Fluctuations and Irreversible Thermodynamics*

LASZLO TISZA AND IRWIN MANNING†
Department of Physics and the Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts
(Received August 16, 1956)

The time-dependent theory of fluctuations is based on a combined application of the phenomenological theory of dissipation and the stochastic theory of random processes. The traditional method of joining these theories into a uniform scheme proceeds by adding a random perturbation to the differential equation of the phenomenological kinetic theory (the Langevin equation in case of Brownian motion). In the present approach the problem is considered as an essentially statistical one. The role of the differential equation is to fix the form of the distribution function over the manifold of fluctuation paths in function space. The solutions of the equation constitute the most probable region in function space and the fluctuations appear with their appropriate probabilities. The connection between the phenomenological equation and the distribution function is stipulated by means of a postulate, the essential ingredient of which is the auxiliary function recently introduced by Onsager and Machlup. Heuristically this postulate was suggested by the kinetic analog of Boltzmann's principle established by these authors. In the present logical structure it is sufficient to join this postulate to the standard assumptions of the phenomenological and the stochastic theories for the derivation of the entire time-dependent fluctuation theory. A number of statements usually postulated appear as theorems in this presentation. The reversible and irreversible aspects of time play an essential part in the argument. The calculation of fluctuations is carried out in the temporal and the spectral descriptions. The relation of the two schemes is discussed along with the scope and limitations of the theory.

1. INTRODUCTION

THE thermodynamic theory of fluctuations deals with two types of problems. On a simpler level one considers the fluctuation of a quantity as registered at a single instant of time over an ensemble of systems.

In the problems of the second type one observes each member of the ensemble over a time interval rather than a single instant, yielding a fluctuation path $\alpha=\alpha(t)$ for each system; we are concerned with the distribution of these paths over an ensemble.

This classification of fluctuation problems corresponds closely with the division of phenomenological thermodynamics into *thermostatics* dealing with equilibrium and *irreversible thermodynamics;* or *kinetics* devoted to the study of time-dependent dissipative processes.

Whenever ambiguity might arise we shall qualify a fluctuation problem as "time-independent," "static" or "thermostatic," on the one hand and as "time-dependent" or "kinetic" on the other.

The time-independent theory is based on Einstein's[1] interpretation of Boltzmann's principle. The definition of the entropy function is extended to a class of nonequilibrium states and used to establish a probability density function (p.d.f.).

Recently Onsager and Machlup[2,3] proved a theorem which is the kinetic analog of the Boltzmann-Einstein principle. They established a probability distribution function (p.d.f.) for fluctuation paths in terms of an auxiliary function (henceforth: the OM function), which in turn was defined by means of the phenomenological kinetic equation.

The Boltzmann principle plays a central role in the time-independent theory. It allows one to compute fluctuation moments and, more generally, serves as a point of departure for the theory of the generalized canonical ensembles.

In this paper we embark on a program in which the Onsager-Machlup theorem is used as a similar organizing principle for the development of the time-dependent theory.

According to the central idea of this theory there is close connection between fluctuation and dissipation phenomena.[2–10] In the usual approach this connection is accounted for by the method of stochastic differential equations. The phenomenological equations are augmented by a random term producing the correct amount of fluctuation. The equation obtained in case of Brownian motion is called the Langevin equation.[2,6]

We shall show in this paper that the same problem can be attacked also from another point of view. Instead of focusing attention on differential equations, one considers the set of fluctuation paths and describes various physical situations in terms of appropriate distribution functions. The statistical point of view

* This paper is a development of the Ph.D. thesis submitted by Irwin Manning to the Department of Physics, Massachusetts Institute of Technology. The work was supported in part by the Army (Signal Corps), the Air Force (Office of Scientific Research, Air Research and Development Command), and the Navy (Office of Naval Research).

† Present address: Physics Department, Syracuse University, Syracuse, New York.

[1] A. Einstein, Ann. Physik **33**, 1275 (1910).
[2] L. Onsager and S. Machlup, Phys. Rev. **91**, 1505 (1953).
[3] S. Machlup and L. Onsager, Phys. Rev. **91**, 1512 (1953).
[4] A. Einstein, Ann. Physik **17**, 549 (1905).
[5] P. Langevin, Compt. rend. **146**, 520 (1908).
[6] H. Nyquist, Phys. Rev. **32**, 110 (1928).
[7] L. Onsager, Phys. Rev. **37**, 405 (1931) and **38**, 2265 (1931).
[8] H. B. Callen and R. F. Greene, Phys. Rev. **86**, 702 (1952).
[9] M. C. Wang and G. E. Uhlenbeck, Revs. Modern Phys. **17**, 323 (1945).
[10] N. Hashitsume, Progr. Theoret. Phys. (Japan) **8**, 461 (1952) and *Proceedings of the International Conference of Theoretical Physics, Kyoto and Tokyo, September, 1953* (Science Council of Japan, Tokyo, 1954), p. 495.

inherent in this approach is similar to the one taken in the static fluctuation theory. However, the situation is complicated by the emergence of time as an independent variable. To clarify some of the problems connected with the role of time is the main purpose of this paper.

Sections 2 and 3 contain a short summary of the basic definitions and postulates of the phenomenological theory of dissipation and the stochastic theory of fluctuations. In Sec. 4 we formulate a postulate which allows us to combine these disparate elements into a unified theory of kinetic fluctuations. We shall arrive at a derivation of the OM p.d.f. which differs from the original proof of Onsager and Machlup. The behavior of the OM function on time reversal plays an essential part in the argument.

We then go on to deal with the calculation of averages. The theory exhibits a duality corresponding to two methods for the description of fluctuation paths. In the *temporal description* the values of the fluctuating quantity are specified within narrow limits at closely spaced time intervals. In the *spectral description* one deals with Fourier coefficients considered as random variables.

We shall show that the OM p.d.f. has two equivalent formulations, one in the temporal, the other in the spectral description. We demonstrate the application of the p.d.f. in both schemes.

The theory will be developed within certain restrictions to be specified in detail later. Some of these assumptions, such as the limitation to a single fluctuating variable, are made primarily in the interest of simplicity. We hope to make extensions to more general situations in a future paper. On the other hand, assumptions such as the limitations to Gaussian Markoff processes and to linear phenomenological equations represent more essential restrictions on the scope of the present method.

2. PHENOMENOLOGICAL RELATIONS

a. Thermostatistics

Let us consider a closed thermodynamic system which is specified in equilibrium by its fundamental equation

$$S=S(Q_1,Q_2,\cdots Q_r), \tag{2-1}$$

where S is the entropy and the Q_i are a set of extensive variables obeying conservation laws. Associated with each variable Q_i is a conjugate thermostatic force,

$$F_i=\partial S/\partial Q_i, \quad i=1, 2, \cdots r. \tag{2-2}$$

The most important pairs of conjugate variables are U, $1/T$; N_γ, $-\mu_\gamma/T$; q, $-\varphi/T$ where U is the internal energy, T is the absolute temperature, N_γ the number of moles of the chemical species γ, μ_γ the chemical potential of this species, q the electric charge, and φ the electrostatic potential.

The thermostatic theory based on Eq. (2-1) applies only to systems which have reached their state of thermodynamic equilibrium. In order to extend the formalism to *processes*, we use the device of *composite systems*. The original system is partitioned into closed subsystems each of which is allowed to reach equilibrium. A process is initiated by the relaxation of an internal constraint; it consists of the redistribution of the quantities Q_i among the subsystems, a transition from a constrained to an unconstrained (or less constrained) equilibrium.

In this paper we shall consider only the one-variable case in which a single quantity Q is redistributed between two subsystems.

The entropy of the composite system is in this case

$$S(Q';Q'')=S'(Q')+S''(Q''), \tag{2-3}$$

where S' and S'' refer to the subsystems. For all processes in the composite system,

$$Q'+Q''=\text{constant}.$$

The deviation from unconstrained equilibrium is described in terms of the variable

$$\alpha=Q'-Q_e'=-(Q''-Q_e''), \tag{2-4}$$

where Q_e, Q_e' are the equilibrium values.

We define the thermodynamic force as

$$X=\frac{\partial S}{\partial \alpha}=\frac{\partial S'}{\partial Q'}-\frac{\partial S''}{\partial Q''}=F'-F''. \tag{2-5}$$

In unconstrained equilibrium $X=0$.

We note that the process corresponding to the differentiation (2-2) leaves the system homogeneous, but not closed, while the process of increasing Q is performed. In contrast, (2-5) refers to an internal displacement within a closed, but no longer homogeneous system.

We shall need the expansion of the total entropy around equilibrium

$$S=S_0-\tfrac{1}{2}s\alpha^2+\cdots. \tag{2-6}$$

Here

$$s=-\frac{\partial^2 S}{\partial \alpha^2}=-\frac{\partial^2 S'}{\partial Q'^2}-\frac{\partial^2 S''}{\partial Q''^2}. \tag{2-7}$$

It is a consequence of the entropy maximum principle that $s>0$. If the primed system is a reservoir then

$$\frac{\partial^2 S'}{\partial Q'^2}=\frac{\partial F'}{\partial Q'}=0.$$

There are two obvious physical interpretations of the single-variable theory. First, we have $Q=U$, $F=1/T$, and $s=1/(T^2C_v)$, where C_v is the constant volume heat capacity; the process is heat conduction

under conditions of negligible thermal expansion ($C_v \simeq C_p$).

Second, we have $Q=q$, $F=\varphi/T$, $s=1/(TC)$, where C is the electric capacity; the process is electric conduction.[11]

In all that follows, we shall assume that the composite system is close enough to unconstrained equilibrium (α small) so that the series (2-6) can be broken off after the second term. To within this approximation the thermodynamic force (2-5) becomes

$$X=-s\alpha. \tag{2-5a}$$

b. Kinetics

We now turn to the temporal aspects of the regression process. From the thermostatic point of view we can merely say that relaxation of a constraint leads eventually to a new equilibrium with an entropy which is not smaller than the initial entropy.

In irreversible thermodynamics, or kinetics, we describe the rate of the regression process in terms of a phenomenological equation connecting the forces and fluxes. We assume this relation to be linear. Hence, in the single variable theory we have

$$R\dot{\alpha}=X. \tag{2-8}$$

Here X is defined by (2-5), $\dot{\alpha}$ is the current or "flux" within the composite system and R is the resistance of the partition or of the connecting wire which establishes the coupling between the two systems. It is assumed that this resistor does not contribute significantly to the entropy of the composite system.

In view of (2-5a) the kinetic equation can be written also as

$$R\dot{\alpha}+s\alpha=0. \tag{2-8a}$$

The form (2-8) or (2-8a) of the kinetic equation involves a number of simplifying assumptions. First, the thermostatic formalism developed for constrained equilibria is assumed to hold even in the absence of constraints. One speaks in this case of *local equilibrium*. This assumption may fail far from equilibrium and in the presence of relaxation phenomena.[12]

The second simplification inherent in (2-8a) is the linearity of the relation. This may fail not only for large values of the current, but also for small ones. An extreme instance for such a behavior is a superconductor in which Ohm's law breaks down as the current falls below a critical value.

Finally, we have assumed also that the forces governing the variations of α are independent of the "velocities" $\dot{\alpha}$, i.e., there are no external magnetic fields or Coriolis forces present, nor is the entropy dependent on $\dot{\alpha}$.

The solutions of (2-8a) are of the form

$$\varphi(t)=Ae^{-\gamma t}, \tag{2-9}$$

with

$$\gamma=s/R, \tag{2-10}$$

and A a constant.

We consider now the function

$$\mathcal{O}=\frac{1}{4R}(R\dot{\alpha}-X)^2=\frac{R}{4}(\dot{\alpha}+\gamma\alpha)^2, \tag{2-11}$$

which is identical within a multiplying factor to that introduced by Onsager and Machlup.[2] We shall refer to it as the OM function.

Obviously $\mathcal{O}\geqq 0$ and $\mathcal{O}_{\min}=\mathcal{O}(\varphi)=0$, i.e., $\mathcal{O}$ vanishes if and only if its argument is a solution of Eq. (2-8a).

The OM function can be written also in the form

$$\mathcal{O}=\tfrac{1}{2}[\phi(\dot{\alpha})+\psi(X)-\dot{S}], \tag{2-12}$$

where

$$\phi\equiv\tfrac{1}{2}R\dot{\alpha}^2, \tag{2-13}$$

$$\psi\equiv\tfrac{1}{2}R^{-1}X^2, \tag{2-14}$$

and

$$\dot{S}=\frac{\partial S}{\partial\alpha}\dot{\alpha}=X\dot{\alpha}. \tag{2-15}$$

The quantities ϕ and ψ are the dissipation functions, and their sum is the rate of total entropy production.

For processes which satisfy the kinetic equation,

$$\phi(\varphi)=\psi(X)=\tfrac{1}{2}\dot{S}. \tag{2-16}$$

On the basis of this relation the quantities 2ϕ and 2ψ are sometimes used interchangably with the entropy production $\dot{S}$. Actually, however, (2-16) is no longer valid if fluctuation paths are taken into consideration.

An extension of the phenomenological theory to include fluctuations will be developed in Sec. 4. Meanwhile, in the next section we summarize the definitions and postulates involving probability functions needed for that purpose.

3. STOCHASTIC PROCESSES

In this section we confine ourselves to a brief formulation of the basic definitions and postulates, appropriate to our case, which will provide the mathematical equipment for dealing with stochastic processes. For an integration of this formalism into a wider context we refer to Onsager and Machlup.[2,3,13]

[11] This statement calls for some qualification. A displacement of electric charge is always associated with a transfer of energy. Hence, we have a coupled-flux problem. This point of view is the basis for the theory of thermoelectric effects. However, if all parts of the composite system are maintained at a constant temperature, the energy transfer need not be made explicit in the theory. This will be justified from the general thermodynamic formalism in a subsequent paper.

[12] It is interesting to note that relaxation effects become amenable to treatment in a limiting case which is in a way complementary to that of local equilibrium. Thus in the case of a spin-lattice system, there may be approximate equilibrium within the spin system and within the phonon system in the face of a local disequilibrium between the two.

[13] Further details concerning the formalism will be found also in Wang and Uhlenbeck's review on the theory of Brownian motion (reference 9). There is a considerable formal analogy between the kinetic fluctuation theory and the theory of Brownian motion paths.

We introduce the notation that

$$W_p\begin{pmatrix}\alpha^{(1)} & \cdots & \alpha^{(p)}\\ t_1 & \cdots & t_p\end{pmatrix}d\alpha^{(1)}\cdots d\alpha^{(p)} \tag{3-1}$$

be the probability that the fluctuation path $\alpha(t)$ satisfies the following specifications: at the times $t_1<t_2\cdots<t_p$

$$\alpha^{(k)}\leq\alpha(t_k)\leq\alpha^{(k)}+d\alpha^{(k)},\quad k=1,\,2,\,\cdots p, \tag{3-2}$$

where $\alpha^{(1)},\,\cdots,\,\alpha^{(p)};\,d\alpha^{(1)},\,\cdots,\,d\alpha^{(p)}$ are given numbers. We say that $\alpha(t)$ is to pass through the "gates" of infinitesmal width specified by (3-2), and we will call the above equations a *finite gate specification*. W_p is called the joint, or p-gate, probability density function (p.d.f.) for the fluctuation path.

We proceed to formulate some basic assumptions, more or less standard in the theory of stochastic processes.

Postulate I: We assume that *the process is Markoffian;* that is, that W_p can be written in the form

$$W_p\begin{pmatrix}\alpha^{(1)} & \cdots & \alpha^{(p)}\\ t_1 & \cdots & t_p\end{pmatrix}=P_2\left(\begin{matrix}\alpha^{(p)}\\ t_p\end{matrix}\middle|\begin{matrix}\alpha^{(p-1)}\\ t_{p-1}\end{matrix}\right)W_{p-1}\begin{pmatrix}\alpha^{(1)} & \cdots & \alpha^{(p-1)}\\ t_1 & \cdots & t_{p-1}\end{pmatrix}, \tag{3-3}$$

where

$$P_2\left(\begin{matrix}\alpha^{(p-1)}\\ t_{p-1}\end{matrix}\middle|\begin{matrix}\alpha^{(p)}\\ t_p\end{matrix}\right)d\alpha^{(p)}$$

is the probability that if $\alpha(t_{p-1})=\alpha^{(p-1)}$ then

$$\alpha^{(p)}\leq\alpha(t_p)\leq\alpha^{(p)}+d\alpha^{(p)}.$$

P_2 is called the conditional p.d.f.

Postulate II: The process is stationary, i.e., W_p and hence, P_2 depend only on the time lags

$$\tau_k=t_k-t_{k-1},\quad k=2,\,\cdots,\,p. \tag{3-4}$$

The above postulate suggests the following simplified notation:

$$W_1(\alpha^{(1)})\equiv W_1\begin{pmatrix}\alpha^{(1)}\\ t_1\end{pmatrix}. \tag{3-5}$$

$$P_2(\alpha^{(2)},\tau_2|\alpha^{(1)})\equiv P_2\left(\begin{matrix}\alpha^{(2)}\\ t_2\end{matrix}\middle|\begin{matrix}\alpha^{(1)}\\ t_1\end{matrix}\right). \tag{3-6}$$

Postulate III: The fluctuations satisfy the principle of *microscopic reversibility:*

$$W_p\begin{pmatrix}\alpha^{(1)} & \cdots & \alpha^{(p)}\\ t_1 & \cdots & t_p\end{pmatrix}=W_p\begin{pmatrix}\alpha^{(p)} & \cdots & \alpha^{(1)}\\ t_1 & \cdots & t_p\end{pmatrix}. \tag{3-7}$$

Since the W's and P's are probability densities, they are normalized, non-negative, and satisfy the consistency relations

$$W_k\begin{pmatrix}\alpha^{(1)} & \cdots & \alpha^{(k)}\\ t_1 & \cdots & t_k\end{pmatrix}=\int\cdots\int W_p\begin{pmatrix}\alpha^{(1)} & \cdots & \alpha^{(p)}\\ t_1 & \cdots & t_p\end{pmatrix}d\alpha^{(k+1)}\cdots d\alpha^{(p)}, \tag{3-8}$$

where $k<p$.

It follows from (3-3) that

$$W_p\begin{pmatrix}\alpha^{(1)} & \cdots & \alpha^{(p)}\\ t_1 & \cdots & t_p\end{pmatrix}=P_2(\alpha^{(p)},\tau_p|\alpha^{(p-1)})\cdots P_2(\alpha^{(2)},\tau_2|\alpha^{(1)})W_1(\alpha^{(1)}). \tag{3-9}$$

Thus the entire process is determined by the two basic probabilities W_1 and P_2. However, these are not independent of each other, but are related according to (3-7) and (3-9) as

$$P_2(\alpha^{(2)},\tau_2|\alpha^{(1)})W_1(\alpha^{(1)})=P_2(\alpha^{(1)},\tau_2|\alpha^{(2)})W_1(\alpha^{(2)}). \tag{3-10}$$

We shall use also the p-gate conditional probability density defined as

$$P_p(\alpha^{(p)},\tau_p;\,\cdots;\alpha^{(2)},\tau_2|\alpha^{(1)})\equiv W_p\begin{pmatrix}\alpha^{(1)} & \cdots & \alpha^{(p)}\\ t_1 & \cdots & t_p\end{pmatrix}\Big/W_1(\alpha^{(1)}). \tag{3-11}$$

Accordingly, we write the Markoff chain relation (3-9) in the form

$$P_p(\alpha^{(p)},\tau_p;\,\cdots\alpha^{(2)},\tau_2|\alpha^{(1)})=P_2(\alpha^{(p)},\tau_p|\alpha^{(p-1)})\cdots P_2(\alpha^{(2)},\tau_2|\alpha^{(1)}). \tag{3-12}$$

Using the terminology of statistical mechanics, we shall also refer to the probability density functions W_p and P_p as "distributions."

The foregoing postulates form a precise basis for further deductions. Before proceeding on this line of thought, however, it is natural to ask the question: Under what conditions can we expect that the distributions defined in terms of the above postulates correspond to the actual behavior of physical systems? We do not wish to discuss this delicate question in any detail at this point, but the following hints will convey an idea of the factors involved.

The strongest assumption is the Markoffian postulate I. It requires in the first place that the system has a short "memory", in the sense that a knowledge of the values of $\alpha(t)$ at the times t_{k-1}, t_k yields information of no greater value in predicting $\alpha(t_{k+1})$ than if we knew $\alpha(t_k)$ alone. This is apparently a good approximation if the time interval t_k-t_{k-1} is long compared to the mean collision time. Systems exhibiting slow relaxation processes and frozen-in molecular disorder are omitted from consideration.

The Markoffian assumption also requires that the number of α variables chosen is adequate for the specification of the statistical properties of the system. This means that the application of the present one-variable theory is severely limited. The rules governing the choice of variables will be considered in another paper dealing with more general systems.

Finally, the fact that the formalism operates in terms of probabilities renders it inadequate for the description of effects in which the quantum mechanical interference of *probability amplitudes* manifests itself in macroscopic phenomena. Hence such phenomena as superconductivity and superfluidity are presumably beyond the scope of the present formalism.[14]

4. DERIVATION OF THE DISTRIBUTION FUNCTION

The previous section contains the formal framework for the description of a general class of random processes. In order to interpret the latter as thermodynamic fluctuations we shall express the basic probability densities W_1 and P_2 in terms of the phenomenological coefficients of Sec. 2.

This junction of the phenomenological and stochastic theories into a unified theory of fluctuation and dissipation calls for the formulation of two postulates. The first of these is well known.[1] It refers to W_1 and we summarize the argument only since it casts light on the subsequent discussion of the p.d.f. P_2.

Postulate IV: The p.d.f. $W_1(\alpha)$ can be expressed as a function of the entropy function:

$$W_1(\alpha)=f(S(\alpha)). \tag{4-1}$$

This postulate is motivated by the intuitively recognized parallel between the one-gate p.d.f. and the entropy function.[15] It is justified by the correctness of its implications.

Once the qualitative relation (4-1) is granted, one concludes in well-known fashion that the unknown function f is of the form

$$W_1(\alpha)\propto\exp[-S(\alpha)/k], \tag{4-2}$$

where k is Boltzmann's constant. On using the expansion (2-6), we obtain the distribution law in the Gaussian approximation. After normalization we have

$$W_1(\alpha)=(s/2\pi k)^{\frac{1}{2}}\exp(-s\alpha^2/2k). \tag{4-3}$$

From here the variance of α is

$$\langle\alpha^2\rangle=ks^{-1}. \tag{4-4}$$

The notation on the left-hand side indicates an ensemble average.

The discussions of this paper will be confined to Gaussian distributions. In particular, we shall use (4-3) rather than the exact distribution (4-2). The scope and limitations of this approximation will be considered in the final discussion.

We proceed now to formulate a postulate which is the kinetic analogue of (4-1) and which expresses a qualitative connection between the conditional p.d.f. P_p and the OM function. In order to render the analogous properties of these functions apparent and, in fact, even to state the postulate in a precise fashion, we need an auxiliary concept to be referred to as the *exponentially smoothed fluctuation path.* This is introduced in a natural fashion in the context of the problem of time reversal.

It is a generally accepted principle[7,16] that fluctuation paths around thermodynamic equilibrium are symmetric with respect to the transformation which reverses the sense of time: $t\rightarrow -t$. We shall denote the corresponding operator by K. This principle finds its expression in the symmetry of the *joint* probability densities W_p (postulate III). On the other hand, assuming that $\alpha^{(1)}>\alpha^{(2)}$, one obtains from (3-10) and (4-3)

$$P_2\left(\begin{matrix}\alpha^{(2)}\\ t_2\end{matrix}\middle|\begin{matrix}\alpha^{(1)}\\ t_1\end{matrix}\right)>P_2\left(\begin{matrix}\alpha^{(1)}\\ t_2\end{matrix}\middle|\begin{matrix}\alpha^{(2)}\\ t_1\end{matrix}\right). \tag{4-5}$$

Hence, the irreversible trend toward equilibrium is expressed in terms of the *conditional* probabilities.

This reconciliation of the reversible aspect of fluctuation and the irreversible trend toward equilibrium within a discrete stochastic model has been known for a long time.[16] However, the situation is less clear within the continuum model of the phenomenological theory. This gap can be filled by utilizing a unique property of the OM function.

We start from the fact that the α variables, considered as averages over molecular parameters, are invariant under time reversal: $K\alpha=\alpha$ (reference 7). The thermodynamic force is defined by (2-5a) as $X\equiv -s\alpha$; hence $KX=X$. In contrast, the fluxes transform according to $K\dot{\alpha}=-\dot{\alpha}$. They are β variables in Casimir's terminology.[17]

We have seen in Sec. 2 that the OM function is minimized by the phenomenological curve $\varphi(t)$ of (2-9). The concept of minimization implies that we may insert into $\mathcal{O}$ virtual paths which do not satisfy the phenomenological equation. We shall construct such virtual paths by means of the time reversal operator. We obtain at once $K\varphi(t)=Ae^{\gamma t}$ and shall refer to $\varphi(t)$ and $K\varphi(t)$ as the forward and reversed paths, respectively.

The dissipation functions ϕ and ψ are positive definite and are invariant under time reversal. In contrast $K\dot{S}=-\dot{S}$; in other words, the reverse paths are associated with a decrease of total entropy. These are "unnatural" processes excluded by the second law

[14] F. London, *Superfluids* (John Wiley and Sons, Inc., New York, 1950), Vol. 1, p. 142.

[15] We note that the entropy function corresponding to the local equilibrium state α is a random function. This is in contrast to the thermodynamic entropy of the equilibrium state.

[16] P. Ehrenfest and T. Ehrenfest, Encykl. Math. Wiss. 4, 2, ii (1911), and Physik. Z. 8, 311 (1907).

[17] H. B. G. Casimir, Revs. Modern Phys. 17, 343 (1945).

of thermodynamics. Therefore we obtain from (2-13) and (2-17)

$$\mathcal{O}(\varphi)=0, \quad \mathcal{O}(K\varphi)=-\dot{S}>0; \tag{4-6}$$

or, expressed in words: *The OM function vanishes for forward paths and is equal to the rate of entropy decrease for the reverse paths.*[3]

This selective property of the OM function with respect to the sense of time remains valid even under more general conditions. Let us consider the virtual path

$$\alpha=Ae^{-\gamma t}+Be^{\gamma t}, \tag{4-7}$$

where A and B are constants, and γ is defined by (2-10). We have

$$\mathcal{O}(\alpha)=\mathcal{O}(Be^{\gamma t})=R\gamma^2B^2e^{2\gamma t}, \tag{4-8}$$

and

$$\int_{t_1}^{t_2}\mathcal{O}(\alpha)dt=\tfrac{1}{2}sB^2(e^{2\gamma t_2}-e^{2\gamma t_1}) =S_B(t_1)-S_B(t_2)=-\Delta S_B, \tag{4-9}$$

where $-\Delta S_B$ is the entropy decrease associated with the reverse component of the path (4-7).

Equations (4-6) to (4-9) show that the OM function has a peculiar "filtering" property. The entropy decrease associated with the reverse component is registered while the forward component is "passed," leaving the corresponding entropy increase unrecorded.

Motivated by the above, we associate with each gate specification (3-2) an *exponentially-smoothed path* $\alpha(t)$, defined by the $(p-1)$ equations

$$\alpha_{\exp}(t)\equiv A_ke^{-\gamma(t-t_{k-1})}+B_ke^{\gamma(t-t_{k-1})} \tag{4-10}$$

in the intervals

$$t_{k-1}\leq t\leq t_k, \quad k=1, \cdots, p-1,$$

where γ is defined by (2-10). The $2p$ constants A_k, B_k are to be determined by the conditions

$$\alpha_{\exp}(t_k)=\alpha^{(k)}, \quad k=1, \cdots, p. \tag{4-11}$$

There is a one-to-one correspondence between the gate specifications and the smoothed paths. One finds

$$\begin{aligned} A_k&=(\rho_k^{-1}-\rho_k)^{-1}(\alpha^{(k-1)}\rho_k^{-1}-\alpha^{(k)}),\\ B_k&=(\rho_k^{-1}-\rho_k)^{-1}(\alpha^{(k)}-\alpha^{(k-1)}\rho_k), \end{aligned} \tag{4-12}$$

where

$$\rho_k=e^{-\gamma\tau_k}, \quad \tau_k=t_k-t_{k-1}. \tag{4-13}$$

The exponentially-smoothed path enables one to take advantage of the remarkable "filtering" property of the OM function expressed by (4-9). We obtain

$$\int_{t_1}^{t_p}\mathcal{O}\{\alpha_{\exp}(t)\}dt=-\sum_{k=2}^{p}\Delta S_B(\tau_k), \tag{4-14}$$

where $S_B(\tau_k)$ is the entropy decrease associated with the reversed component of $\alpha(t)$ in the time interval τ_k.

The expression (4-14) therefore, is a kind of measure for the deviation of the specification (3-2) from phenomenological behavior. This, together with the remark following (4-5), motivates us to formulate the following postulate, in analogy to (4-1).

Postulate V: The conditional p.d.f. P_p of (3-11), (3-12) can be written as a function of the left-hand side of (4-14):

$$P_p(\alpha^{(p)},\tau_p; \cdots; \alpha^{(2)},\tau_2|\alpha^{(1)})=g\left(\int_{t_1}^{t_p}\mathcal{O}\{\alpha_{\exp}(t)dt\}\right). \tag{4-15}$$

Added to the four standard postulates introduced earlier, this postulate is sufficient for the derivation of the entire fluctuation theory. This circumstance provides its justification.

We proceed now to determine the function g. According to (3-12),

$$P_3(\alpha^{(3)},\tau_3;\alpha^{(2)},\tau_2|\alpha^{(1)}) =P_2(\alpha^{(3)},\tau_3|\alpha^{(2)})P_2(\alpha^{(2)},\tau_2|\alpha^{(1)}). \tag{4-16}$$

From (4-15) and (4-16)

$$g\left(\int_{t_1}^{t_2}\mathcal{O}(\alpha_{\exp})dt+\int_{t_2}^{t_3}\mathcal{O}(\alpha_{\exp})dt\right) =g\left(\int_{t_1}^{t_2}\mathcal{O}(\alpha_{\exp})dt\right)g\left(\int_{t_2}^{t_3}\mathcal{O}(\alpha_{\exp})dt\right). \tag{4-17}$$

The only nonsingular solution of this functional equation is

$$P_2(\alpha^{(2)},\tau_2|\alpha^{(1)})\propto\exp\left(a\int_{t_1}^{t_2}\mathcal{O}(\alpha_{\exp})dt\right), \tag{4-18}$$

where a is a constant which remains to be determined.

We write (4-18) in a more explicit form by using (4-10), (4-12), and (4-9).

$$P_2(\alpha^{(2)},\tau_2|\alpha^{(1)})\propto\exp[\tfrac{1}{2}as(\alpha^{(2)}-\alpha^{(1)}\rho_2)^2(1-\rho_2^2)^{-1}]. \tag{4-19}$$

The ratio of time-reversed conditional probabilities is

$$P_2(\alpha^{(2)},\tau_1|\alpha^{(1)})/P_2(\alpha^{(1)},\tau_1|\alpha^{(2)}) =\exp\{\tfrac{1}{2}as[(\alpha^{(2)})^2-(\alpha^{(1)})^2]\}. \tag{4-20}$$

The same ratio is obtained also from (3-10) and (4-3) and the comparison yields $a=-1/k$. Thus we have, finally,

$$P_p(\alpha^{(p)},\tau_p; \cdots; \alpha^{(2)},\tau_2|\alpha^{(1)}) \propto\exp\left[-\frac{1}{k}\int_{t_1}^{t_p}\mathcal{O}(\alpha_{\exp})dt\right], \tag{4-21}$$

and

$$W_p\begin{pmatrix}\alpha^{(1)} & \cdots & \alpha^{(p)}\\ t_1 & \cdots & t_p\end{pmatrix} \propto\exp\left[-\frac{1}{2k}s(\alpha^{(1)})^2-\frac{1}{k}\int_{t_1}^{t_p}\mathcal{O}(\alpha_{\exp})dt\right]. \tag{4-22}$$

Although phrased somewhat differently, Eqs. (4-21) and (4-22) describe essentially the distributions of Onsager and Machlup.[2] The latter expressed their result in a variational language. It is evident, however, from their discussion that the exponentially smoothed paths are the extremal functions of their problem. The concept of smoothed path is an artifice which enables one to formulate the distribution functions in closed form. Further discussion of this point is postponed to Sec. 8.

5. AVERAGES IN THE TEMPORAL REPRESENTATION

We will use the notation that $\bar{y}$ is the time average of the parameter y for a single system taken over an infinite time interval, and $\langle y\rangle$ is the average taken over an ensemble of systems at a particular instant of time. Since we are dealing with equilibrium ensembles, the ergodic property yields

$$\bar{y}=\langle y\rangle. \tag{5-1}$$

The distribution functions obtained in the last section can be used to compute averages of various types. For this purpose the explicit form (4-19) is a convenient starting point.

We write the conditional p.d.f. in the normalized form[18]

$$P_2(\alpha^{(2)},\tau\,|\,\alpha^{(1)})=(s/2\pi k)^{\frac{1}{2}}(1-\rho^2)^{-\frac{1}{2}} \times\exp[-(s/2k)(1-\rho^2)^{-1}(\alpha^{(2)}-\alpha^{(1)}\rho)^2]. \tag{5-2}$$

From P_2 and W_1 the W_p can be built up stepwise. We have

$$W_2\begin{pmatrix}\alpha^{(1)} & \alpha^{(2)}\\ t_1 & t_2\end{pmatrix}=(s/2\pi k)(1-\rho_{12}{}^2)^{-\frac{1}{2}} \times\exp\left[-\frac{s}{2k}\frac{(\alpha^{(1)})^2-2\alpha^{(1)}\alpha^{(2)}\rho_{12}+(\alpha^{(2)})^2}{1-\rho_{12}{}^2}\right], \tag{5-3}$$

where

$$\rho_{ik}=e^{-\gamma|t_i-t_k|}, \tag{5-4}$$

and

$$W_3\begin{pmatrix}\alpha^{(1)} & \alpha^{(2)} & \alpha^{(3)}\\ t_1 & t_2 & t_3\end{pmatrix}=\left(\frac{s}{2\pi k}\right)^{\frac{3}{2}}(1-\rho_{12}{}^2)^{-\frac{1}{2}}(1-\rho_{23}{}^2)^{-\frac{1}{2}} \times\exp\left\{-\frac{s}{2k}\left[\frac{(\alpha^{(1)})^2}{1-\rho_{12}{}^2}+\frac{(\alpha^{(2)})^2(1-\rho_{13}{}^2)}{(1-\rho_{12}{}^2)(1-\rho_{23}{}^2)} +\frac{(\alpha^{(3)})^2}{1-\rho_{23}{}^2}-\frac{2\alpha^{(1)}\alpha^{(2)}\rho_{12}}{1-\rho_{12}{}^2}-\frac{2\alpha^{(2)}\alpha^{(3)}\rho_{23}}{1-\rho_{23}{}^2}\right]\right\}. \tag{5-5}$$

From (5-5) the structure of the distribution function is apparent and it is easy to write down the general expression for W_p. Instead of doing this we prefer to characterize W_p by means of its characteristic function.

The characteristic function of the variables $z_1,\cdots z_p$ with respect to the distribution W_p is defined as the Fourier transform of the joint p.d.f.

$$g(z_1\cdots z_p)=\langle\exp(i\textstyle\sum_k\alpha^{(k)}z_k)\rangle =\displaystyle\int_{\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\exp(i\textstyle\sum_k\alpha^{(k)}z_k)W_p\,d\alpha^{(1)}\cdots d\alpha^{(p)}. \tag{5-6}$$

It is easy to verify that

$$g(z_1\cdots z_p)=\exp\left[-\frac{1}{2}\left(\frac{k}{s}\right)\left(\sum_1^p z_k{}^2+2\sum_{i>k}z_iz_k\rho_{ik}\right)\right]. \tag{5-7}$$

The proof consists essentially in showing that the matrix of the quadratic form in the parenthesis in (5-7) is the inverse of the form in the bracket of Eq. (5-5).

One makes use of the obvious relation

$$\rho_{ik}\rho_{kl}=\rho_{il},\quad t_i<t_k<t_l. \tag{5-8}$$

The use of the generating function has the advantage of allowing the straightforward calculation of averages.

We have

$$\langle\alpha^{(i)}\alpha^{(k)}\rangle=-\left(\frac{\partial^2 g}{\partial z_i\partial z_k}\right)_{z_1=\cdots=z_p=0}. \tag{5-9}$$

Hence

$$\langle\alpha^{(i)}\alpha^{(k)}\rangle=ks^{-1}\rho_{ik}; \tag{5-10}$$

or, using (5-4),

$$\langle\alpha(t)\alpha(t+\tau)\rangle=\langle\alpha(t)\alpha(t-\tau)\rangle=\langle\alpha^2\rangle e^{-\gamma\tau}. \tag{5-11}$$

The equality of the first two terms of Eq. (5-11) is the usual expression of the principle of microscopic reversibility.

The conditional p.d.f. (5-2) immediately yields a description of the regression of fluctuations within an equilibrium ensemble. Focusing attention on those members of the ensemble for which

$$\alpha(t_1)=\alpha^{(1)}, \tag{5-11a}$$

we have

$$\langle\alpha(t_2)\,|\,\alpha^{(1)}(t_1)\rangle=\alpha^{(1)}\rho_{12}, \tag{5-12}$$

where the vertical bar indicates that the averaging is to be performed only over those ensemble members satisfying (5-11a).

Introduce the quantity

$$\Delta\alpha(t_2)=\alpha(t_2)-\alpha^{(1)}\rho_{12} \tag{5-13}$$

to denote the deviation of an actual path from the average regression curve (5-12) at the instant t_2. We find

$$\langle\Delta\alpha(t)^2\rangle=ks^{-1}(1-\rho_{12}{}^2). \tag{5-14}$$

We turn now to the discussion of systems undergoing an irreversible process. We consider an ensemble of

[18] This expression was first obtained by G. E. Uhlenbeck and L. S. Ornstein, Phys. Rev. **36**, 823 (1930). It served as a point of departure of Onsager and Machlup[2] for the derivation of (4-21).

systems in constrained equilibrium, each of whose members has $\alpha=\alpha^{(1)}$. At a given instant of time t_1 a "shutter" is opened in each system, and the entire simultaneously starts its approach to unconstrained equilibrium.

We see that the phenomenological regression given by (2-9) is identical to the average regression (5-12) of fluctuation theory. This statement, frequently advanced as a postulate,[2,7] is implied by our postulate V.

6. FLUCTUATIONS IN THE SPECTRAL REPRESENTATION‡

We turn now to the spectral description of fluctuation paths. Consider a very long time interval θ, and expand the fluctuation path of each member of the ensemble in a Fourier series[19]

$$\alpha(t)=\mathrm{Re}\sum_{n=1}^{\infty}\alpha_n e^{i\omega_n t}. \tag{6-1}$$

Here "Re" means "real part," and

$$\omega_n=(2\pi/\theta)n, \tag{6-2}$$

$$\alpha_n=a_n+ib_n. \tag{6-3}$$

In order to describe the statistical properties of the paths $\alpha(t)$ of (6-1), introduce the notation that

$$f_N(a_1,\cdots,a_N;b_1,\cdots,b_N)da_1,\cdots da_N db_1,\cdots db_N \tag{6-4}$$

be the probability that an observed $\alpha(t)$ satisfies the specifications that the real and imaginary parts of the first N Fourier coefficients of (6-1) lie within the intervals of width da_n, db_n centered around the values appearing in the argument of f_N. (We will call such a specification a *finite Fourier specification.*)

Analogous to the procedure of Sec. 4, we find it convenient to associate with every such path specification a *Fourier-smoothed path* $\tilde{\alpha}(t)$, defined by the equation

$$\tilde{\alpha}(t)=\mathrm{Re}\sum_{n=1}^{N}\alpha_n e^{i\omega_n t}. \tag{6-5}$$

The first N Fourier coefficients are determined by the specification associated with (6-4) and all other α_n's are set equal to zero.

On account of the differences in the associated specifications, it is not obvious that f_N should be very similar to W_p of the temporal description. However, it will be shown in the next section that for the stochastic process of Secs. 4 and 5

$$f_N=C_N\exp\left[-\frac{1}{k}\int_0^{\theta}\mathcal{O}\{\tilde{\alpha}(t)\}dt\right]. \tag{6-6}$$

In the above, C_N is a coefficient of normalization.

In this section we derive a few of the implications of the above formalism. On using (2-11), (6-1) and (6-4) we write explicitly for this probability density

$$f_N=C_N\exp\left[-\frac{\theta}{8k}\sum_{n=1}^{N}\frac{(a_n^2+b_n^2)\omega_n^2}{\sigma(\omega_n)}\right], \tag{6-7}$$

where

$$C_N=\left(\frac{\theta}{8\pi k}\right)^N\prod_{n=1}^{N}\omega_n^2/\sigma(\omega_n), \tag{6-7a}$$

and

$$\sigma(\omega)=R/(R^2+s^2/\omega^2) \tag{6-8}$$

is the real part of the admittance in the language of the electric analogy.

We concern ourselves at first with the autocorrelation coefficient. According to (5-1) we need not distinguish between time and ensemble averages. We have

$$\langle\alpha(t)\alpha(t+\tau)\rangle=\sum_{nm}\langle(\mathrm{Re}\ \alpha_n e^{i\omega_n t})(\mathrm{Re}\ \alpha_m e^{i\omega_m(t+\tau)})\rangle. \tag{6-9}$$

Computing expectation values with the probability densities (6-7), we find

$$\begin{aligned}&\langle a_n^2\rangle=\langle b_n^2\rangle=4k\sigma(\omega_n)/\theta\omega_n^2,\\ &\langle a_n a_m\rangle=\langle b_n b_m\rangle=0,\quad n\neq m,\\ &\langle a_n b_m\rangle=0.\end{aligned} \tag{6-10}$$

The above, therefore, becomes in the limit $\theta\to\infty$,

$$\langle\alpha(t)\alpha(t+\tau)=\frac{2k}{\pi}\int_0^{\infty}\frac{\sigma(\omega)}{\omega^2}\cos\omega\tau d\omega. \tag{6-11}$$

The integration yields

$$\langle\alpha(t)\alpha(t+\tau)\rangle=ks^{-1}e^{-\gamma\tau} \tag{6-12}$$

in agreement with (5-11). On setting $\tau=0$ we obtain a result identical with (4-4), which was derived by averaging in terms of W_1.

The spectral density for α^2 can be obtained by inspection of (6-11) with $\tau=0$ or, alternatively, by a direct calculation, using the p.d.f. (6-7), of the expectation value of the component of α^2 lying within a narrow frequency band. We have

$$G_\alpha(\omega)=(2k/\pi)\sigma(\omega)/\omega^2. \tag{6-13}$$

This is the fluctuation dissipation theorem of Callen and Greene.[8]

Let us introduce the quantity $\epsilon(t)$ defined by

$$R\dot{\alpha}+s\alpha=\epsilon. \tag{6-14}$$

On substituting (6-1) into the above and noting (6-7),

‡ *Note added in proof.*—The connection of the Onsager-Machlup formalism with the method of Rice (spectral representation) has been recently developed also by Hashitsume, Progr. Theoret. Phys. (Japan) **15**, 369 (1956).

[19] There should be a term $\alpha_0=(1/\theta)\int_0^{\theta}\alpha(t)dt$ included in (6-1) which is omitted for simplicity. The end results will remain unaffected since we will eventually allow $\theta\to\infty$.

we find that the coefficients in the Fourier expansion of $\epsilon(t)$ have Gaussian distributions. Furthermore, on making use of (6-13), we obtain the spectral density of ϵ as

$$G_\epsilon = 2kR/\pi. \tag{6-15}$$

Equation (6-14) is the Langevin equation, and the above properties are formal expressions of the Langevin postulates. The fact that these appear as theorems in our formulation is essentially a justification of our postulate V.

In the electric analogy (6-15) is the Nyquist relation[6] of thermal noise for an RC circuit. Changing from the entropy to the energy representation, we have

$$\epsilon = V/T, \quad R = \mathfrak{R}/T, \quad \omega = 2\pi\nu, \tag{6-16}$$

where V is the voltage and $\mathfrak{R}$ the conventional resistance. Hence

$$G_V(\nu) = 4kT\mathfrak{R}. \tag{6-17}$$

7. EQUIVALENCE OF THE TEMPORAL AND SPECTRAL DISTRIBUTION FUNCTIONS

The averages obtained with the probability densities W_p from (3-1), (4-22) and f_N defined by (6-4), (6-6) were found to be consistent with each other. However, it still remains to show that the probability densities W_p actually imply the p.d.f.'s f_N. We first show that the inverse statement is true.

Using our f_N, we shall compute the characteristic function

$$h(z_1,\cdots z_p) = \left\langle \exp\left[i\sum_{k=1}^{p} z_k\alpha^{(k)} \right] \right\rangle, \tag{7-1}$$

and shall prove that the result is identical to the characteristic function $g(z_1,\cdots,z_p)$(5-7) computed by means of W_p. It is well known, however, that the characteristic function uniquely determines the distribution.[20] Therefore the process governed by the distributions f_N implies the process described by the W_p's.

For the sake of simplicity we carry out the calculation for the case of $p=2$. The generalization is entirely straightforward.

Let us set $t_1 = t$ and $t_2 = t+\tau$. We obtain from (6-1)

$$\begin{aligned} \alpha^{(1)} &= \alpha(t) = \mathrm{Re}\sum_{n=1}^{\infty} \alpha_n e^{i\omega_n t}, \\ \alpha^{(2)} &= \alpha(t+\tau) = \mathrm{Re}\sum_{n=1}^{\infty} \alpha_n e^{i\omega_n(t+\tau)}. \end{aligned} \tag{7-2}$$

We first compute

$$h_N(z_1z_2) = \langle \exp[iz_1\alpha^{(1)} + iz_2\alpha^{(2)}] \rangle_N, \tag{7-3}$$

where the subscript N indicates that the averaging takes place only over the first N Fourier coefficients. On using for f_N the explicit form (6-7), we obtain

$$h_N = C_N \prod_{n=1}^{N} A_n B_n \exp\left[iz_1 \,\mathrm{Re}\sum_{N+1}^{\infty} \alpha_n e^{i\omega_n t} + iz_2 \,\mathrm{Re}\sum_{N+1}^{\infty} \alpha_n e^{i\omega_n(t+\tau)} \right], \tag{7-4}$$

where

$$\begin{aligned} A_n &= \int_{-\infty}^{+\infty} \exp\left\{ -\frac{\theta a_n^2\omega_n^2}{8k\sigma(\omega_n)} + ia_n[z_1\sin\omega_n t + z_2\sin\omega_n(t+\tau)] \right\} da_n, \\ B_n &= \int_{-\infty}^{+\infty} \exp\left\{ -\frac{\theta b_n^2\omega_n^2}{8k\sigma(\omega_n)} + ib_n[z_1\cos\omega_n t + z_2\cos\omega_n(t+\tau)] \right\} db_n, \end{aligned} \tag{7-5}$$

and C_N is given by (6-7a). We have

$$C_N \prod_{1}^{N} A_n B_n = \exp\left\{ \frac{-2k}{\theta} \sum_{n=1}^{N} [z_1^2 + z_2^2 + 2z_1z_2\cos\omega_n t] \frac{\sigma(\omega_n)}{\omega_n^2} \right\}. \tag{7-6}$$

We go now to the limit $\theta\to\infty$ and $N/\theta \approx \omega_N \to \infty$. Under this limiting process the residual factor in (7-4) can be replaced by unity,[21] and the negative of the exponent in (7-6) becomes

$$\begin{aligned} &\lim \frac{2k}{R\theta} \sum_{n=1}^{N} (z_1^2 + z_2^2 + 2z_1z_2\cos\omega_n t)(\omega_n^2 + \gamma^2)^{-1} \\ &= \frac{k}{\pi R}\int_0^{\infty} (z_1^2 + z_2^2 + 2z_1z_2\cos\omega t)(\omega^2+\gamma^2)^{-1} d\omega \\ &= \tfrac{1}{2}ks^{-1}(z_1^2 + z_2^2 + 2z_1z_2e^{-\gamma\tau}). \end{aligned} \tag{7-7}$$

Hence

$$h = \lim_{N\to\infty} h_N = \exp[-\tfrac{1}{2}ks^{-1}(z_1^2 + z_2^2 + 2z_1z_2e^{-\gamma\tau})]. \tag{7-8}$$

This is indeed identical to (5-7) for $p=2$. Since the extension to arbitrary p offers no difficulty, we have shown that the f_N imply the W_p.

That the inverse statement is also true follows from

[20] See, e.g., B. V. Gnedenko and A. N. Kolmogorov, *Limit Distributions for Sums of Independent Random Variables*, translation by K. L. Chung (Addison-Wesley Publishing Company, Inc., Cambridge, 1954), p. 50.

[21] Choose any $\epsilon>0$ and consider the collection of all paths which make this factor differ from unity by an amount greater than ϵ. It can be shown that the probability measure of this collection tends to zero as $N\to\infty$. This follows directly from theorem 4.1 of J. L. Doob [*Stochastic Processes* (John Wiley and Sons, New York, 1953), p. 155] (make the substitution $y_n = \alpha_k e^{i\omega_k t}$), and from a property of convergence in the mean [Doob, p. 9, Eq. (4.6)].

a theorem due to Kolmogoroff.[22] It is a direct consequence of this theorem that any complete set of finite-dimensional p.d.f.'s (such as our f_N's) *uniquely* determines any other set of finite-dimensional p.d.f.'s for the same stochastic process (in our case, the W_p's).

8. DISCUSSION

a. Variational Principles and Smoothed Paths

The distributions in function space discussed in this paper can be considered as generalizations of the situation encountered in variational problems. Given an appropriate set of initial and boundary conditions a variational principle determines a single extremal function, i.e., a δ-function distribution in function space.

While variational problems are equivalent to their deterministic Euler-Lagrange equations, the distribution problem is equivalent to a stochastic differential equation.

It is interesting to consider variational principles which can be formulated in terms of the OM function. A variational principle which implies the phenomenological equation has been mentioned in Sec. 2b. We shall consider now another class of variational problems defined as follows: let us select the path $\alpha(t)$ which minimizes the time integral of the OM function

$$\int_{t_1}^{t_p} \mathcal{O}\{\alpha(t)\}dt = \min. \tag{8-1}$$

Depending on the constraints imposed on the family of trial functions, (8-1) generates a whole family of variational principles.

For example, impose the constraints that every $\alpha(t)$ in the family of trial functions satisfies the conditions

$$\alpha(t_k)=\alpha^{(k)}, \quad k=1, \cdots p, \tag{8-2}$$

where $t_1<\cdot<t_p$ are p instants of time and the $\alpha^{(k)}$ are given numbers. It is easily shown that the solution of this variational problem is the exponentially-smoothed path $\alpha_{\exp}(t)$ of (4-10).[23]

On the other hand, one can choose as trial functions all functions of the form (6-1) subject to the constraints that the first N Fourier coefficients have specified values. In this case one finds that the solution of the variational problem is the Fourier-smoothed path $\tilde{\alpha}(t)$ of (6-5).

We therefore have the following property common to both the temporal and Fourier descriptions. To every (gate, Fourier) specification there corresponds a set of constraints for (8-1). The solution of this variational problem is the (exponentially, Fourier) smoothed path. The probability density associated with the (gate, Fourier) specification is a function of the quantity obtained by inserting the solution of the variational problem back into the functional which is varied:

$$\left.\begin{aligned} P_p &= P_p\left\{\int_{t_1}^{t_p} \mathcal{O}(\alpha_{\exp}(t))dt\right\} \\ f_N &= f_N\left\{\int_0^{\theta} \mathcal{O}(\tilde{\alpha}(t))dt\right\} \end{aligned}\right\}. \tag{8-3}$$

In the foregoing discussion the exponentially smoothed fluctuation path has only a formal meaning; it serves to formulate the probability of a particular gate specification.

In addition, however, this concept can be given a more physical interpretation: the true fluctuation paths are clustered around the exponentially smoothed path. This statement can be cast into a precise form, as was pointed out to us by Siegel,[24] by computing the following problem:

Let $t_1<t_2<t_3$ be three instances of time and consider the specification $\alpha(t_1)=\alpha^{(1)}$, $\alpha(t_3)=\alpha^{(3)}$.

One can show by straightforward calculation that $\alpha(t_2)$ obeys a Gaussian distribution centered around $\alpha_{\exp}(t_2)$.

b. Conclusion

The basic problem of the time dependent theory of fluctuations is to combine the phenomenological kinetic theory and the stochastic theory of random processes into a consistent scheme. The main contribution of this paper is to achieve this junction in a novel fashion.

The conventional approach is characterized by certain assumptions which we call the Langevin postulates. The procedure is to augment the kinetic equation by a term corresponding to a random force specified in terms of statistical assumptions: the force is assumed to be a Gaussian random function with a white spectrum, the density of which is fixed by Boltzmann's principle [see Eq. (6-14)]. The existence of fluctuations makes it evident that the phenomenological kinetic equation is not exact, and the Langevin method "saves" the validity of the equation by adding to it a random perturbation. This is the approach taken, for example, by Onsager and Machlup in their derivation of the density functions W_p.

The point of view taken in this paper calls for a shift in emphasis from the differential equation to the set of functions which describe the fluctuation paths. Concrete physical situations are accounted for in terms of distribution functions over this set of paths, or in other words through p.d.f. in function space. These distributions are closely related to the kinetic differential equation; the solutions of the latter constitute the most probable region in function space. The connection

[22] A. N. Kolmogoroff, *Foundations of the Theory of Probability* (Chelsea Publishing Company, New York, 1950), Chap. III, Sec. 4. This theorem is also described in footnote 7 of reference 2.

[23] This result and its connection with p.d.f. W_p of (4-22) was discovered by Onsager and Machlup.[2]

[24] Armand Siegel (private communication).

between the kinetic equation and the p.d.f. in function space is stipulated by postulate V [Sec. 4].

Heuristically, this postulate was suggested by the Onsager-Machlup distribution function derived snythetically from the Langevin postulates. In the present logical structure it is sufficient to join postulate V to the standard assumption of the kinetic and the stochastic theories to develop the kinetic theory of fluctuations. In particular, we obtain both the OM distribution and the Langevin postulates as theorems. So far as results are concerned this presentation is found to be equivalent to the one conventionally followed. However, we believe that the present method has some appealing conceptual and formal features.

In the first place our formalism gives justice both to the reversible and to the irreversible aspects of time. The manifold of fluctuation paths is invariant with respect to time reversal[25] while the bias in favor of the forward direction is introduced by means of OM function, which exhibits a peculiar selectivity with respect to the sense of time.

Although the Langevin theory is consistent with the microscopic reversibility of fluctuation paths, it fails to provide any formal counterpart to this principle. Thus, the regression of fluctuations is described by the phenomenological equation and the reverse process, the building up of fluctuations is attributed to the random force.

A second point we may bring up in favor of the present approach is its economy. This is apparent in the postulational basis, where postulate V replaces the Langevin postulates and a number of others usually advanced without proofs. The economy appears also in the formal structure built on the postulational basis. We refer, for example, to the not inconsiderable simplification in the derivation of the Callen-Greene fluctuation dissipation theorem.[8]

Throughout this paper we have emphasized the parallel aspects of the time-dependent distributions with the statistical thermodynamics of equilibrium. It is in order to point out, therefore, that in its present stage of development the kinetic fluctuation theory is subject to considerable limitations. We shall indicate the crucial point in connection with the characteristic function defined by Eq. (5-6).

This relation is seen to be the analogue of the definition of the canonical ensemble in statistical mechanics.[26]

It can indeed be shown that this analogy has its merits although it is limited by the fact that our distributions are Gaussian. This approximation leads to the correct second fluctuation moments, but does not have the formally exact character of the canonical distribution.[27] In order to achieve such a generality it would be necessary to go beyond the Gaussian approximation.

While the solution of the last-mentioned problem would require essentially new ideas, there are a number of other generalizations which can be carried out with minor extensions of the present framework.

It is possible to consider systems with many variables, exhibiting inertial effects, spatial expansion and those which are subject to external forces. We hope to return to these questions later.

The authors are indebted to Dr. D. F. Falkoff, Dr. Bayard Rankin, and Dr. Armand Siegel for commenting on the manuscript.

[25] This means that time reversal transforms fluctuation paths into admissible fluctuation paths. The paths which are individually invariant under time reversal form a special subset.

[26] A. I. Khinchin, *Statistical Mechanics* (Dover Publications, New York, 1949), p. 76.

[27] R. F. Green and H. R. Callen, Phys. Rev. **83**, 1231 (1951).