from inside to outside the element. This in turn proves and justifies the statement of the minimum K-entropy for the case that probabilistic kinetics is governed by the Pauli master equation, since the net transition probability per unit time is just the rate of losing information about the location of the representative point, apart from a numerical factor.

Macroprocesses without microscopic reversibility asymptotically constitute a macrostructure in which the representative point is found mostly in a small element of Γ space accompanying the minimum rate of losing information. On the other hand, macroprocesses subject to microscopic reversibility give a macrostructure in which the point is found everywhere inside Γ space with equal probability, that is, the net transition probability per unit time from inside to outside a partitioned element is the same as that for the reverse transition. This would give an intuitive ground for understanding why a nonequilibrium macrostructure without microscopic reversibility carries specific information compared with structures satisfying the reversibility which would be common to most near-equilibrium phenomena.[6]

[6] The master equation resulting from the Hermitian ansatz of the interaction Hamiltonian exhibits microscopic reversibility (cf. Ref. 7). However, this would not be quite true since, if either a completely or an almost isolated system is the case, the adiabatic interaction on a macroscopic scale would no longer be Hermitian. Further discussion is given in Section 8 and in footnote 10.

One more specific aspect is that macroscopic fluctuations of structure around its asymptote do not obey the central limit theorem in the case of broken microscopic reversibility.

Applications and discussions of a general nature are presented from Section 5 on. Macroprocess with one degree of freedom are discussed in Section 5 within the scheme of the Fokker–Planck equation with one variable, as a specific example of the Pauli master equation. If the Fokker–Planck equation does not exhibit microscopic reversibility in its associated Γ space subject to coarse graining, macrofluctuations whose correlation times are not less than some finite value would never be excluded. One thus finds that such macrofluctuations are not subject to the central limit theorem and that these non-Gaussian fluctuations are extraordinarily enhanced near a critical point. Agreement with experiments is also confirmed to a certain extent. Macrofluctuations are divided into two categories, linear and nonlinear.[20] Nonlinear macrofluctuations associated with nonlinear decay exhibit a logarithmic singularity in their spectral representation, showing a distinct contrast with linear macrofluctuations.

When macroprocesses include two degrees of freedom or more a certain phase relationship will generally appear among them. We discuss the case of two degrees of freedom in Section 6. Absolute phase is not a directly observable quantity since one can measure only a relative phase difference from a certain reference point. In particular, phase dynamics can be described in terms of its time derivative, which is nothing but frequency. Macrofluctuations in a relative phase difference or, equivalently, in frequency could thus be fixed with the aid of the probabilistic kinetics of the macroprocesses without microscopic reversibility.

As macroprocesses increase the number of coupled degrees of freedom to much greater than unity, a limited observer could not precisely follow each phase relation, because of the highly intricate nature of the couplings. One way in which a certain macroscopic observation could be made would be not to try to obtain information about each phase relation but to regard the macroprocesses as being in a random phase relation with each other. A brief discussion on this subject is presented in Section 7. One must note, however, that the random phase characteristic thus introduced is artificial and depends entirely upon the finiteness of the observer.

One may think that probabilistic kinetics of macroprocesses having a large number of degrees of freedom could also deal with near-equilibrium phenomena. Consider a completely isolated system with N ($\gg 1$) degrees of freedom. If one can think of an appropriate projection which could produce a probabilistic kinetic equation with N' ($\lesssim N$) degrees of freedom eliminating only $N - N'$ ($\ll N$) independent variables, the equation would describe the near-equilibrium phenomena of the *almost* isolated system after Landau.[21]

We discuss in Section 8 that the probabilistic kinetic equation of the *almost* isolated system could not exhibit microscopic reversibility because adiabatic processes on a macroscale are characteristic of the completely isolated system. This would suggest that even in the near-equilibrium case there exists a certain kinetic process which does not exhibit symmetry between conjugate transition probabilities. The microscopic reversibility of Onsager[3] is not an only account for the macrokinetics even near equilibrium. Its drastic exemplification would be expected near the critical point of an equilibrium phase transition.

All of the results thus obtained are based upon the viewpoint that an observer making theoretical predictions and practicing experiments is not unlimited with respect to both types of performance. When macroprocesses with many coupled degrees of freedom are prepared, one might wonder how each observer could fix his own way of coarse graining, and how common and objective agreements could be attained among different observers. Difficulties may lie not with the nonlinearity between observer and object as pointed out by Wigner[22] but rather with the finiteness of the observer. In spite of its fundamental importance, this problem is left unsolved.

2. GENERAL THEOREM

We first introduce a phase space, that is, Γ space of classical macroprocesses having N degrees of freedom. Because of the finiteness of the observer, the resulting coarse graining divides Γ space into a set of partitions

$$\{S_i\} = \Gamma, \qquad i = 1, 2, \ldots \tag{1}$$

Although the way of partitioning is arbitrary depending upon the manner of observation, a particular partition in (1) will be supposed for now. When the representative point is found inside a partitioned element S_i the observer assigns to the point only the coordinate number i and nothing else. Hence the subscript of the element is necessary and sufficient for describing the state of the representative point.

We next introduce a coarse graining of time which could apply only to the observer. For this purpose a measure

$$\mu[S_i; T, \tau_0] \equiv [T]/T \tag{2}$$

which is specific to the observer, is defined. The meanings of the symbols are as follows: We denote as a unit-continuous-event (UCE) of coordinate i the period that the representative point remains at i from the moment the point transits to i until the moment it leaves from i; $[T]$ is the sum of the time durations of the UCE's of coordinate i during the time T. The time duration of a UCE of i is counted as zero if it is less than the coarse-grained time unit

τ_0 employed by the observer. $[T]$ is consecutively counted in units of τ_0 over a sequence of time T. If two successive UCE's of i occur with time durations t_1 and t_2, respectively, such that the time interval t' between them is $< \tau_0$, the successive events are regarded as if a single UCE of i with the time duration $t_1 + t' + t_2$ had occurred only in the case that $[T]$ is counted.

Since the asymptotic behavior of a macrostructure with N degrees of macroscopic freedom and the associated fluctuations will be treated, we must define the term macrostructure in a strict form.

Definition. A macrostructure S_i which continues to exist during the time $T\ (\geq \tau_0)$ is fixed by the statement

$$\mu[S_i; T, \tau_0] = 1 \tag{3}$$

where the asymptotic macrostructure is obtained in the limit $T \to \infty$.

This definition means that even if the representative point once at i leaves i, it returns by some means within a time interval less than τ_0, whose size will be discussed for each practical case in the following sections. The observer is supposed to be unable to follow the fast transition dynamics of the point within such a *small* time interval less than τ_0.

Let us introduce the probability measure $\varphi_i(t_1)P_i(t_2, t_1)$ for the two successive events. Namely, the event that the representative point is at coordinate i at time $t = t_1$ is followed by another event that the point remains at the same coordinate until $t = t_2\ (> t_1)$ without undergoing any transition, where

$$\varphi_i(t_1) \equiv \begin{cases} 1 & \text{if the representative point is found at } i \text{ at } t = t_1 \\ 0 & \text{otherwise} \end{cases}$$

We implicitly assume the presence of probabilistic kinetics of the representative point in the coarse-grained Γ space. In general, the probability measure $P_i(t_2, t_1)$ is not unaffected by the previous memory before $t = t_1$. Furthermore, we introduce another quantity

$$P_c(S_j|S_i; T, \tau_0) \equiv \varphi_j(t')P_j(T + t', t')/P_i(T + t'', t'') \tag{4}$$

where t' and t'' are arbitrarily chosen.

As a result, one sees the following lemmas:

Lemma A. $0 \leqslant \mu[S; T, \tau_0] \leqslant 1$ for $\forall S \subset \{S_i\}$.

Lemma B. $\mu[S; T, \tau_0] + \mu[S'; T, \tau_0] \leqslant \mu[S \cup S'; T, \tau_0]$ for $\forall S, \forall S' \subset \{S_i\}$ satisfying $S \cap S' = \varnothing$.

Lemma C. $P_c(S_j|S_i; T, \tau_0) = 0$ for $\forall\{\mu[S_j; T, \tau_0] = 0\}$.

Lemma D. $P_c(S_j|S_i; T, \tau_0) \neq 0$ for $\exists\{\mu[S_j; T, \tau_0] = 1\}$.

Lemma A is straightforward from the original definition (2). Since the representative point cannot be found simultaneously in both the regions S and S' without any overlapping between the two when $[T]$ is counted in a consecutive way, Lemma B also follows. For any event satisfying the condition $\mu[S_j; T, \tau_0] = 0$ it never occurs that the representative point remains at coordinate j any longer than τ_0 without undergoing transition. This proves Lemma C. On the other hand, for a certain event satisfying $\mu[S_j; T, \tau_0] = 1$ the possibility is not excluded that the representative point stays at j during the time T supposing that t' is appropriately chosen so as to fulfill $\varphi(t') = 1$. Thus, Lemma D results.

With the aids of Lemmas A–D one can prove the following theorem:

Theorem. Suppose $\mu[S_x; T, \tau_0] = 1$, where x is an arbitrary coordinate among $\{i\}$. Then

$$\mu[S_i; T, \tau_0] = 1 \rightleftarrows P_c(\forall S_{j \neq i} | S_i; T, \tau_0) = 0 \tag{5}$$

Proof. If $\mu[S_i; T, \tau_0] = 1$, then $\mu[\forall S_{j \neq i}; T, \tau_0] = 0$ because of Lemmas A and B. Thus, one sees $P_c(\forall S_{j \neq i} | S_i; T, \tau_0) = 0$ from Lemma C. Next, suppose $P_c(\forall S_{j \neq i} | S_i; T, \tau_0) = 0$ under the constraint $\mu[S_x; T, \tau_0] = 1$. On the other hand, $P_c(S_x | S_i; T, \tau_0) \neq 0$ follows for a certain event satisfying $\mu[S_x; T, \tau_0] = 1$ because of Lemma D. If $x \neq i$, contradiction would occur. Thus, $\mu[S_i; T, \tau_0] = 1$ results. Q.E.D.

3. K-ENTROPY

3.1. Macrostructure at Its Asymptote

Suppose that an asymptotic macrostructure S_i is realized in the form

$$\mu[S_i; T \to \infty, \tau_0] = 1 \tag{6}$$

Any asymptote is so defined that it has a certain invariance with respect to temporal translation. Hence the probability measure $P_i(t_2, t_1)$ that the representative point found at i at $t = t_1$ remains until $t = t_2$ $(> t_1)$ without undergoing any transition satisfies

$$P_i(t_2 + T_i, t_1 + T_i) = P_i(t_2, t_1) \tag{7}$$

where T_i is a characteristic time representing the translational invariance. Here it is understood that T_i is a possible least time among those that satisfy (7) for any set of t_1 and t_2 $(> t_1)$. Because of the property (7), the relation

$$P_i(nT_i + t_1, (n-1)T_i + t_1) = P_i(T_i + t_1, t_1) \tag{8}$$

follows, where n is a positive integer.

If one observes the event $\mu[S_i; nT_i, \tau_0] = 1$, then n-fold same events will successively appear at each time interval of T_i. Even probabilistic events are no exception. Hence the probability that the representative point stays at i all through the time nT_i is just equal to the n-fold products of $P_i(T_i + t_1, t_1)$ as follows:

$$P_i(nT_i + t_1, t_1) = [P_i(T_i + t_1, t_1)]^n \tag{9}$$

One must realize, however, that expression (9) comes solely from the assumed translational invariance (7) and not from an additional Markovian-like assumption, although the assertion for the presence of translational invariance, that is, asymptotic macrostructure, might either include or belong to the Markovian postulate. In fact, it is observed that any Markov process reaching its asymptote will satisfy the condition of the form (9) if a sufficiently large T_i is chosen.

The probability function $P_i(T_i + t_1, t_1)$ is independent of t_1 as

$$P_i(T_i + t_1, t_1) \equiv P_i(T_i) \tag{10}$$

since the macrostructure S_i is subject to the translational invariance with period T_i. Otherwise, an asymptotic structure with such an invariant property could not be realized. Hence one finds

$$P_i(t_2, t_1) = \{P_i^*(1)\}^{(t_2 - t_1)} \tag{11}$$

for $t_2 - t_i = nT_i$ $(n = 1, 2,...)$ with

$$P_i^*(1) \equiv \{P_i(T_i)\}^{1/T_i} \tag{12}$$

Expression (12) leads to

$$\{P_i(mT_i)\}^{1/mT_i} = \{P_i(T_i)\}^{1/T_i} = P_i^*(1) \tag{13}$$

with the aid of (10), where m is a positive integer. This property will be used later. One more characteristic of the probability function $P_i(t_2, t_1)$ is

$$\{P_i^*(1)\}^{[t_2 - t_1]_i + T_i} < P_i(t_2, t_1) \leqslant \{P_i^*(1)\}^{[t_2 - t_1]_i} \tag{14}$$

with

$$[t_2 - t_1]_i \equiv \{nT_i; nT_i \leqslant t_2 - t_1 < (n + 1)T_i \text{ for an integer } n\}$$

which is a generalization of expression (11).

Now, consider a macroscopic constraint which makes it possible to realize an asymptotic structure comprising macroprocesses with N degrees of freedom. The constraint will be supposed to be so macroscopic that any macrostructure S_i', i.e.,

$$\mu[S_i'; T \to \infty, \tau_0] = 1 \tag{15}$$

belonging to a local subset $\{S_i'\}$ of the partitioned Γ space $\{S_i\}$ as

$$\{S_i'\} \subset \{S_i\} \tag{16}$$

may not contradict the constraint. Here the auxiliary set $\{S_i''\}$ leading to

$$\{S_i'\} \cup \{S_i''\} = \{S_i\} \tag{17}$$

with

$$\{S_i'\} \cap \{S_i''\} = \varnothing$$

is supposed to yield

$$\mu[S_j''; T \to \infty, \tau_0] < 1 \qquad \text{for} \quad \forall j'' \in \{i''\} \tag{18}$$

under the same constraint. Thus, one may raise the following question: Which elements among the local set $\{S_i'\}$ would be realizable as a physically stable structure in the sense of (15)? We shall try to answer to this question in the latter half of this section.

On substituting (14) into the right-hand side of (4), one obtains the inequalities

$$\frac{\varphi(t')\{P_{j'}^*(1)\}^{T+T_{j'}}}{\{P_{i'}^*(1)\}^T} < P_c(S_{j'}|S_{i'}; T, \tau_0) < \frac{\{P_{j'}^*(1)\}^T}{\{P_{i'}^*(1)\}^{T+T_{i'}}} \tag{19}$$

Furthermore, we introduce a subset $\{\alpha\}$ belonging to $\{i'\}$ as follows:

$$\{\alpha\} = \{\{i'\}; \max[P_{i'}^*(1)]\} \tag{20}$$

in which the number of coordinates included in $\{\alpha\}$ maximizing $P_{i'}^*(1)$ is not necessarily unity. For the coordinate $i' = \beta$ with $\beta \in \{\alpha\}$, the second inequality of (19) leads to

$$\lim_{T \to \infty} P_c(S_{j'}|S_\beta; T, \tau_0) = 0 \qquad \text{for} \quad \forall j' \notin \{\alpha\} \tag{21}$$

since

$$P_{j'}^*(1) < P_\beta{}^*(1) \tag{22}$$

and

$$\lim_{T \to \infty} \{P_{j'}^*(1)/P_\beta{}^*(1)\}^T = 0 \tag{23}$$

If one defines the local subset $S_{\{\alpha\}}$ as

$$S_{\{\alpha\}} \equiv \bigcup_{\beta \in \{\alpha\}} S_\beta \tag{24}$$

the relation

$$\lim_{T \to \infty} P_c(\forall S_{j' \in \{i'\} \setminus \{\alpha\}}|S_{\{\alpha\}}; T, \tau_0) = 0 \tag{25}$$

follows, where $\{i'\}\setminus\{\alpha\}$ denotes a difference set ($\{i'\} - \{\alpha\}$). Hence, if the condition

$$\mu[S_{\{x\}}; T \to \infty, \tau_0] = 1 \qquad \text{for} \quad \exists\{x\} \in \{i'\} \tag{26}$$

is the case, the theorem of Section 2 yields

$$\mu[S_{\{\alpha\}}; T \to \infty, \tau_0] = 1 \tag{27}$$

The condition (26) simply states that there exists a certain macrostructure which could be realized asymptotically. We have already supposed that such an asymptotic structure really exists as expressed in (15). The macrostructure $S_{\{\alpha\}}$ is thus shown to be the only candidate that is stable and asymptotic.

Next, it must be exhibited that the structure $S_{\{\alpha\}}$ is truly realizable. For this purpose, the size of the coarse-grained time unit τ_0 will be examined. On introducing the coordinate γ satisfying

$$\gamma \equiv \{j'; \min_{j' \in \{i'\} \setminus \{\alpha\}} [P^*_{\{\alpha\}}(1) - P^*_{j'}(1)]\} \tag{28}$$

with $P^*_{\{\alpha\}}(1) \equiv P_\beta{}^*(1)$ for $\forall \beta \in \{\alpha\}$, one can make the quantity

$$P_c(S_{j'}|S_{\{\alpha\}}; \tau_0, \tau_0) \equiv \varphi_{j'}(t')P_{j'}(\tau_0 + t', t')/P_{\{\alpha\}}(\tau_0 + t'', t'') \tag{29}$$

with $j' \in \{i'\}\setminus\{\alpha\}$ arbitrarily small for a sufficiently large τ_0 in the range

$$\tau_0 \gg \tau_\gamma \equiv \{\ln[P^*_{\{\alpha\}}(1)/P_\gamma{}^*(1)]\}^{-1} \tag{30}$$

and

$$\tau_0 \gg T_{j'} \qquad \text{for} \quad \forall j' \in \{i'\} \tag{31}$$

On the other hand, the ansatz (26) will lead to

$$\mu[S_{\{\alpha\}}; \tau_0, \tau_0] = 1 \tag{32}$$

for almost every interval τ_0 in an infinite time $T \to \infty$. Since the right-hand side of (29) becomes vanishingly small for

$$\tau_0 \gg \max(\tau_\gamma, T_{j'} \text{ for } \forall j' \in \{i'\}) \tag{33}$$

the choice of τ_0 shown in (33) finally yields

$$\mu[S_{\{\alpha\}}; \tau_0, \tau_0] = 1 \tag{34}$$

for almost every interval τ_0 with the aid of the theorem in Section 2.

3.2. K-Entropy

The entropy density of partition introduced by Kolmogorov has the form

$$H \equiv -\log_2 P \tag{35}$$

for an event which occurs with probability P, where we refer only to the entropy of a particular event and not to that of the set of every event covering the whole probability space. This is just a simple generalization of Shannon's

entropy density of information.[19] Also, the conditional entropy density of partition is definable as

$$H(A|B) \equiv -\log_2[P(A|B)/P(B)] \tag{36}$$

where $P(B)$ is the probability of the event B, and $P(A|B)$ is that of the event B which is followed by the event A. Hence one can introduce the conditional entropy density of partition $H_i(t_2, t_1)$ that the representative point found at coordinate i at time $t = t_1$ in the coarse-grained Γ space keeps its coordinate until $t = t_2$ $(> t_1)$ without undergoing any transition as follows:

$$H_i(t_2, t_1) \equiv -\log_2 P_i(t_2, t_1) \tag{37}$$

The conditional entropy density (37) is found to yield an invariant quantity

$$h_i \equiv \sup\left\{\lim_{(t_2-t_1)\to\infty} [H_i(t_2, t_1)/(t_2 - t_1)]\right\} \tag{38}$$

$$= -\log_2 P_i^*(1) \tag{39}$$

with the aid of (14).

The quantity h_i is a measure of the increase in the conditional entropy density of partition per unit time. A similar quantity has already been investigated in the measure-preserving transformation with a real coefficient group, that is, flow.[23] The K-entropy of Krylov, Kolmogorov, and Sinai is just the quantity by which the increase in rate of entropy of partition could be measured for a given dynamical system.[17,18] The original K-entropy, however, refers to the whole phase space. On the other hand, the invariant quantity h_i of (38) is characteristic of only a particular element of Γ space. Consequently, the quantity h_i should be understood as the K-entropy of the element S_i or, equivalently, as a K-entropy density. In general, the K-entropy density h_i would not remain constant all over the Γ space even if the related dynamics is ergodic in the whole phase space. We shall call the K-entropy density simply the K-entropy when there is no chance of confusion.

Following the result presented in (20), one can establish the relation

$$h_{j'} \geqslant h_{\{\alpha\}} \quad \text{for} \quad j' \in \{i'\} \quad \text{with} \quad h_{\{\alpha\}} \equiv -\log_2 P^*_{\{\alpha\}}(1) \tag{40}$$

This predicts that the stable macrostructure which could asymptotically be realized among the various candidates, which satisfy

$$\mu[S_{j'}; T \to \infty, \tau_0] = 1 \quad \text{for} \quad \forall j' \in \{i'\} \ (\subset \{i\}) \tag{41}$$

necessarily minimizes the associated K-entropy.

The K-entropy $h_{i'}$ is an invariant quantity of the macrostructure $\mu[S_{i'}; T \to \infty, \tau_0] = 1$. This implies that the K-entropy could be regarded as a state function whose minimum fixes the structure to be realized asymptotically. The only ansatz for this statement is the presence of the asymptote.

There is an argument that a nonequilibrium potential could be defined at each finite time interval.[8,24] Examples have been found in linear[25] and nonlinear cases.[11] Given a time interval τ_{pot} for defining the potential, the distribution of fluctuations whose characteristic times are less than τ_{pot} would completely be fixed by the potential, as in fluctuations around equilibrium. If macrofluctuations whose characteristic times are greater than τ_{pot} are concerned, one must assume an additional macrokinetics to deal with such macrofluctuations. However, there is no unique way of determining the form of macrokinetics,[6,24] although hydrodynamics has frequently been tried.[3,11],[7] As an attempt to avoid the indefiniteness associated with the macrokinetics, one could derive a state function which may include any macrofluctuations. In fact, the K-entropy is just such a quantity.

The minimum K-entropy, which is identical to the least irreversible decay rate (cf. Section 4 and Ref. 16), has been argued to apply to any stable structure realized in the asymptotic time limit. This in turn provides a selection rule that the structure with less irreversible decay rate is more stable in the course of time evolution, of course, on a coarse-grained time scale.

3.3. Fluctuations Around the Minimum K-Entropy

A macrostructure $S_{i'}$ has its own K-entropy $h_{i'}$ as a state function. Although the asymptotic structure is fixed by the statement $min\{h_{i'}\}$, this

[7] It is always possible to fix the probability distribution function of fluctuations whose characteristic times are less than, say, τ_{pot}, at each interval of τ_{pot}. One might sometimes be allowed to relate the distribution function to a sort of potential. However, the question would still remain as to how the potential, if one wishes to call it so, which is necessarily local in its time scale, may have its correlation with a macrokinetics or macrofluctuations whose characteristic time is greater than τ_{pot}. In principle, the potential which is characteristic of only the less macroscopic fluctuations cannot tell anything about macrofluctuations or, equivalently, slow process. If one holds the view that the potential should be related to a probability distribution function of relevant fluctuations, macrofluctuations whose characteristic times are greater than the time unit of fixing the distribution function will always be missed in the discussion since the time unit remains finite in any case. In order to overcome the present incompleteness, Glansdorff and Prigogine[25] stress certain a priori nonlinear kinetic laws on a heuristic basis. Also, van Kampen[6] simply assumes a phenomenological law, pointing out the indefiniteness of nonequilibrium potentials. Hence it would be rather misleading to assert the presence of a potential as a probability distribution function available to any macrofluctuation. All one can do with such macrofluctuations is to follow them as time goes on. The K-entropy is a state function fixed to a macrofluctuation which appears as a once and for all event along the irreversible time axis.

Even if one succeeds in obtaining a probabilistic kinetic equation capable of any slow process, it would be inappropriate to say that a potential as a probability distribution function of any fluctuation exists unless the contribution from macrofluctuations certainly vanishes. For example, the time-dependent Ginzburg–Landau equation

never excludes fluctuations with finite duration around the point of minimum K-entropy. This is because even if

$$\mu[S_{\{\alpha\}}; T, \tau_0] = (T - T')/T \qquad \text{with} \quad T' > 0 \tag{42}$$

the asymptotic condition

$$\lim_{T \to \infty} \mu[S_{\{\alpha\}}; T, \tau_0] = 1 \tag{43}$$

will hold so long as T' is finite.

We shall consider fluctuations in the K-entropy, though these might not always be observable unless a certain effort is exerted in this direction.

The probability function $P_{i'}(t + t_1, t_1)$ leads to

$$2^{-h_{i'}([t]_{i'} + T_{i'})} < P_{i'}(t + t_1, t_1) \leq 2^{-h_{i'}[t]_{i'}} \tag{44}$$

because of (14) and (39). Expression (44) gives the upper and the lower bounds of the probability that the state with the K-entropy $h_{i'}$ continues to exist during the time t without undergoing any transition. Hence the average of the deviation $(h_{i'} - h_{\{\alpha\}})$, which continues over the interval of t, over various possible states has the form

$$\langle (h_{i'} - h_{\{\alpha\}}) \rangle_t \equiv \frac{\sum_{i'} (h_{i'} - h_{\{\alpha\}}) P_{i'}(t + t', t')}{\sum_{i'} P_{i'}(t + t'', t'')} \tag{45}$$

This expression is a kind of ensemble average only over the events that remain unchanged during the time t. If one chooses as such a time interval

$$t \gg \max(\tau_{\gamma}, T_{j'} \text{ for } \forall j' \in \{i'\}) \tag{46}$$

expression (45) gives

$$\langle \delta h \rangle_t \equiv \langle (h_{i'} - h_{\{\alpha\}}) \rangle_t = \frac{\int_0^{\infty} (\delta h) 2^{-\delta h t} \, d(\delta h)}{\int_0^{\infty} 2^{-\delta h t} \, d(\delta h)} \tag{47}$$

where the degeneracy of the macrostates having the same K-entropy is supposed not to affect the observation. The final result yields

$$\langle \delta h \rangle_t = 1/(t \ln 2) \tag{48}$$

where the constraint of time in (46) must be noticed.

A principal feature of the present result (48) is that fluctuations whose correlation times are not less than some finite value would never be excluded, exhibiting distinct contrast with fluctuations subject to the central limit theorem. This suggests that the asymptotic structure of the macroprocess is realized as an accumulation of macrofluctuations in the low-frequency limit.

---

assumes a free energy functional.[26] It is, however, by no means evident why one could neglect macrofluctuations in deriving the TDGL. Macrofluctuations gain their crucial importance near critical points. See Section 8.

The only condition for the appearance of lasting fluctuations is that

$$\{\alpha\} \subset \{i'\} \qquad \text{and} \qquad \{\alpha\} \neq \{i'\} \tag{49}$$

in expression (20), otherwise any relative difference $h_{i'} - h_{\{\alpha\}}$ would vanish. In fact, the requirement (49) will be seen not to contradict broken microscopic reversibility as discussed in the next section.

One can also imagine fluctuations in the characteristic time of the translational invariance. The characteristic time $T_{\{\alpha\}}$ which is supposed to exist might be equal to any $T_{i'}$ for $i' \in \{\alpha\}$ or a common multiple among some of them. If macrofluctuations of the characteristic time $T_{\{\alpha\}}$ occur, one could express the fluctuation phenomena as

$$T_{\{\alpha\}} = T^*_{\{\alpha\}} + \delta T_{\{\alpha\}} \tag{50}$$

where $T^*_{\{\alpha\}}$ is the characteristic time representing an asymptotic macrostructure with the minimum K-entropy, and $\delta T_{\{\alpha\}}$ is a slowly varying part of the instantaneous characteristic time $T_{\{\alpha\}}$ satisfying

$$|\delta T_{\{\alpha\}}| \ll T^*_{\{\alpha\}} \tag{51}$$

The ratio of the probability that the deviation $\delta T_{\{\alpha\}}$ continues over the time t ($\gg T^*_{\{\alpha\}}$) to that of fluctuation-free event is found to be

$$P_c(\delta T_{\{\alpha\}}, t) = \{P^*_{\{\alpha\}}(1)\}^{t/T_{\{\alpha\}}} / \{P^*_{\{\alpha\}}(1)\}^{t/T^*_{\{\alpha\}}} \qquad \text{for} \quad t \gg T^*_{\{\alpha\}} \tag{52}$$

with the aid of (12) and (20), where the strong inequality

$$|(\partial/\partial T_{\{\alpha\}}) P^*_{\{\alpha\}}(1)\ \delta T_{\{\alpha\}}| \ll P^*_{\{\alpha\}}(1) \tag{53}$$

is noticed because of (51). Since the fluctuation-free state is an asymptotic one, we have the inequality

$$P_c(\delta T_{\{\alpha\}}, t) \leqslant 1 \tag{54}$$

which, in turn, gives the constraint

$$\delta T_{\{\alpha\}} \geqslant 0 \tag{55}$$

Hence the stochastic average of the square of the fluctuation $\delta T_{\{\alpha\}}$ which continues over the time t yields

$$\langle \delta T^2_{\{\alpha\}} \rangle_t \equiv \frac{\int_0^\infty (\delta T_{\{\alpha\}})^2 P_c(\delta T_{\{\alpha\}}, t)\, d(\delta T_{\{\alpha\}})}{\int_0^\infty P_c(\delta T_{\{\alpha\}}, t)\, d(\delta T_{\{\alpha\}})}$$

$$= \frac{2(T^*_{\{\alpha\}})^4}{[\ln P^*_{\{\alpha\}}(1)]^2} \frac{1}{t^2} \qquad \text{for} \quad t \gg T^*_{\{\alpha\}}, \tau_\gamma \tag{56}$$

One can write expression (56) in terms of relative fluctuations as

$$\langle(\delta T_{\{\alpha\}}/T^*_{\{\alpha\}})^2\rangle_t = [2(T^*_{\{\alpha\}})^2/(h_{\{\alpha\}} \ln 2)^2](1/t^2) \tag{57}$$

where we used the expression in (40).

4. PAULI MASTER EQUATION

Given a partition of Γ space as

$$\{S_i\} = \Gamma, \qquad i = 1, 2,... \tag{58}$$

let us suppose that the representative point of the finite system with N degrees of freedom is well described in terms of only the set of discrete coordinates $\{i\}$ and time t within the coarse graining presently employed. We shall impose on a prospective macrostructure appearing as its asymptote a stronger condition than that of a simple translational invariance. A Markov process will be supposed which is subject to the Pauli master equation

$$(\partial/\partial t)P(i, t) = \sum_{j \neq i} W(i \leftarrow j)P(j, t) - \sum_{j \neq i} W(j \leftarrow i)P(i, t) \tag{59}$$

concerning only the discrete coordinates $\{i\}$ and time, where $P(i, t)$ is the probability density of finding the representative point at coordinate i and $W(i \leftarrow j)$ is the probability of transition from j to i per unit time. The symmetry property giving the microscopic reversibility

$$W(i \leftarrow j) = W(j \leftarrow i) \tag{60}$$

will not be assumed.

The probability function $P_i(t_2, t_1)$ introduced in Section 2 leads to

$$P_i(t_2, t_1) = \exp\left[-\int_{t_1}^{t_2} W(i)\, dt\right] \tag{61}$$

with

$$W(i) \equiv \sum_{j \neq i} W(j \leftarrow i) \tag{62}$$

since the representative point found at coordinate i follows the kinetic equation

$$(\partial/\partial t)P(i, t) = -\sum_{j \neq i} W(j \leftarrow i)P(i, t) \tag{63}$$

with

$$P(j, t) = 0 \qquad \text{for} \quad \forall j \neq i \tag{64}$$

The quantity $P_i^*(1)$ of (12), which plays a role of figure of merit for determining a macrostructure as asymptotically realizable, is expressible as

$$P_i^*(1) = \lim_{T\to\infty} \{P_i(T)\}^{1/T} \tag{65}$$

because of the property (13). As a result, we obtain

$$P_i^*(1) = \exp[-\langle W(i)\rangle] \tag{66}$$

with

$$\langle W(i)\rangle \equiv \lim_{T\to\infty} (1/T) \int_0^T W(i)\, dt \tag{67}$$

If the transition probabilities $\{W(i \leftarrow j)\}$ are independent of time, as will often be the case, expression (66) reduces to

$$P_i^*(1) = \exp[-W(i)] \tag{68}$$

The result (20) predicts that if there exists a macrostructure

$$\mu[S_{\{\alpha\}}; T\to\infty, \tau_0] = 1 \qquad \text{for} \quad \{\alpha\} \subset \{i\} \tag{69}$$

the subset $\{\alpha\}$ must be

$$\{\alpha\} = \{\{i\}; \max[P_i^*(1)]\} \tag{70a}$$

That is to say

$$\{\alpha\} = \{\{i\}; \min[\langle W(i)\rangle]\} \tag{70b}$$

Next, we shall investigate a sufficient condition under which the macrostructure (69) subject to (70) is truly realizable. For this purpose, we introduce the quantity

$$\langle \tilde{W}(i)\rangle \equiv \lim_{T\to\infty} (1/T) \int_0^T \left\{ \sum_{j\neq i} W(i \leftarrow j) \right\} dt \tag{71}$$

which measures the probability of transition from anywhere except i to i per unit time. If the inequality

$$\langle W(i)\rangle < \langle \tilde{W}(i)\rangle \tag{72}$$

is satisfied, the transition from i to anywhere else would occur less frequently than the reverse transition from anywhere else to i. Any transition process must be balanced by its counterpart in the asymptotic limit. This, however, by no means implies that each transition must be in a strict detailed balance.[5]

Now, suppose two characteristic times

$$\tau(i \to i) \qquad \text{and} \qquad \tau(i \leftrightarrow i) \tag{73}$$

Here the quantity $\tau(i \to i)$ is the mean recurrence time from the moment the representative point transits from anywhere else to i until the moment the point returns to i after having undergone transition from i to anywhere else, and $\tau(i \leftrightarrow i)$ is the average holding time over which the representative point stays

at i within one period $\tau(i \to i)$. With the aid of the characteristic time $\tau(i \leftrightarrow i)$, the asymptotic balance between the transition from i to anywhere else and its reverse is expressible as

$$\langle W(i) \rangle = \langle \tilde{W}(i) \rangle \exp[-\langle W(i) \rangle \tau(i \leftrightarrow i)] \tag{74}$$

or, equivalently,

$$\tau(i \leftrightarrow i) = (1/\langle W(i) \rangle) \ln(\langle \tilde{W}(i) \rangle / \langle W(i) \rangle) \tag{75}$$

Needless to say, the characteristic time $\tau(i \leftrightarrow i)$ vanishes for any i if the symmetry relation (60) resulting in

$$\langle W(i) \rangle = \langle \tilde{W}(i) \rangle \qquad \text{for} \quad \forall i \in \{i\} \tag{76}$$

is the case. One also observes $\tau(i \leftrightarrow i) = 0$ in the case that

$$\langle W(i) \rangle > \langle \tilde{W}(i) \rangle \tag{77}$$

In order to estimate the magnitude of $\tau(i \to i)$ in (73), let us consider the probability

$$P^{(\nu)}_{\text{cont}}(\forall j \neq i; t_2, t_1) \tag{78}$$

for an event ν $(=1, 2,...)$ that the representative point is never found at coordinate i from the moment t_1 until t_2 $(> t_1)$. Hence it follows that

$$\frac{P^{(\nu)}_{\text{cont}}(\forall j \neq i; t + t', t')}{P_i(t + t', t')} < C \frac{[\max_{j \neq i}\{P_j^*(1)\}]^t}{\{P_i^*(1)\}^t} \tag{79}$$

where C is a positive constant independent of time t as seen from (14). As a particular case of the inequality (79), one obtains

$$\frac{P^{(\nu)}_{\text{cont}}(\forall j \notin \{\alpha\}; t + t', t')}{P_{\{\alpha\}}(t + t', t')} < C \frac{\{P_\gamma^*(1)\}^t}{\{P^*_{\{\alpha\}}(1)\}^t} \tag{80}$$

in which the quantity $P_\gamma^*(1)$, which is less than $P^*_{\{\alpha\}}(1)$, has already been introduced in (28).

If the time t satisfies the inequality

$$t \gg \{\ln[P^*_{\{\alpha\}}(1)/P_\gamma^*(1)]\}^{-1} \ (\equiv \tau_\gamma) \tag{81}$$

the right-hand side of (80) almost vanishes. The probability that the representative point can never be found on any of the coordinates belonging to $\{\alpha\}$ of (70) throughout the time interval much greater than τ_γ would thus become vanishingly small compared with the probability that the representative point can be found on any one of $\{\alpha\}$ for even a moment during the interval. Hence if one chooses the coarse-grained time unit τ_0 in the range

$$\tau_0 \gg \max(\tau_\gamma, T_j \text{ for } \forall j \in \{i\}) \tag{82}$$

where T_j is the fundamental period of the time-dependent transition probability $W(j)$, the mean recurrence time $\tau(\{\alpha\} \to \{\alpha\})$ will satisfy $\tau(\{\alpha\} \to \{\alpha\}) \lesssim \tau_0$. A similar discussion has already been developed with regard to (28)–(34). This will prove the statement

$$\mu[S_{\{\alpha\}}; T \to \infty, \tau_0] = 1 \tag{83}$$

if the inequality $\langle \tilde{W}(\beta) \rangle - \langle W(\beta) \rangle > 0$ holds. In fact, once the representative point is found at any one of $\{\alpha\}$, the point is expected to remain at $\{\alpha\}$ during the time

$$\tau(\{\alpha\} \leftrightarrow \{\alpha\}) \sim \frac{1}{\langle W(\beta) \rangle} \ln \frac{\langle \tilde{W}(\beta) \rangle}{\langle W(\beta) \rangle} \; (>0) \qquad \text{for} \quad \beta \in \{\alpha\} \tag{84}$$

on average, as seen from (75) (cf. Ref. 16).

As a result, one finds the asymptotic solution $P_{\text{asymp}}(i, t)$ to the Pauli master equation (59) of the form

$$P_{\text{asymp}}(x, t) = \begin{cases} 1 & \text{for } x = \{\alpha\} \\ 0 & \text{for } x \in \{i\} \backslash \{\alpha\} \end{cases} \tag{85}$$

only in the sense of

$$\mu[S_x; T \to \infty, \tau_0] = \begin{cases} 1 & \text{for } x = \{\alpha\} \\ 0 & \text{for } x \in \{i\} \backslash \{\alpha\} \end{cases} \tag{86}$$

where the following two constraints must be satisfied:

$$\langle \tilde{W}(\beta) \rangle - \langle W(\beta) \rangle > 0 \qquad \text{for} \quad \beta \in \{\alpha\} \tag{87}$$

$$\tau_0 \gg \max \left\{ \left[\min_{j \in \{i\} \backslash \{\alpha\}} (\langle W(j) \rangle - \langle W(\beta) \rangle) \right]^{-1}, T_j \text{ for } \forall j \in \{i\} \right\} \tag{88}$$

If the inequality (87) is violated, a trial by which the solution of the form (85) could be retained is to employ another way of partitioning Γ space as

$$\{S_i^*\} = \Gamma, \qquad i^* = 1, 2, \ldots \tag{89}$$

instead of (58) so that the inequality $\langle \tilde{W}(\beta^*) \rangle - \langle W(\beta^*) \rangle > 0$ could survive in the revised coordinate system $\{i^*\}$.

If

$$\min_{j \in \{i\} \backslash \{\alpha\}} \{\langle W(j) \rangle - \langle W(\beta) \rangle\} \to 0 \tag{90}$$

in (88) due to a change in external constraints which fix the set of transition probabilities $\{W(i \leftarrow j)\}$, the structure S_j satisfying (90) could also participate in the stable asymptotic structure. In fact, the critical-point behavior giving rise to a rearrangement of asymptotic macrostructure would be common to phase changes on a macroscale.

The discussions on the K-entropy and its fluctuations around the minimum developed in the previous section are also applicable to the present Markovian system. In fact, if microscopic reversibility (60) does not hold, one can obtain

$$\{\alpha\} \subset \{i\} \qquad \text{and} \qquad \{\alpha\} \neq \{i\} \tag{91}$$

following (70). This is the very condition under which fluctuations of macrostructure not subject to the central limit theorem could be observed, as seen from (45)–(48).

5. APPLICATION I: ONE DEGREE OF FREEDOM

We shall apply the general method developed in the scheme of the Pauli master equation to a specific example found in the Fokker–Planck equation with one degree of freedom. Its relevance to experiments will also be argued.

5.1. General Discussion

Consider the Fokker–Planck equation with one degree of freedom

$$\frac{\partial}{\partial T}P(A, t) = \frac{\partial}{\partial A}[g(A)P(A, t)] + \frac{\partial^2}{\partial A^2}[h(A)P(A, t)] \tag{92}$$

where $P(A, t)$ is the probability density of finding a macroscopic variable A, which is real, at time t, and $g(A)$ and $h(A)$ are functions of the variable A. The minimum time unit

$$\tau_{\mathrm{FP}} \tag{93}$$

for observing the Fokker–Planck behavior never vanishes.

On introducing a coarse graining denoted as ΔA as a unit for measuring the variable A, expression (92) within $O(\Delta A)$ yields

$$\begin{aligned}\frac{\partial}{\partial t}P(A, t) = {} & \frac{1}{\Delta A}[g(A + \Delta A)P(A + \Delta A, t) - g(A)P(A, t)] \\ & + \frac{1}{\Delta A^2}[h(A + \Delta A)P(A + \Delta A, t) - 2h(A)P(A, t) \\ & + h(A - \Delta A)P(A - \Delta A, t)] \\ = {} & W(A \leftarrow A + \Delta A)P(A + \Delta A, t) \\ & + W(A \leftarrow A - \Delta A)P(A - \Delta A, t) \\ & - W(A, \Delta A)P(A, t)\end{aligned} \tag{94}$$

with

$$W(A \leftarrow A + \Delta A) \equiv \frac{g(A + \Delta A)}{\Delta A} + \frac{h(A + \Delta A)}{\Delta A^2} \tag{95}$$

$$W(A \leftarrow A - \Delta A) \equiv \frac{h(A - \Delta A)}{\Delta A^2} \tag{96}$$

$$W(A, \Delta A) \equiv \frac{g(A)}{\Delta A} + \frac{2h(A)}{\Delta A^2} \tag{97}$$

We shall further suppose the inequalities

$$W(A \leftarrow A + \Delta A) > 0, \qquad W(A \leftarrow A - \Delta A) > 0,$$
$$W(A, \Delta A) > 0 \tag{98}$$

which will turn out to be quite plausible in practical cases.

The Fokker–Planck equation (94) is a specific example of the Pauli master equation (59). If one applies the results (85)–(88) to the present case, the asymptotic solution to Eq. (94) leads to

$$P_{\text{asymp}}(A, t) = \begin{cases} 1 & \text{for } |A - A_m| \leqslant \frac{1}{2}\Delta A \\ 0 & \text{for } |A - A_m| > \frac{1}{2}\Delta A \end{cases} \tag{99}$$

in the sense of

$$\mu[A_m \pm \tfrac{1}{2}\Delta A; T \to \infty, \tau_0] = \begin{cases} 1 & \text{for } |A - A_m| \leqslant \frac{1}{2}\Delta A \\ 0 & \text{for } |A - A_m| > \frac{1}{2}\Delta A \end{cases} \tag{100}$$

with

$$W(A_m + \Delta A, \Delta A) > 0, \qquad W(A_m - \Delta A, \Delta A) > 0 \tag{101}$$

where the following two constraints must be taken into account:

$$\tilde{W}(A_m, \Delta A) - W(A_m, \Delta A) > 0 \tag{102}$$

with

$$\tilde{W}(A, \Delta A) \equiv W(A \leftarrow A + \Delta A) + W(A \leftarrow A - \Delta A) \tag{103}$$

and

$$\tau_0 \gg [W(A_m + \Delta A, \Delta A) - W(A_m, \Delta A)]^{-1} \tag{104}$$

The argument $A_m \pm \frac{1}{2}\Delta A$ of the measure $\mu[A_m \pm \frac{1}{2}\Delta A, T \to \infty, \tau_0]$ implies that the macroscopic variable A is found in the range

$$A_m - \tfrac{1}{2}\Delta A < A \leqslant A_m + \tfrac{1}{2}\Delta A \tag{105}$$

If both the functions $g(A)$ and $h(A)$ are differentiable, the conditions (101) and (104) will be written as

$$\left.\frac{\partial W(A, \Delta A)}{\partial A}\right|_{A=A_m} = 0, \qquad \left.\frac{\partial^2 W(A, \Delta A)}{\partial A^2}\right|_{A=A_m} > 0 \tag{106}$$

and

$$\tau_0 \gg \left[\frac{1}{2}\left.\frac{\partial^2 W(A, \Delta A)}{\partial A^2}\right|_{A=A_m} (\Delta A)^2\right]^{-1} \tag{107}$$

respectively.

One more consideration which is specific to the present Fokker–Planck equation concerns the size of the coarse graining ΔA. The probability measure which must be minimized in the asymptotic limit is also dependent upon the size ΔA which the observer employs. If the macrostructure

$$\mu[A_m \pm \tfrac{1}{2}\Delta A_m, T \to \infty, \tau_0] = 1 \tag{108}$$

is realized for $\Delta A = \Delta A_m$, the theorem of Section 2 yields

$$\lim_{T\to\infty} P_c(A_m + p\,\Delta A_m \pm \tfrac{1}{2}\Delta A_m | A_m \pm \tfrac{1}{2}\Delta A_m; T, \tau_0) = 0 \tag{109}$$

with $p = \pm 1, \pm 2, \ldots$. Hence the asymptote which should be observed must satisfy

$$\left.\frac{\partial W(A_m, \Delta A)}{\partial(\Delta A)}\right|_{\Delta A=\Delta A_m} = 0 \qquad \text{(if any)} \tag{110}$$

with

$$\left.\frac{\partial^2 W(A_m, \Delta A)}{\partial(\Delta A)^2}\right|_{\Delta A=\Delta A_m} > 0 \tag{111}$$

Here the size of the coarse graining ΔA_m must be sufficiently small since the Fokker–Planck equation (94) follows by neglecting terms of the order $O(\Delta A)$ or higher.

Fluctuations around $A = A_m$ are expressed in terms of the function $P_{A,\Delta A}(t)$, which measures the probability that the variable A keeps its value without any fluctuations over the time t, as

$$P_{A,\Delta A}(t) \equiv \exp[-W(A, \Delta A)t] \tag{112}$$

as seen from (61). The mean square of fluctuations in which A deviates from A_m by δA during the time t can be expressed as

$$\langle \delta A^2 \rangle_t \equiv \frac{\int_{-\infty}^{\infty} (\delta A)^2 P_c(\delta A, t)\, d(\delta A)}{\int_{-\infty}^{\infty} P_c(\delta A, t)\, d(\delta A)} \tag{113}$$

with

$$P_c(\delta A, t) \equiv \frac{P_{A_m+\delta A, \Delta A_m}(t)}{P_{A_m, \Delta A_m}(t)} \tag{114}$$

This is an ensemble average of macroscopic events each of which continues during the time t. If the probability function $W(A, \Delta A_m)$ is differentiable at $A = A_m$, the averaged fluctuation intensity $\langle \delta A^2 \rangle_t$ gives

$$\langle \delta A^2 \rangle_t = \frac{1}{[\partial^2 W(A\ \Delta A_m)/\partial A^2]|_{A=A_m}} \frac{1}{t} \tag{115}$$

On the other hand, if $W(A, \Delta A_m)$ is not differentiable at $A = A_m$, expression (113) yields

$$\langle \delta A^2 \rangle_t = \left\{ \frac{2}{[W_A(A_m{}^+, \Delta A_m)]^2} + \frac{2}{[W_A(A_m{}^-, \Delta A_m)]^2} \right\} \frac{1}{t^2} \tag{116}$$

with

$$W_A(A_m{}^{\pm}, \Delta A_m) \equiv \lim_{A \to A_m \pm 0} [\partial W(A, \Delta A_m)/\partial A] \tag{117}$$

The time t appearing in (113) must satisfy the inequality

$$t \gg \tau_{\mathrm{FP}} \tag{118}$$

along with

$$t \gg \tau_0 \tag{119}$$

since a probabilistic kinetics whose characteristic time is less than τ_{FP} of (93) has already been averaged out.

Next it will be shown that the mean square of fluctuations $\langle \delta A^2 \rangle_t$ can be related to a correlation function of $\delta A(t')$, although in a very restricted sense. One can write the fluctuation $\delta A(t')$ in the form

$$\delta A(t') = \delta A_{>s}(t') + \delta A_{<s}(t') \tag{120}$$

with

$$\lim_{T \to \infty} (1/T) \int_0^T [\delta A_{>s}(t') + \delta A_{<s}(t')]\, dt' = 0 \tag{121}$$

Here the fluctuation $\delta A_{>s}(t')$ satisfies

$$\delta A_{>s}(t'' + t') - \delta A_{>s}(t') = 0 \qquad \text{for} \quad 0 < t'' < s \tag{122}$$

if

$$\delta A_{>s}(t' + 0) \neq \delta A_{>s}(t' - 0) \tag{123}$$

and $\delta A_{<s}(t')$ represents other less macroscopic fluctuations than those involved in $\delta A_{>s}(t')$.

When a time average is involved, one could always imagine the presence of a low-pass filter because of an inevitable high-frequency cutoff associated

with any measurement. If the cutoff frequency is of the order of $1/s$, one can approximate expression (120) by

$$\delta A(t') = \delta A_{>s}(t') \tag{124}$$

only from an observational viewpoint.

The time correlation function

$$\delta A_t^2 \equiv \lim_{T\to\infty} (1/T) \int_0^T \delta A(t' + t)\, \delta A(t')\, dt' \tag{125}$$

would be replaced approximately by

$$\lim_{T\to\infty} (1/T) \int_0^T \delta A_{>s}(t' + t)\, \delta A_{>s}(t')\, dt' \tag{126}$$

if one measures the quantity δA_t^2 considering only the low-frequency components which can pass through such a low-pass filter with a cutoff frequency of $\sim 1/s$. Hence, if the strong inequality

$$t \ll s \tag{127}$$

holds, the measurable quantity δA_t^2 as a time average of $\delta A(t' + t)\, \delta A(t')$ could be interpreted as a stochastic average of δA^2 keeping its value during the time t since δA itself follows a probabilistic kinetics. This is also because any member constituting the ensemble of stochastic events would necessarily last over a much longer time than the fixed time t. Only under the constraint (127) could one expect an approximate equality

$$\langle \delta A^2 \rangle_t = \delta A_t^2 \tag{128}$$

An immediate consequence obtained from expression (115) is that if the probability function $W(A, \Delta A)$ is differentiable twice at $A = A_m$, as is often satisfied in nonlinear kinetics, the fluctuation intensity of macrofluctuations $\langle \delta A^2 \rangle_t$, which would sometimes be interpreted as the time correlation function in the sense of (128), will have an inverse time characteristic[20] except for the trivial case of $W(A, \Delta A) = \text{const}$. This would give a low-frequency singularity of logarithmic divergence in the spectrum. On the other hand, it is seen from expression (116) that if the probability function $W(A, \Delta A)$ ($\neq$const) is not differentiable twice at $A = A_m$, as will be experienced in linear kinetics, the fluctuation intensity has a t^{-2} characteristic instead of t^{-1}. One readily notes that the t^{-2} characteristic in the correlation function does not lead to any low-frequency singularity in its spectral representation. Since the intensity proportional to t^{-1} will become much greater than that proportional to t^{-2} for increasing t, it is concluded that any nonlinear contribution to a sufficiently long-time behavior cannot be neglected

no matter how small its strength might be compared with that of linear process. The further implication of this will be discussed in Section 5.5.

5.2. Brownian Motion

If one puts in the Fokker–Planck equation

$$g(A) = U \tag{129}$$

$$h(A) = D \tag{130}$$

where both U and D are positive constants, the resulting equation

$$(\partial/\partial t)P(A, t) = (\partial/\partial A)[UP(A, t)] + (\partial^2/\partial A^2)[DP(A, t)] \tag{131}$$

governs a particular Brownian motion. In the coarse-grained representation one obtains (94) with

$$W(A \leftarrow A + \Delta A) = (U/\Delta A) + (D/\Delta A^2) \tag{132}$$

$$W(A \leftarrow A - \Delta A) = (D/\Delta A^2) \tag{133}$$

$$W(A, \Delta A) = (U/\Delta A) + (2D/\Delta A^2) \tag{134}$$

An immediate consequence is

$$\tilde{W}(A, \Delta A) - W(A, \Delta A) = 0 \tag{135}$$

for any A and ΔA, where $\tilde{W}(A, \Delta A)$ is given in (103). Relation (135), as a specific form of (76), does not contradict the presence of microscopic reversibility in the coarse-grained Γ space, although this is one dimensional in the present case. If the one-dimensional space is not unlimited, the asymptotic solution will have the well-known form

$$P_{\text{asymp}}(A, t) = \text{const} \tag{136}$$

as a realization of the principle of equal weight. Fluctuations with the inverse time characteristic as shown in (45)–(48) never appear.

5.3. Linear Fokker–Planck Equation

If one replaces the function (129) by

$$g(A) = a_1 A \qquad \text{with} \quad a_1 > 0 \tag{137}$$

a linear Fokker–Planck equation

$$(\partial/\partial t)P(A, t) = (\partial/\partial A)[a_1 AP(A, t)] + (\partial^2/\partial A^2)[DP(A, t)] \tag{138}$$

in the sense of van Kampen[13] will follow. The set of transition probabilities (95)–(97) will be expressed as

$$W(A \leftarrow A + \Delta A) = \frac{a_1(A + \Delta A)}{\Delta A} + \frac{D}{\Delta A^2}$$

$$W(A \leftarrow A - \Delta A) = \frac{D}{\Delta A^2} \tag{139}$$

$$W(A, \Delta A) = \frac{a_1 A}{\Delta A} + \frac{2D}{\Delta A^2}$$

Since one readily observes

$$\tilde{W}(A, \Delta A) - W(A, \Delta A) = a_1 \tag{140}$$

the right-hand side of which is always positive, the probabilistic kinetics of the macrovariable A in the coarse-grained phase space does not obey microscopic reversibility.

One more specific feature about the Fokker–Planck equation (138) is that it is invariant under the reflection $A \rightarrow -A$. Hence we shall consider only the case

$$A \geqslant 0 \tag{141}$$

The reflection symmetry enables us to find

$$A_m = 0 \tag{142}$$

with the aid of (101). The size of coarse graining ΔA_m remains indefinite,

$$\Delta A_m: \quad \text{indefinite} \tag{143}$$

since one cannot find a finite ΔA_m which would lead to the minimization in (110). Hence the asymptotic solution to the linear Fokker–Planck equation (138) gives

$$P_{\text{asymp}}(A \pm \tfrac{1}{2}\Delta A_m, t) = \begin{cases} 1 & \text{for } |A| \leqslant \frac{1}{2}\Delta A_m \\ 0 & \text{for } |A| > \frac{1}{2}\Delta A_m \end{cases} \tag{144}$$

in the sense of

$$\mu[A \pm \tfrac{1}{2}\Delta A_m; T \rightarrow \infty, \tau_0] = \begin{cases} 1 & \text{for } |A| \leqslant \frac{1}{2}\Delta A_m \\ 0 & \text{for } |A| > \frac{1}{2}\Delta A_m \end{cases} \tag{145}$$

under the constraint

$$\tau_0 \gg 1/a_1 \tag{146}$$

The size of ΔA_m could be fixed arbitrarily depending upon what sort of

measurement may be taken by each relevant observer. The fluctuation intensity $\langle \delta A^2 \rangle_t$ in (113) leads to

$$\langle \delta A^2 \rangle_t = 4(\Delta A_m^2/a_1{}^2)(1/t^2) \tag{147}$$

following expression (116), in which the time interval t must satisfy the inequality

$$t \gg \tau_{\mathrm{FP}}, \tau_0 \tag{148}$$

being similar to (118) and (119).

We shall concentrate only on fluctuations. If one considers the probabilistic kinetics of only microscopic fluctuations averaged over the time interval τ_{FP} and neglects any macrofluctuation whose characteristic time is greater than τ_{FP}, one obtains an asymptotic solution $P_{\mathrm{asymp}}(A, t)$ to (138) having the Gaussian form

$$P_{\mathrm{asymp}}(A, t) = (a_1/2\pi D)^{1/2} \exp[-(a_1/2D)A^2] \tag{149}$$

The mean square of fluctuations δA averaged over the Gaussian probabilistic space (149) leads to

$$(\delta A^2)^G = D/a_1 \tag{150}$$

in comparison with expression (147). One must, however, distinguish between the macrofluctuations, which give the fluctuation intensity of the form (147) with the characteristic time much greater than τ_{FP}, and the microscopic ones subject to a sort of central limit theorem resulting in the Gaussian distribution (149).

The linear Fokker–Planck equation (138) is related to a linear Langevin equation of the form[27]

$$(d/dt)A(t) = -a_1 A(t) + R(t) \tag{151}$$

where $R(t)$ is a random force whose correlation time is less than τ_{FP}. Following an ensemble average, which might be equivalent to an infinite-time average, one obtains the spectral representation of (151) as

$$\langle A^2(\omega) \rangle_{\mathrm{ens}} = \langle R^2(\omega) \rangle_{\mathrm{ens}}/(\omega^2 + a_1{}^2) \tag{152}$$

with

$$\langle A^2(\omega) \rangle_{\mathrm{ens}} \equiv \int_{-\infty}^{\infty} \langle A(t + t')A(t') \rangle_{\mathrm{ens}} e^{-i\omega t}\, dt \tag{153}$$

$$\langle R^2(\omega) \rangle_{\mathrm{ens}} \equiv \int_{-\infty}^{\infty} \langle R(t + t')R(t') \rangle_{\mathrm{ens}} e^{-i\omega t}\, dt \tag{154}$$

Since the correlation time of $R(t)$ is always less than τ_{FP}, the spectral intensity (153) gives

$$\langle A^2(\omega) \rangle_{\mathrm{ens}} \sim \text{const} \qquad \text{for} \quad \omega \ll \min\{a_1, 1/\tau_{\mathrm{FP}}\} \tag{155}$$

On the other hand, the macrofluctuations with the correlation function (147) in the sense of (128) will yield a linear frequency dependence for

$$\omega \ll \min\{1/\tau_0, 1/\tau_{\mathrm{FP}}\} \tag{156}$$

because of

$$\langle \delta A^2 \rangle_t \propto 1/t^2 \qquad \text{for} \quad t \gg \tau_{\mathrm{FP}}, \tau_0 \tag{157}$$

Hence one observes that the Lorentzian-shaped spectrum (152) based upon an ensemble average of any event appearing in an infinite time interval never coincides with the spectrum based upon the correlation function (147) in the sense of (128) following the ensemble average of only macroevents each of which keeps its own structure during a finite time interval. This, however, is by no means a contradiction since there is no reason for an ensemble average of any event appearing in an infinite time interval to agree with the similar average of only macroevents continuing over a finite time. Since it would not be possible to prepare an ensemble of macroevents the duration t of each of which approaches infinity, the resulting fluctuations among the macroevents could no longer be expected in the limit $t \to \infty$. In fact, the intensity (147) of macrofluctuations, which are characteristic only of the ensemble average of macroevents continuing over the time t, would vanish for $t \to \infty$. However, the time needed for fixing each macroevent does remain finite for any experimental observation.

One should note that the linear Fokker–Planck equation is not consistent with the statement of the central limit theorem. Only after the ansatz is employed that macrofluctuations with characteristic time greater than τ_{FP} in (93) and τ_0 can be discarded will a sort of central limit theorem be reproduced. So long as microscopic reversibility does not hold in the coarse-grained Γ space, as exemplified in the case of the linear Fokker–Planck equation (138), macrofluctuations whose correlation times are not less than some finite value could never be excluded. This observation also agrees with the fact that the linear Fokker–Planck equation as a Sturm–Liouville operator has a uniform distribution of its eigenvalues.[14]

5.4. Nonlinear Fokker–Planck Equation

We shall replace the linear function in (137) by

$$g(A) = a_1 A + a_2 A^2 + a_3 A^3 \tag{158}$$

where nonlinear terms higher than A^4 will not be considered. The corresponding nonlinear Fokker–Planck equation yields

$$\frac{\partial}{\partial t} P(A, t) = \frac{\partial}{\partial A} [(a_1 A + a_2 A^2 + a_3 A^3) P(A, t)] + \frac{\partial^2}{\partial A^2} [D P(A, t)] \tag{159}$$

Although the present choice by no means includes all possible nonlinear phenomena following Fokker–Planck equations, certain experimental results could be explained within the framework of (159), as will be discussed.

Case I. $a_1 \neq 0$, $a_2 > 0$, and $a_3 = 0$.
The asymptotic value of A gives

$$A_m = -a_1/2a_2 \tag{160}$$

following (106). The size of coarse graining ΔA must satisfy

$$\Delta A < 8a_2 D/a_1^2 \tag{161}$$

because of the condition $W(A_m, \Delta A) > 0$. In fact, ΔA must always be a sufficiently small quantity. The intensity of fluctuations $\langle \delta A^2 \rangle_t$ in (113) reduces in the present case to

$$\langle \delta A^2 \rangle_t = (\Delta A/2a_2)(1/t) \tag{162}$$

Non-Gaussian fluctuations, whose correlation function in the sense of (128) has an inverse time characteristic, are seen to diverge as a_2^{-1} for $a_2 \to +0$.

Case II. $a_1 = 0$, $a_2 \neq 0$, and $a_3 > 0$.
Following the same procedures as before, one obtains

$$A_m = 0, \qquad \langle \delta A^2 \rangle_t = (\Delta A/2a_2)(1/t) \qquad \text{for} \quad a_2 > 0 \tag{163}$$

and

$$A_m = -2a_2/3a_3, \qquad \langle \delta A^2 \rangle_t = -(\Delta A/2a_2)(1/t) \qquad \text{for} \quad a_2 < 0 \tag{164}$$

with

$$\Delta A < -27a_3^2 D/2a_2^3$$

The fluctuation intensity $\langle \delta A^2 \rangle_t$ diverges as $|a_2|^{-1}$ for $a_2 \to \pm 0$.

Case III. $a_1 \neq 0$, $a_2 = 0$, and $a_3 > 0$.
This particular case follows from the nonlinear Langevin equation of van der Pol type

$$(d/dt)A(t) = -a_1 A(t) - a_3 A^3(t) + R(t) \tag{165}$$

where $R(t)$ is a random force. As has been shown in the case of a single-mode laser oscillation,[8] the Langevin equation (165) leads to the Fokker–Planck equation[27]

$$\frac{\partial}{\partial t} P(A, t) = \frac{\partial}{\partial A} [(a_1 A + a_3 A^3) P(A, t)] + \frac{\partial^2}{\partial A^2} [DP(A, t)] \tag{166}$$

The coefficient D depends upon the random force $R(t)$.

The asymptotic value of A and the associated macrofluctuations yield

$$A_m = 0, \qquad \langle \delta A^2 \rangle_t = \frac{4\,\Delta A^2}{a_1^2}\frac{1}{t^2} \qquad \text{for} \quad a_1 > 0 \tag{167}$$

and

$$A_m = \left(-\frac{a_1}{3a_3}\right)^{1/2}, \qquad \langle \delta A^2 \rangle_t = \frac{\Delta A}{6a_3}\left(-\frac{3a_3}{a_1}\right)^{1/2}\frac{1}{t} \qquad \text{for} \quad a_1 < 0 \tag{168}$$

with $\Delta A < -(3D/a_1)(-3a_3/a_1)^{1/2}$. Macrofluctuations increase their intensity as a_1^{-2} for $a_1 \to +0$ and as $|a_1|^{-1/2}$ for $a_1 \to -0$. In either case, non-Gaussian fluctuations diverge at the critical point $a_1 = \pm 0$. This will determine a critical threshold for laser oscillation if the Fokker–Planck equation (166) applies to a single-mode laser.

Furthermore, one could suppose that the nonlinear Fokker–Planck equation of van der Pol type (166) would govern macroscopic kinetics of oscillations in many cases, specifically for amplitude fluctuations. Non-Gaussian fluctuations have been observed near the threshold of electrical oscillation.[28] There is also an observation that current fluctuations are critically enhanced near the onset of Gunn oscillations.[29] If one follows the ansatz that probabilistic kinetics of oscillations on a macroscale be subject to the nonlinear Fokker–Planck equation (166) derived from the modified van der Pol equation (165), the observations of both non-Gaussian fluctuations and their divergences near the critical point might be explained in terms of macrofluctuations whose correlation functions have a long-time tail of inverse-time or inverse-squared-time characteristics. However, there is no justification for saying that the single-mode kinetics would be sufficient for describing the cooperative phenomena near the critical point.[8]

5.5. Remarks on Small-Parameter Expansions

It sometimes occurs that the Fokker–Planck equation includes a small parameter ϵ as in

$$\frac{\partial}{\partial t}P(A,t) = \frac{\partial}{\partial A}[g(A)P(A,t)] + \epsilon\frac{\partial^2}{\partial A^2}[h(A)P(A,t)] \tag{169}$$

In fact, this expression can be derived from the Kramers–Moyal expansion of the Pauli master equation in continuous phase space by neglecting higher than

[8] In case of oscillations, one needs at least two degrees of freedom, which sometimes reduce to amplitude and phase[5] (see Section 6). The coupling between the two modes

third-order differentials.[13] One observes two approximations in solving (169) which would seem to be rather independent of each other. One is to express the stochastic variable $A(t)$ as

$$A(t) = y(t) + \epsilon^{1/2}x(t) \tag{170}$$

with

$$\dot{y} = -g(y) \tag{171}$$

where $y(t)$ represents an orbit of secular motion with the neglect of the terms $O(\epsilon)$.[13] The other is to include the contribution of order ϵ in the secular motion by supposing the Gaussian distribution[14]

$$P_G(A, t) = [2\pi\epsilon\sigma(t)]^{-1/2} \exp\{-[A - y(t) - \epsilon u(t)]^2/2\epsilon\sigma(t)\} \tag{172}$$

Following van Kampen,[13] the Fokker–Planck equation with respect to the stochastic variable x in (170) around the secular motion $y(t)$ is expressible as

$$\frac{\partial}{\partial t} P_y(x, t) = \frac{\partial}{\partial x}\left\{\left[g'(y)x + \frac{1}{2}\epsilon^{1/2}g''(y)x^2 + \frac{1}{6}\epsilon g'''(y)x^3 + \cdots\right] \times P_y(x, t)\right\} + \frac{\partial^2}{\partial x^2}\{[h(y) + \epsilon^{1/2}h'(y)x + \cdots]P_y(x, t)\} \tag{173}$$

For simplicity, we discard the terms greater than x^3 in the first bracket on the right-hand side of (173) and those greater than x in the second one,

$$\frac{\partial}{\partial t} P_y(x, t) = \left\{\frac{\partial}{\partial x}\left[g'(y)x + \frac{1}{2}\epsilon^{1/2}g''(y)x^2\right]\cdot + \frac{\partial^2}{\partial x^2} h(y)\cdot\right\}P_y(x, t) \tag{174}$$

In the linear approximation, neglecting the term x^2, one obtains the intensity of macrofluctuations in x as

$$\langle x^2\rangle_t^{(1)} = \{4\,\Delta x^2/[g'(y)]^2\}(1/t^2) \tag{175}$$

with the aid of (147), where Δx is the size of coarse graining. On the other hand, if the nonlinear term x^2 is also considered, the fluctuation intensity in x will be

$$\langle x^2\rangle_t^{(2)} = [\Delta x/\epsilon^{1/2}g''(y)](1/t) \qquad \text{for} \quad g''(y) > 0 \tag{176}$$

---

would become sufficiently influential near the critical point since a coherent relationship between the two modes is finally established beyond that point. Hence it may be tempting to regard the collective dynamics of the two coupled modes near the critical point as that of interacting clusters each of which is a locally coherent state of the two modes. Both the size and the lifetime of such a cluster would diverge on approaching the critical point. See Ref. 37.

following (162). Hence the time interval t during which the linear Fokker–Planck equation may be influential compared with the nonlinear one in (174) must satisfy

$$\langle x^2 \rangle_t^{(1)} \gg \langle x^2 \rangle_t^{(2)} \tag{177}$$

This gives

$$t \ll \{4g''(y)\, \Delta x/[g'(y)]^2\}\epsilon^{1/2} \sim \epsilon^{1/2} \tag{178}$$

As a result, the linear Fokker–Planck equation obtained by employing the decoupling between the secular motion and fluctuations as shown in (170) with the use of the measure $\epsilon^{1/2}$ would remain valid only in a small time interval of order $\epsilon^{1/2}$ and could not deal with macrofluctuations with characteristic times greater than a time of $\sim\epsilon^{1/2}$. This also suggests that any nonlinear effect, no matter how small its strength might be, could finally overcome a linear effect so long as macrofluctuations lasting for a long time are concerned. A principal result remains unchanged even if a cubic term in (173) is taken into account.

The Gaussian ansatz (172) of Kubo *et al.*,[14] however, includes a certain correction to the secular motion within the order ϵ. The secular motion $y(t)$ and its correction $\epsilon u(t)$, within order ϵ, and the variance $\epsilon\sigma(t)$ are subject to

$$\dot{y}(t) = -g(y) \tag{179}$$

$$\dot{\sigma}(t) = -2g'(y)\sigma + h(y) \tag{180}$$

$$\dot{u}(t) = -g'(y)u - \tfrac{1}{2}g''(y)\sigma \tag{181}$$

Although the time needed for fixing the probability function (172) is of order of τ_{FP} in (93), another characteristic time $\tau^{\mathrm{G}}_{\mathrm{macro}}$ of macrofluctuations in the correction term $u(t)$ to the secular motion and in the variance $\epsilon\sigma(t)$ will be seen in the form

$$\tau^{\mathrm{G}}_{\mathrm{macro}} \sim 1/|g'_{\mathrm{asymp}}(y)| \tag{182}$$

from expressions (180) and (181), where $g'_{\mathrm{asymp}}(y)$ represents a value of $g'(y)$ in the asymptotic limit. Certain macrofluctuations are included in the scheme of (172), specifically those with the characteristic time $\tau^{\mathrm{G}}_{\mathrm{macro}}$ of (182). However, the Fokker–Planck equation (169) or (92) never excludes macrofluctuations whose correlation times are not less than some finite value so long as the broken microscopic reversibility shown in (101) and in (102) holds. Macrofluctuations with characteristic time greater than $\tau^{\mathrm{G}}_{\mathrm{macro}}$ of (182) are not covered by the Gaussian ansatz (172).

6. APPLICATION II: TWO DEGREES OF FREEDOM

Consider the coupled Langevin equations with two degrees of macroscopic freedom as

$$(d/dt)X(t) = -g(X, Y) + R_x(t) \tag{183}$$

$$(d/dt)Y(t) = -h(X, Y) + R_y(t) \tag{184}$$

in which the random forces $R_x(t)$ and $R_y(t)$ acting on macrovariables $X(t)$ and $Y(t)$, both of which are real, exhibit themselves as contributions from microscopic variables imbedded in a heat reservoir. We shall employ a coarse graining in time whose unit

$$\tau_{xy} \tag{185}$$

is supposed to be so coarse that the macrokinetics by the unit of τ_{xy} may be Markovian undergoing Gaussian fluctuations. Here the functions $g(X, Y)$ and $h(X, Y)$ represent systematic parts.

The Langevin equation is a slight modification of a nonlinear kinetic equation which one frequently meets in, say, work on autocatalytic reactions,[24] though in the latter, random forces are not usually considered. Heat reservoirs always play two roles. One is as an absorber for various irreversible processes and as a supplier of macroscopic currents. The other is as an agent providing microscopic and incoherent stimuli to relevant macroprocesses. Hence the latter contribution should also be taken into account. At this point one must note that random forces may not be totally arbitrary provided that a certain constraint is assumed such as the conservation law encountered in the Lotka–Volterra system.[30]

On averaging the coupled Langevin equations (183) and (184) over a short time of the order of τ_{xy} in (185), one obtains the Fokker–Planck equation[27]

$$\frac{\partial}{\partial t} P(X, Y; t) = \left\{ \frac{\partial}{\partial X} g(X, Y)\cdot + \frac{\partial}{\partial Y} h(X, Y)\cdot \right. \\ \left. + \frac{\partial^2}{\partial X^2} D_x\cdot + \frac{\partial^2}{\partial Y^2} D_y\cdot \right\} P(X, Y; t) \tag{186}$$

where $P(X, Y; t)$ is a probabilitiy density at (X, Y), and D_x and D_y are positive constants depending on the random forces R_x and R_y, respectively. Suppose that the unit ΔU for measuring X and Y is sufficiently small that (186) may be written as a difference equation as (94); an asymptotic macrostructure which is in broken microscopic reversibility would be realized at a maximum of

$$P^*_{(2)}(1) \equiv \exp\left[-(1/T_{(2)}) \int_{t_1}^{T_{(2)}+t_1} W(X, \Delta U; Y, \Delta U)\, dt \right] \tag{187}$$

with

$$W(X, \Delta U; Y, \Delta U) \equiv \frac{g(X, Y)}{\Delta U} + \frac{2D_x}{\Delta U^2} + \frac{h(X, Y)}{\Delta U} + \frac{2D_y}{\Delta U^2} \tag{188}$$

because of the minimum K-entropy (70), where the time $T_{(2)}$ is a character-

istic time representing a translational invariance expressed in (7). We have not written the right-hand side of (187) in the reduced form

$$\exp\left[\lim_{T\to\infty} -(1/T)\int_0^T W(X, \Delta U;\ Y, \Delta U)\, dt\right] \tag{189}$$

since fluctuations in the characteristic time $T_{(2)}$ will be examined, if any.

The system with two degrees of freedom keeps a phase relation between the macrovariables X and Y. If a certain invariance holds for temporal translation, the phase relation will also exhibit such an invariance.[5] Hence fluctuations in the characteristic time $T_{(2)}$ would correlate with those in the phase relation.

If the period $T_{(2)}$ deviates only slightly from its asymptotic value $T^*_{(2)}$ by δT,

$$T_{(2)} = T^*_{(2)} + \delta T, \qquad T^*_{(2)} \gg |\delta T| \tag{190}$$

during the time t ($\gg T_{(2)}$), one can calculate the mean square of fluctuations $\langle \delta T^2 \rangle_t$ following the general formula (56). The result is

$$\langle \delta T^2 \rangle_t = \frac{2(T^*_{(2)})^4}{[\ln P^*_{(2)m}(1)]^2}\frac{1}{t^2} \qquad \text{for} \quad t \gg T^*_{(2)} \tag{191}$$

with

$$P^*_{(2)m}(1) \equiv \max[P^*_{(2)}(1)] \tag{192}$$

If one introduces the quantities

$$\omega_0 \equiv 2\pi/T^*_{(2)} \tag{193}$$

and

$$\delta\omega \equiv (2\pi/T_{(2)}) - \omega_0 \tag{194}$$

expression (191) reduces to

$$\langle \delta\omega^2 \rangle_t = \{8\pi^2/[\ln P^*_{(2)m}(1)]^2\}(1/t^2) \tag{195}$$

in which $\delta T \ll T_{(2)}$. The newly defined quantity $\delta\omega$ is just a measure of the difference between the instantaneous frequency $2\pi/T_{(2)}$ and the fluctuation-free one ω_0. Hence phase fluctuations can be observed as frequency fluctuations. The resulting fluctuations in frequency around the asymptote do not obey the central limit theorem since their correlation function in the sense of (128) has the t^{-2} characteristic, suggesting that the largest correlation time is by no means restricted to a finite value. A specific example of experiments for measuring relative phase differences and their fluctuations is discussed in Ref. 31.

7. APPLICATION III: N DEGREES OF FREEDOM ($N \gg 1$)

Let us investigate a Markov process with N ($\gg 1$) coupled degrees of macroscopic freedom

$$\{X_1(t), X_2(t),..., X_N(t)\} \tag{196}$$

Following the general discussion developed in Section 4, the asymptotic macrostructure $S_{\{\alpha\}}$ of (69) having the least amount of K-entropy imposes on the N macrovariables the constraint

$$\{X_1(t), X_2(t),..., X_N(t)\} \in S_{\{\alpha\}} \tag{197}$$

In the μ-space representation one could perform observations more microscopically than in the Γ-space representation, in which only the representative point is followed. We shall consider the problem of how the macrokinetics of the representative point is related to that of each individual of N variables, all of which are supposed to be real.

One could consider $N(N-1)/2$ independent phase factors $\Theta_{ij}(t)$ defined as

$$\tan \Theta_{ij}(t) \equiv X_j(t)/X_i(t) \tag{198}$$

with

$$\Theta_{ji}(t) = \tfrac{1}{2}\pi - \Theta_{ij}(t) \tag{199}$$

Since we are interested in an asymptotic macrostructure in which a certain invariance of temporal translation holds, the phase factor $\Theta_{ij}(t)$ satisfies

$$\Theta_{ij}(t'') - \Theta_{ij}(t') = \Theta_{ij}(T_{ij} + t'') - \Theta_{ij}(T_{ij} + t') \tag{200}$$

with

$$T_{ji} \equiv T_{ij} \tag{201}$$

where T_{ij} is an unknown characteristic time exhibiting the invariance. Each phase factor $\Theta_{ij}(t)$ interacts with all the remaining phase factors because of the assumed nonvanishing couplings among them.

We consider the cross product $X_i(t')X_j(t')$ averaged over a sufficiently large time $\tilde{T}$ as

$$R_{ij}(t) \equiv (1/\tilde{T}) \int_t^{t+\tilde{T}} X_i(t')X_j(t')\, dt' \tag{202}$$

The product is also expressible as

$$X_i(t')X_j(t') = \tfrac{1}{2}A_{ij}^2(t') \sin 2\Theta_{ij}(t') \tag{203}$$

with

$$A_{ij}(t') \equiv [X_i^2(t') + X_j^2(t')]^{1/2} \tag{204}$$

with the use of (198).

On the other hand, the translational invariance expressed in (200) leads to the reduced form

$$\Theta_{ij}(nT_{ij} + t') - \Theta_{ij}(t') = (n - 1)[\Theta_{ij}(T_{ij} + t') - \Theta_{ij}(t')] \tag{205}$$

with $n = 1, 2, \ldots$. Hence for a sufficiently large n the quantity R_{ij} in (202) leads approximately to

$$R_{ij} = (1/n)\left[\sum_{k=1}^{n} \tfrac{1}{2}A_{ij}^2(kT_{ij} + t') \sin 2\Theta_{ij}(kT_{ij} + t')\right] \tag{206}$$

in which it must be noted that even if the limit $n \to \infty$ is taken, this should be understood in the sense of

$$\lim_{T \to \infty} (nT_{ij}/\tilde{T}) \sim 1 \tag{207}$$

On recalling the Jacobi's theorem,[9] one readily finds

$$R_{ij} \to 0 \qquad \text{for} \quad (1/2\pi)[\Theta_{ij}(T_{ij} + t') - \Theta_{ij}(t')] \quad \text{an irrational number} \tag{208}$$

since $A_{ij}^2(t)$, which is always positive and finite, exhibits a certain translational invariance whose characteristic time might not necessarily be T_{ij}.

The result expressed in (208) is just the statement for the presence of phase mixing, although it is a weak one in the language of ergodic theory.[17] If one considers the subset of macrovariables

$$\{X_{i'}(t')\} \qquad \text{with} \quad \{i'\} \subset (1, 2, 3, \ldots, N) \tag{209}$$

satisfying

$$(1/2\pi)[\Theta_{i'j'}(T_{i'j'} + t') - \Theta_{i'j'}(t')] \quad \text{an irrational number for } \forall i', \forall j' \ (\neq i') \in \{i'\} \tag{210}$$

a random phase characteristic resulting from

$$R_{i'j'} \to 0 \qquad \text{for} \quad i', j' \ (\neq i') \in \{i'\} \tag{211}$$

will necessarily follow for the subset so long as a sufficiently large $\tilde{T}$ in (202) is taken. Phase mixing among macrovariables could be regarded as turbulence.

Phase relations among the macrovariables $\{X_i(t)\}$ are not always involved in every observation. In some cases only the set of power

$$\bar{X}_i^2(t) \equiv (1/\tilde{T}) \int_t^{t+\tilde{T}} X_i(t')X_i(t')\, dt', \qquad \text{for} \quad i = 1, 2, 3, \ldots, N \tag{212}$$

may be investigated, where $\bar{X}_i^2(t)$ is defined at each interval of a sufficiently large time $\tilde{T}$. Kinetics of the power $\{\bar{X}_i^2(t)\}$ will be examined in the following. The time unit of the kinetics is $\tilde{T}$.

[9] See Appendix I of Ref. 17.

Let us introduce the probability measure

$$P(\Delta, t) \qquad \text{for} \quad \Delta \subset \bar{\Gamma} \tag{213}$$

in the phase space $\bar{\Gamma}$ spanned by the set of coordinates $\{\bar{X}_i^2\}$, in which $P(\Delta, t)$ denotes the probability of finding the power with the structure $\{\bar{X}_i^2\} \subset \Delta$ at time t. Since the set $\{\bar{X}_i^2\}$ is obtained as a result of a certain projection eliminating irrelevant variables such as

$$\langle X_i(t)X_j(t)\rangle \equiv (1/\tilde{T})\int_t^{t+\tilde{T}} X_i(t')X_j(t')\,dt', \qquad i \neq j \tag{214}$$

the probabilistic kinetics of $P(\Delta, t)$ will necessarily result because of a lack of fine information. The time unit for defining $P(\Delta, t)$ is also $\tilde{T}$.

A macroscopic property of the probability function $P(\Delta, t)$ is its conservation in the form

$$(d/dt)\int_{\bar{\Gamma}} P(\Delta', t)\,d\mu(\Delta') = 0 \tag{215}$$

where $\mu(\Delta')$ is the volume measure of the element Δ'. On introducing a coarse graining for the volume measure, one obtains

$$(d/dt)P(\Delta, t) + Y = 0 \tag{216}$$

with

$$Y \equiv \frac{1}{\mu(\Delta)}\frac{d}{dt}\int_{\Gamma-\Delta} P(\Delta', t)\,d\mu(\Delta') \tag{217}$$

where $\mu(\Delta)$ is a sufficiently small quantity compared with $\mu(\bar{\Gamma})$. The first term on the left-hand side of (216) leads to

$$\frac{d}{dt}P(\Delta, t) = \frac{\partial}{\partial t}P(\Delta, t) + \sum_i \frac{\partial}{\partial \bar{X}_i^2}P(\Delta, t)\frac{\partial \bar{X}_i^2}{\partial t} \tag{218}$$

Furthermore, we shall introduce a small unit $\hat{X}^2$ for measuring any member of the set $\{\bar{X}_i^2\}$, leading to the approximate equality

$$(\partial/\partial \bar{X}_i^2)P(\Delta, t) = (1/\hat{X}^2)[P(\Delta + \Delta_i, t) - P(\Delta, t)] \tag{219}$$

where the arguments of $P(\Delta + \Delta_i, t)$ are identical to those of $P(\Delta, t)$ except for replacing $\bar{X}_i^2$ by $\bar{X}_i^2 + \hat{X}^2$.

Hence one can write (216) in the form

$$(\partial/\partial t)P(\Delta, t) = \sum_i W(\Delta \leftarrow \Delta + \Delta_i)P(\Delta + \Delta_i, t) - W(\Delta)P(\Delta, t) - Y \tag{220}$$

with

$$W(\Delta \leftarrow \Delta + \Delta_i) \equiv -(1/\hat{X}^2)\,\partial \bar{X}_i^2/\partial t \tag{221}$$

$$W(\Delta) \equiv -\sum_i (1/\hat{X}^2)\,\partial \bar{X}_i^2/\partial t \tag{222}$$

If the probabilistic kinetics presented in (220) is Markovian, the quantity $W(\Delta)$ in (222) is just the irreversible decay rate of the structure $\{\bar{X}_i^2\} \subset \Delta$ since Y does not include the term $P(\Delta, t)$. This process is equivalent to a so-called diffusion approximation. We have chosen $\tilde{T}$ in (212) as a time unit for describing the kinetics. So long as $\tilde{T}$ is sufficiently large, the phase mixing property shown in (208) would give plausible arguments for the establishment of Markovian kinetics.

If the Markovian postulate is approved, the least irreversible decay rate will determine the most probable asymptotic structure. In fact, turbulence as an asymptotic structure of many coupled degrees of freedom is argued within the present scheme of phase mixing.

8. APPLICATION IV: EQUILIBRIUM STATISTICAL MECHANICS

A completely isolated system S_{comp} with N_{comp} ($\gg 1$) degrees of freedom is sometimes referred to when one considers statistical mechanics. Since the isolated system has a Hamiltonian H_{comp} which is Hermitian, its eigenvalue is real. For a given eigenvalue E there is a set of degenerate eigenstates, which are, of course, macroscopic,

$$\{|i\colon E\rangle_{\text{comp}}\}, \qquad i = 1, 2, 3,... \tag{223}$$

with

$$_{\text{comp}}\langle E\colon i|j\colon E\rangle_{\text{comp}} = 0 \qquad \text{for} \quad i \neq j \tag{224}$$

where $_{\text{comp}}\langle E\colon i|$ is the Hermitian conjugate to $|i\colon E\rangle_{\text{comp}}$. The orthogonality (223) prevents any transition between $|i\colon E\rangle_{\text{comp}}$ and $|j\colon E\rangle_{\text{comp}}$ if $i \neq j$. This forbiddenness would not reflect the real physical characteristics of macroscopic states. Hence one would be motivated to introduce an almost isolated system instead of the former completely isolated one, after Landau.[21]

Let us suppose an almost isolated system S_{alm} with N_{alm} degrees of freedom inside the system S_{comp} satisfying

$$(N_{\text{comp}} - N_{\text{alm}})/N_{\text{comp}} \ll 1 \tag{225}$$

and

$$N_{\text{aux}} \equiv N_{\text{comp}} - N_{\text{alm}} \gg 1 \tag{226}$$

where the auxiliary system S_{aux} has N_{aux} degrees of freedom. We shall further introduce the Hamiltonians H_{alm} and H_{aux} for the systems S_{alm} and S_{aux}, respectively. If there were no interaction between S_{alm} and S_{aux}, the total Hamiltonian H_{comp} for the completely isolated system would be

$$H_{\text{comp}} = H_{\text{comp}}^{(0)} \tag{227}$$

with

$$H_{\text{comp}}^{(0)} \equiv H_{\text{alm}} + H_{\text{aux}} \tag{228}$$

Since the energy of the system S_{comp} must be conserved, one obtains

$$E = E_{\text{alm}}(i) + E_{\text{aux}}(j) \tag{229}$$

with

$$H_{\text{alm}}|i\rangle = E_{\text{alm}}(i)|i\rangle, \qquad i = 1, 2, 3,... \tag{230}$$

$$H_{\text{aux}}|j\rangle_x = E_{\text{aux}}(j)|j\rangle_x, \qquad j = 1, 2, 3,... \tag{231}$$

where $\{|i\rangle\}$ and $\{|j\rangle_x\}$ are the orthonormal eigenstates with the energies $\{E_{\text{alm}}(i)\}$ and $\{E_{\text{aux}}(j)\}$ for the two independent systems S_{alm} and S_{aux}, respectively.

One should, however, note that interactions between S_{alm} and S_{aux} could never be excluded. Furthermore, such interactions always occur through adiabatic processes since the total system S_{comp} is completely isolated. This would make it feasible to write the total Hamiltonian in the form

$$H_{\text{comp}} = H^{(0)}_{\text{comp}} + H_{\text{int}}(t, \epsilon) \tag{232}$$

with

$$H_{\text{int}}(t, \epsilon) \equiv e^{-|t|\epsilon} H_{\text{int}} \qquad \text{for} \quad \epsilon \to +0 \tag{233}$$

where H_{int} is an interaction Hamiltonian between S_{alm} and S_{aux}. The factor $e^{-|t|\epsilon}$ means that the interaction could gradually be switched on and off for *each* transition of macroscopic state from, say, $|i\rangle|j\rangle_x$ to $|i'\rangle|j'\rangle_x$, where $|i\rangle|j\rangle_x$ denotes the direct product of $|i\rangle$ and $|j\rangle_x$.

One also notes that the interaction Hamiltonian H_{int} is not Hermitian since the total energy in (229) must be conserved.[10] If H_{int} were Hermitian, a certain eigenstate $|i''\rangle|j''\rangle_x$ of the total system would give

$$_x\langle j''|\langle i''|H_{\text{int}}|i''\rangle|j''\rangle_x \neq 0 \tag{234}$$

[10] The interaction Hamiltonian H_{int} giving rise to transitions in the system imbedded in a heat reservoir will be a Hermitian operator if the interaction energy is observable. In fact, the time-dependent perturbation theory always requires that the interaction energy may be observable. Hence the master equation based upon the time-dependent perturbation theory satisfies the microscopic reversibility between conjugate transition probabilities because of the Hermitian characteristic of the interaction. However, any perturbation theory has a drawback in that transition probabilities may not be conserved. This tells us that if and only if a reaction from the heat reservoir with the present system is ignored will the master equation with microscopic reversibility keep its validity.

On the other hand, if an almost isolated system is the case, the conservation of transition probabilities must be borne in mind because of the finiteness of the system no matter how large it may be. The time-dependent perturbation theory, which admits both the breakdown in the conservation of transition probabilities and a Hermitian interaction Hamiltonian, cannot work any more. Because of the isolatedness of S_{comp}, the interaction between S_{alm} and S_{aux} must be adiabatic. The interaction energy between S_{alm} and S_{aux} may not be observable once the complete sets $\{|i\rangle\}$ and $\{|j\rangle_x\}$ are defined for S_{alm} and S_{aux}, respectively, since the interaction would always cause

This would apparently lead to contradiction, since the energy eigenvalue of the Hamiltonian (232)

$$ {}_x\langle j''|\langle i''|H_{\text{comp}}|i''\rangle|j''\rangle_x \tag{235} $$

would not equal the conserved energy E, leading to

$$ E = E_{\text{alm}}(i'') + E_{\text{aux}}(j'') \tag{236} $$

if the Hermitian property of the form (234) held. One can write such a non-Hermitian characteristic as

$$ {}_x\langle j'|\langle i'|H_{\text{int}}|i\rangle|j\rangle_x \neq {}_x\langle j|\langle i|H_{\text{int}}|i'\rangle|j'\rangle_x \tag{237} $$

for certain pair states which are orthogonal,

$$ {}_x\langle j'|\langle i'|i\rangle|j\rangle_x = 0 \tag{238} $$

One more characteristic about the interaction H_{int} is its weakness. Since one can choose the auxiliary system S_{aux} as small as possible as expressed in (225), the strength of the interaction may be made arbitrarily small though it would never vanish. Hence one may think that the dynamics of each transition between different macroscopic states would be subject to a certain adiabatic interaction of sufficiently low strength.

The standard S-matrix formulation with iteration procedure or, more specifically, the covariant perturbation theory is capable of accounting for a process with the adiabatic switching of weak interactions. Within the lowest iteration expansion the S-matrix element

$$ {}_x\langle j'|\langle i'|S|i\rangle|j\rangle_x $$

for the transition from the macroscopic state $|i\rangle|j\rangle_x$ to $|i'\rangle|j'\rangle_x$ is found to be

$$ \begin{aligned} {}_x\langle j'|\langle i'|S|i\rangle|j\rangle_x &= 2\pi {}_x\langle j'|\langle i'|H_{\text{int}}|i\rangle|j\rangle_x \\ &\quad \times \delta(E_{\text{alm}}(i') + E_{\text{aux}}(j') - E_{\text{alm}}(i) - E_{\text{aux}}(j)) \end{aligned} \tag{239} $$

---

some changes in the reference states $\{|i>\}$ and $\{|j>_x\}$ so as to satisfy the conservation of both transition probability and energy. The states $\{|i>\}$ and $\{|j>_x\}$ could be defined only for in and out states of each interaction. On admitting that we are interested in the states of S_{alm} before and after each interaction with S_{aux}, the present observation will force the interaction Hamiltonian H_{int} to be non-Hermitian.

Consequently, the master equation with microscopic reversibility, which would in turn lead to the ergodic hypothesis (cf. Ref. 4), could be established only in a small system surrounded by a heat reservoir.

The corresponding transition probability per unit time can thus be expressed as

$$W(i', j_x' \leftarrow i, j_x) = (2\pi/\hbar)|_x\langle j'|\langle i'|H_{\text{int}}|i\rangle|j\rangle_x|^2 \times \delta(E_{\text{alm}}(i') + E_{\text{aux}}(j') - E_{\text{alm}}(i) - E_{\text{aux}}(j)) \tag{240}$$

where $\hbar$ is the Planck's constant divided by 2π. Finally, one finds the non-symmetric property

$$W(i', j_x' \leftarrow i, j_x) \neq W(i, j_x \leftarrow i', j_x') \tag{241}$$

because of the non-Hermitian characteristic (237).

The broken symmetry in (241) suggests that the almost isolated system S_{alm} does not exhibit microscopic reversibility for an adiabatic transition process which occurs on a macroscale.

The probabilistic kinetics of the transition process can be written in the form of the kinetic equation of the density matrix, that is, the Pauli master equation, as

$$(\partial/\partial t)\rho(i, t) = \sum_{i' \neq i} W(i \leftarrow i')\rho(i', t) - \sum_{i' \neq i} W(i' \leftarrow i)\rho(i, t) \tag{242}$$

with

$$W(i' \leftarrow i) \equiv \sum_{j,j'} W(i', j_x' \leftarrow i, j_x)$$

where the diagonal element of the density matrix $\rho(i, t)$ includes the trace as for the states of the auxiliary system S_{aux}. On introducing a certain coarse graining in the space of macroscopic states, the probabilistic equation (242) reduces to

$$(\partial/\partial t)\rho^c(i, t) = \sum_{j \neq i} W^c(i \leftarrow j)\rho^c(j, t) - \sum_{j \neq i} W^c(j \leftarrow i)\rho^c(i, t) \tag{243}$$

with

$$W^c(i \leftarrow j) \neq W^c(j \leftarrow i) \tag{244}$$

where a superscript c means that a certain coarse graining is taken, and where the index i runs over each coarse-grained element of the space of those macroscopic states. In the present coarse graining a fixed number n_c ($\gg 1$) of macroscopic quantum states is identified with a coarse-grained macroscopic state. One can proceed with further discussion by following a similar analysis as developed in Section 4, replacing the term phase space by the space of macroscopic quantum states.

The asymptotic macrostructure of the almost isolated system will be found to be associated with the coarse-grained element $S_{\{\alpha\}}$ of (69), which

necessarily leads to the minimization of the K-entropy as expressed in (70b) or, more generally, in (40). The volume measure of the local element $S_{\{\alpha\}}$ would be less than that of the total macroscopic states available to the system S_{alm}. This is due to the fact that the probabilistic kinetics presented in (242) and (243) does not exhibit the microscopic reversibility between a pair of macroscopic states as exemplified in (241) and (244). The macrostructure in which the representative quantum state is found anywhere inside the local space $S_{\{\alpha\}}$ cannot be further clarified if the coarse-grained macroscopic state $S_{\{\alpha\}}$ is asymptotically recognized in the sense of (86) by an observer employing the time unit τ_0 of (88) for measurement. If one wishes to investigate the more microscopic kinetics of the representative quantum state inside the coarse-grained macroscopic state $S_{\{\alpha\}}$, finer grainings for both the space of the macroscopic states and the time unit for measurement must be taken.

Suppose that the characteristic τ_γ originally defined in (30) is evaluated from the probabilistic kinetic equation (243) by following the discussion developed in Section 4. The time unit τ_0 for the coarse graining must be much greater than τ_γ, as seen from (88). If the observer concerns himself only with time-dependent phenomena, or fluctuations whose characteristic time is much less than τ_γ, it would not always be strictly required to follow fluctuations within a scheme of the adiabatic processes leading to the all-purpose probabilistic kinetics as shown in (243). Although the presence of adiabatic processes on a macroscale must always be kept in mind particularly when macrokinetics is concerned, it is often noted that any small subsystem imbedded in the almost isolated system S_{alm} exhibits rather microscopic fluctuations which would seem to be caused by a sort of thermodynamic interaction with its outside acting like a heat reservoir. Such thermodynamic fluctuations include more than just adiabatic fluctuations, e.g., they also include isothermal fluctuations.

Hence it is appropriate to assume, in the sense of a certain approximation, nonadiabatic fluctuations as if they originated in an interaction with a heat reservoir only if fluctuations whose characteristic time is much less than τ_γ are concerned. A small local portion inside the almost isolated system which is always subject to adiabatic processes on a macroscale could sometimes be viewed as if undergoing nonadiabatic fluctuations by an observer who does not care about the details of the kinetics outside the local region. However, one could not imagine nonadiabatic fluctuations whose characteristic times are of the order of or much greater than τ_γ, since all kinetics with respect to the almost isolated system as a whole must be adiabatic. If and only if one considers a rather small-scaled kinetics which would not be greatly influenced by the totality of the almost isolated system would it be admissible to relax to a certain extent the strict condition on adiabatic processes. When one employs a statistical mechanical description of an almost isolated system,

as is often the case, it should be remarked that nonadiabatic fluctuations could not become sufficiently macroscopic, particularly on the time scale involved.

At first sight it might seem to be of purely academic interest to investigate macrofluctuations, whether adiabatic or not, inside an almost isolated system. This, however, becomes of practical importance if phase transitions in equilibrium are considered. It is commonly believed that as the critical point is approached both the divergence of the correlation length and the critical slowing down would take place (see, e.g., Ref. 32). In fact, these views have strong support from experiments (see, e.g., Ref. 33) and have been verified to a certain extent in some model systems.[34,35] Nevertheless, one may raise the question of whether or not such macrofluctuations leading to both the divergence of the correlation length and the critical slowing down would be nonadaiabatic. In an almost isolated system the most macroscopic fluctuations have been argued to be adiabatic instead of nonadiabatic. If one follows the canonical ensemble theory of statistical description which reproduces equilibrium thermodynamics, such a violation of the conjugate property in fluctuations could not be expected since thermodynamics never excludes nonadiabatic fluctuations. Although it might be thought that if one makes the almost isolated system infinitely large, both adiabatic and nonadiabatic fluctuations could consistently be discussed within the canonical ensemble theory, one may not be allowed to take such a limit to infinity, particularly from a phenomenological viewpoint accounting for real experiments.

One should stress the dominance of adiabatic macrofluctuations over the nonadiabatic ones near the critical point for a phase transition. Furthermore, the adiabatic macrofluctuations could not follow the canonical ensemble theory, since it would be inappropriate to suppose an infinitely large heat reservoir outside the almost isolated system. There have been a few attempts to account for the adiabatic macrofluctuations on a heuristic basis. Near the critical point one might describe the adiabatic process, which is macroscopic, as the appearance of an unstable phase as local clusters in an overwhelmingly large area of stable phase.[36,37] There is an observation that the structural transition encountered in ferroelectrics and antiferroelectrics would exhibit a collective excitation, which is sometimes referred to as a central peak of the excitation spectrum located at an almost vanishing wave number,[38,39] manifesting the presence of the macroscopically adiabatic process.[40,41]

The critical dynamics of the phase transition is an example to indicate how the canonical ensemble theory differs from the probabilistic kinetics without microscopic reversibility. In general, the ensemble theory of statistical mechanics employs an ensemble average of any event appearing in an infinite time interval. Even macroscopic events lasting over a sufficiently large time could not be excluded as members constituting the ensemble.

However, if one follows the canonical ensemble theory based upon the principle of a priori equal weight, this would impose a certain restriction on events appearing as a member of the relevant ensemble. If the principle is examined from the viewpoint of probabilistic kinetics, one must assume microscopic reversibility by any means. Otherwise, a certain imbalance in the probability density of the macroscopic state would result as seen in Sections 3 and 4. This would apparently contradict the principle of a priori equal weight.

It has already been pointed out that the probabilistic kinetics of an almost isolated system which is subject to fluctuations solely from adiabatic interactions with a small auxiliary system remaining inside a completely isolated system does not exhibit symmetry between conjugate transition probabilities. A characteristic property of broken microscopic reversibility is seen in the macrokinetics. If the probabilistic kinetic equation without microscopic reversibility is linear with respect to a certain variable, the stochastic average of the squared intensity of its fluctuations lasting over the time t will lead to an inverse-squared-time characteristic t^{-2} as discussed in Section 5. On the other hand, if the equation is nonlinear, the averaged squared intensity of the corresponding fluctuations will yield an inverse-time characteristic t^{-1}. The averaged intensity of squared fluctuations could be viewed as a time correlation function if a sort of low-pass filter with a sufficiently small cutoff frequency is used for extracting only macroscopic events. In the linear case the macrofluctuations with correlation function proportional to t^{-2} for a large time t would not display any singular spectrum. However, in the nonlinear case one expects a long-time singularity or a logarithmic divergence in the low-frequency spectrum since the correlation function of macrofluctuations is proportional to t^{-1} in a long-t limit.[11] One might expect that a nonlinear dynamics characteristic of a phase transition exhibiting a low-frequency singularity known as the critical slowing down could also be discussed within the present adiabatic kinetics leading to broken microscopic reversibility.

ACKNOWLEDGMENTS

The author wishes to thank Prof. Kazuhisa Tomita for sending and explaining his result before its publication, Dr. Ki-ichi Nakamura for

[11] A nonlinear system in which linear decay processes are absent shows a distinct contrast with a linear one even near equilibrium. If one wishes to establish a certain relation between the fluctuation intensity of a fast process and the kinetic behavior of a slow process, the characteristic time distinguishing between the fast and slow processes must not diverge since nonlinear fluctuations have a logarithmic singularity in their low-frequency spectrum. The principle of detailed balance would not be obeyed in a nonlinear system since the principle excludes slow processes.

conversations on various aspects of the present subject, and Dr. Yoshihiro Iwadare for helpful discussions on nonlinear processes.

REFERENCES

1. H. Mori, *Progr. Theor. Phys.* **33**:423 (1965).
2. H. Mori, *Progr. Theor. Phys.* **49**:1516 (1973).
3. L. Onsager, *Phys. Rev.* **37**:405 (1931); **38**:2265 (1931).
4. D. ter Haar, *Elements of Statistical Mechanics*, Holt, Rinehart and Winston, New York (1961).
5. K. Tomita and H. Tomita, *Phys. Letters A* **46**:265 (1973); *Progr. Theor. Phys.* **51**:1731 (1974).
6. N. G. van Kampen, *Physica* **67**:1 (1973).
7. H. Ueyama, *Progr. Theor. Phys.* **48**:1090 (1972).
8. R. Graham and H. Haken, *Z. Phys.* **237**:31 (1970).
9. S. F. Edwards and W. D. McComb, *J. Phys. A* **2**:157 (1969).
10. M. Nelkin, *Phys. Rev. Letters* **30**:1028 (1973).
11. R. Graham, *Phys. Rev. Letters* **31**:1479 (1973).
12. K. Nakamura, private communication.
13. N. G. van Kampen, *Canad. J. Phys.* **39**:551 (1961).
14. R. Kubo, K. Matsuo, and K. Kitahara, *J. Stat. Phys.* **9**:51 (1973).
15. K. Matsuno, *Progr. Theor. Phys.* **50**:327 (1973).
16. K. Matsuno, *J. Math. Phys.* **15**:No. 11 (1974).
17. V. I. Arnold and A. Avez, *Ergodic Problems of Classical Mechanics*, Benjamin, New York (1968).
18. G. M. Zaslavskii and B. V. Chirikov, *Usp. Fiz. Nauk* **105**:3 (1971) [English transl.: *Soviet Phys.—Uspekhi* **14**:549 (1972)].
19. A. N. Kolmogorov, *Doklady Akad. Nauk SSSR* **119**:861 (1958); **124**:754 (1959).
20. K. Matsuno, *Phys. Letters A* **47**:99 (1974).
21. L. D. Landau and E. M. Lifshitz, *Statistical Physics*, Pergamon Press, Oxford (1969).
22. E. P. Wigner, *Symmetries and Reflection: Scientific Essays of Eugene P. Wigner*, MIT Press, Cambridge, Massachusetts (1970), p. 171.
23. Ya. G. Sinai, *Doklady Akad. Nauk SSSR* **124**:768 (1959); *Izv. Math. Nauk* **25**:899 (1961); **30**:15 (1966).
24. P. Glansdorff and I. Prigogine, *Thermodynamic Theory of Structure, Stability and Fluctuations*, Interscience, New York (1971).
25. P. Glansdorff and I. Prigogine, *Physica* **46**:344 (1970).
26. J. R. Tucker and B. I. Halperin, *Phys. Rev. B* **3**:3768 (1971); B. I. Halperin, P. Hohenberg, and S. K. Ma, *Phys. Rev. Letters* **29**:1548 (1972).
27. R. L. Strantonovich, *Topics in the Theory of Random Noise*, Vol. 1, Gordon and Breach, New York (1963).
28. T. Kawakubo, S. Kabashima, and K. Nishimura, *J. Phys. Soc. Japan* **34**:1460 (1973).
29. K. Matsuno, *Phys. Letters* **31A**:335 (1970).
30. E. H. Kerner, *Bull. Math. Biophys.* **19**:121 (1957); **21**:217 (1959); **26**:333 (1964).
31. K. Matsuno, *IEEE Trans. on Electron Devices* **ED-16**:1025 (1969).
32. H. E. Stanley, *Introduction to Phase Transitions and Critical Phenomena*, Oxford University Press, Oxford (1971).
33. P. Heller, *Rept. Prog. Phys.* **30**:731 (1967).
34. L. Onsager, *Phys. Rev.* **65**:117 (1944).

35. E. Lieb, *Phys. Rev. Letters* **18**:1046 (1967); **19**:108 (1967).
36. K. Matsuno and H. E. Stanley, *Phys. Letters* **33A**:425 (1970).
37. K. Matsuno, *J. Phys. Soc. Japan* **33**:1230 (1972); **35**:367 (1973); *Progr. Theor. Phys.* **50**:764 (1973).
38. T. Riste, E. J. Samuelson, K. Otnes, and J. Feder, *Solid State Commun.* **9**:1455 (1971).
39. S. M. Shapiro, J. D. Axe, G. Shirane, and T. Riste, *Phys. Rev. B* **6**:4332 (1972).
40. J. Feder, *Solid State Commun.* **9**:2021 (1971).
41. T. Schneider and E. Stoll, *Phys. Rev. Letters* **31**:1254 (1973).