Fluctuations and Irreversible Processes*

L. ONSAGER, *Sterling Chemistry Laboratory, Yale University, New Haven, Connecticut*

AND

S. MACHLUP, *Sloane Physics Laboratory, Yale University, New Haven, Connecticut,*

and

Bell Telephone Laboratories, Murray Hill, New Jersey

(Received May 13, 1953)

The probability of a given succession of (nonequilibrium) states of a spontaneously fluctuating thermodynamic system is calculated, on the assumption that the macroscopic variables defining a state are Gaussian random variables whose average behavior is given by the laws governing irreversible processes.

This probability can be expressed in terms of the dissipation function; the resulting relation, which is an extension of Boltzmann's principle, shows the statistical significance of the dissipation function. From the form of the relation, the principle of least dissipation of energy becomes evident by inspection.

1. INTRODUCTION

THE connection between the theory of irreversible processes and the spontaneous fluctuations of thermodynamic variables of equilibrium systems was discussed in two earlier papers by one of us.[1] There, the fluctuation theory was brought in to prove a theorem for irreversible processes, the reciprocal relations: the symmetry of the matrix of coefficients of the set of linear equations relating thermodynamic "forces" and "fluxes." The connection was made by postulating that the decay of a system from a given nonequilibrium state produced by a spontaneous fluctuation obeys, *on the average*, the (empirical) law for the decay from the same state back to equilibrium, when it has been produced by a constraint which is then suddenly removed.

The present paper will establish a further connection between the two fields, along the lines suggested in Sec. 5 of RRIP II. We add the following to the assumptions made there: that the (fluctuating) thermodynamic variables be *Gaussian* random variables. We are then in a position to calculate the probability that a system, in undergoing spontaneous thermal fluctuation, passes through a given set of thermodynamic states at given instants. Two kinds of coefficients will enter this probability expression: those (s_{ij}) which give the entropy of a given (nonequilibrium) state as a function of the chosen set of thermodynamic variables, and those (R_{ij}) which give the rates of the irreversible processes, i.e., the rates of change of the thermodynamic variables of state when the state itself is known.

The argument proceeds somewhat as follows: if we are interested in only a single instant, the probability of a given state Γ is given by Boltzmann's principle,

$$k \log \text{Prob}(\Gamma) = S(\Gamma) + \text{const}, \qquad (1\text{-}1)$$

which relates the probability of finding the system in the state Γ to the entropy of Γ. If we are interested in two instants widely separated in time, the probability of given states at each instant is equal to the product of the individual probabilities. A long time lapse makes the states statistically independent.[2] Hence the joint probability of the succession is related to the sum of the two entropies. But if the time lapse is not long, the states will be statistically correlated. It is precisely the laws for irreversible behavior which tell us the correlations, by virtue of the postulate connecting these laws with the average course of fluctuations.

These correlations, together with the means (equilibrium values) of the fluctuating variables, are sufficient to determine the joint probability for two instants, by virtue only of the assumption that the fluctuations are Gaussian processes. If we have the correlations for all pairs of instants, the joint probability for any number of instants is determined.

If the thermodynamic variables are properly chosen (see Sec. 2), they also turn out to be Markoffian random variables. This means that the statistics of the future of the system do not depend on its past but only on its present state. The "proper choice" of thermodynamic variables actually implies limitations on the type of systems for which our treatment is valid. The proof of the reciprocal relations (RRIP I and II) was based on the hypothesis of microscopic reversibility, which we retain here. This excludes rotating systems (Coriolis forces) and systems with external magnetic fields. The assumption of Gaussian random variables is also restrictive: Our system must consist of many "sufficiently" independent particles, and equilibrium must be stable at least for times of the order of laboratory measuring times. This is also required in order that

---

[1] L. Onsager, "Reciprocal Relations in Irreversible Processes," Phys. Rev. **37**, 405 (1931); **38**, 2265 (1931). Henceforth referred to as RRIP I and II.

[2] This statement is, of course, charged with meaning, and requires elaborate precautions about ergodicity, etc. It may be said to hold for system which "forget" their initial states in a "reasonably" short time. It is, however, precisely the choice of time scale that matters. In a sufficiently long time, all physical systems "forget."

we can consider our systems as "aged"—left alone long enough to have "forgotten" their initial states.

The mathematics, then, will involve the treatment of stationary Gaussian Markoff processes in a finite number of variables, which is by now standard.[3] What we believe to be original, in this domain, is the introduction of variational expressions of the type of Eq. (4-7) into the description of statistical distributions. The physics—the theory of thermodynamic fluctuations, and the theory of irreversible processes—involves the concepts dealt with in RRIP (see especially II, pp. 2265–2274, also the review article by Casimir[4]), using the methodological foundation for fluctuation theory given by Einstein.[5]

In RRIP, a *dissipation function* was defined, a quadratic form in the "fluxes" which gives the rate of entropy production for irreversible processes. The calculation of the probability for a preassigned succession of thermodynamic states yields a statistical interpretation of the dissipation function.[6]

As is well known, distribution in function space can not be described by a density, but its projection upon any finite number of dimensions can be so described.[7] Accordingly, it makes sense to ask for the probability that a function—or a finite set of functions—assume values within specified limits at an arbitrarily selected finite set of successive instants. We shall arrive at a description of the probability distribution for "fluctuation paths" in terms of an *auxiliary functional,* such that the maximum of this functional, for a given finite set of specifications, determines the appropriate probability density in a space, whose dimensions correspond to the separate specifications of fluctuating functions at each selected instant.

2. IRREVERSIBLE PROCESSES

The irreversible processes we consider are of the type treated in RRIP II: flows of matter, heat, and electricity whose rates are linearly related to the corresponding *thermodynamic forces*—concentration gradients, temperature gradients emf's. The extension to systems with chemical reactions and local relaxation effects usually involves simply a change of language. We are thinking

of closed systems, systems displaced from equilibrium and then "released" and allowed to return. But again, the extension to open systems (and steady states) is purely formal.

The thermodynamic *state* of a system will be defined by a set of extensive variables $\alpha_1, \alpha_2 \cdots \alpha_n$, quantities like volume, energy, mass of substance, electric charge. We choose extensive variables because of the conceptual difficulties encountered in speaking of fluctuations of intensive quantities. But defining thermodynamic variables at all for a system not in equilibrium requires some sophistication. The standard device is to assume that the system is built up of small sub-systems, each considered in "local equilibrium." This fails far from equilibrium and in systems undergoing such slow relaxations that the concept of equilibrium itself is hard to define.[8] It should be recalled that the thermodynamic functions (e.g., entropy) can be empirically measured only for states—or sets of sub-system states—which are accessible by reversible manipulations.

One may formulate rules for choosing α variables for which the theory here developed will be applicable:

(1) They must be macroscopic variables, i.e., thermodynamically measurable, and must refer to sub-systems containing a large number of molecules. This is always necessary for thermodynamics.

(2) On a kinetic model, they must be algebraic sums of molecular variables. Their fluctuations will then have Gaussian distributions about the equilibrium values, by virtue of a kind of Central Limit Theorem. The rule also eliminates pathological variables (e.g., nonadditive functions of extensive variables).

(3) They must be even functions of those molecular variables which are odd functions of the time, like molecular velocities. This means that a reversal of the sense of time will not change their values and is necessary for the assumption of microscopic reversibility.

We shall assume that the thermodynamic functions are completely determined by specification of the α's. Rule (3) implies that the kinetic energy of the flows is negligible, since this depends on velocity-type variables. The case in which inertia (and hence kinetic energy) is important—in which time derivatives of the α's enter into the thermodynamic functions—will be treated in a subsequent paper.[9]

Thus the *entropy* is a function of the α's:

$$S = S(\alpha_1 \cdots \alpha_n) = S(\alpha).$$

Its maximum (equilibrium) value is denoted by S_0, and the α's will be redefined to vanish for the equilibrium state:

$$S_0 = S(0, \cdots 0).$$

The tendency of the system to seek equilibrium is measured by the *thermodynamic forces,*

$$X_i = \partial S / \partial \alpha_i, \tag{2-1}$$

[3] See, for instance, M. C. Wang and G. E. Uhlenbeck, Revs. Modern Phys. **17**, 323 (1945) or J. E. Moyal, J. Roy. Statist. Soc. **B11**, 150 (1949), for applications in physics. The recent appearance of J. L. Doob, *Stochastic Processes* (J. Wiley and Sons, New York, 1953) makes unnecessary further reference to the mathematical literature.

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

[5] A. Einstein, Ann Physik **33**, 1275 (1910), especially Sec. I, "General Matters Relating to Boltzmann's Principle."

[6] In RRIP II, Sec. 5, a statistical interpretation is suggested in terms of the joint probability for a pair of instants closely spaced in time. An extension of this type of differential expression, using Markoff methods (Fokker-Planck equation), has recently been made by N. Hashitsume, Progr. Theoret. Phys. **8**, 461 (1952).

[7] A. N. Kolmogoroff, *Foundations of the Theory of Probability* (Chelsea Publishing Company, New York, 1950), Chap III, Sec. 4, proves that the set of all finite-dimensional distribution functions uniquely determines the probability functions for all Borel sets in the infinite-dimensional space.

[8] P. W. Bridgman, Revs. Modern Phys. **22**, 56 (1950).

[9] S. Machlup and L. Onsager, following paper [Phys. Rev. **91**, 1512 (1953)].

"restoring forces," which evidently vanish with the α's.[10]

The fluxes (of matter, heat, electricity) are measured by the time derivatives of the α's. The essential physical assumption about the irreversible processes is that they are linear; i.e., that the fluxes depend linearly on the forces that "cause" them:

$$\Sigma_j R_{ij}\dot\alpha_j = X_i, \quad i=1\cdots n; \quad (2\text{-}2)$$

or

$$\Sigma_j L_{ij} X_j = \dot\alpha_i, \quad (2\text{-}3)$$

where the matrices **L** and **R** are mutually reciprocal. These equations express, for instance, Ohm's law for electrical conduction, Fourier's law for heat conduction, Fick's law for diffusion, and the extension of these laws to interacting flows, e.g., anisotropic conduction (heat, electricity), thermoelectric effects, thermal diffusion. For systems for which microscopic reversibility holds (to which this work is confined), we have the *reciprocal relations*

$$R_{ij} = R_{ji}[\mathbf{R} = \mathbf{R}_{tr}], \quad (2\text{-}4)$$

where the subscript *tr* means transpose.

The rate of production of entropy is

$$\dot S = \Sigma_j(\partial S/\partial\alpha_j)\dot\alpha_j = \Sigma_j X_j\dot\alpha_j, \text{ by Eq. (2-1)} \quad (2.5)$$

$$= \Sigma_{ij}R_{ij}\dot\alpha_i\dot\alpha_j, \text{ by Eq. (2-2)} \quad (2\text{-}6)$$

$$= \Sigma_{ij}L_{ij}X_iX_j, \text{ by Eq. (2-3).} \quad (2\text{-}7)$$

The quadratic form in the fluxes,

$$\Phi(d\alpha/dt, d\alpha/dt) = \tfrac{1}{2}\Sigma_{ij}R_{ij}\dot\alpha_i\dot\alpha_j, \quad (2\text{-}8)$$

the dissipation function, acts as a potential for the thermodynamic forces. Its specification is therefore equivalent to knowledge of the phenomenological equations. The corresponding function of the forces,

$$\Psi(\mathbf{X}, \mathbf{X}) = \tfrac{1}{2}\Sigma_{ij}L_{ij}X_iX_j, \quad (2\text{-}9)$$

has a similar property, but it should be noticed that it is a function of the *state*, whereas the numerically equal $\Phi(d\alpha/dt, d\alpha/dt)$ is a function of its rate of change.

If we expand the entropy in a Taylor series about equilibrium, we have

$$S = S_0 - \tfrac{1}{2}\Sigma_{ij}s_{ij}\alpha_i\alpha_j + \text{higher terms.} \quad (2\text{-}10)$$

Neglect of the higher terms means the assumption that fluctuations are Gaussian: for Boltzmann's principle states that the logarithm of the probability of a given fluctuation is proportional to its entropy, or

$$\text{Prob.}\{\alpha\} \propto \exp(S/k) \propto \exp\left[-\left(\frac{\tfrac{1}{2}\Sigma_{ij}s_{ij}\alpha_i\alpha_j}{k}\right)\right]. \quad (2\text{-}11)$$

The thermodynamic forces then become linear restoring

forces,

$$X_i = -\Sigma_j s_{ij}\alpha_j, \quad (2\text{-}12)$$

and the phenomenological laws (2-2) become

$$\Sigma_j(R_{ij}\dot\alpha_j + s_{ij}\alpha_j) = 0. \quad (2\text{-}13)$$

3. FLUCTUATIONS

Stochastic Processes

The deterministic Eqs. (2-2) are modified to include fluctuations by the addition of a *random force* term,

$$\Sigma_j R_{ij}\dot\alpha_j = X_i + \epsilon_i, \quad (3\text{-}1)$$

and thereby become stochastic equations. For the present, the only specification on the statistics of the ϵ's is that they have zero means. Another way of looking at the equations is to call the rhs of (3-1) random forces with means X_i.

We shall be concerned with the path of the α's in time under the "influence" of these random forces. Our aim is to calculate the probability of "any path." We must, therefore, define a probability measure for paths. For simplicity we consider first a single variable α, and ask for the probability of a path $\alpha(t)$ over some finite interval of t. To obtain a nonzero probability, we have to phrase our question: Given a *set* of functions over the given interval, what is the probability that the spontaneously fluctuating variable α is given by a function of this set during the interval? There is a difficulty in specifying the functions: each function is specified by an infinite number of coordinates. If we confine ourselves to well-behaved functions and a finite interval, a denumerable infinity will suffice, e.g., the value of the function at all rational points of t, or all the Fourier coefficients. But the probability of an event which is specified by an infinite number of specifications is always zero, unless the set of specifications is "asymptotically trivial." So we must confine ourselves to specifying a finite number of coordinates.

As far as observation is concerned, this is quite sufficient. For although we may obtain a "continuous record" from a recording instrument, corresponding to an infinite number of observations, every instrument essentially takes a time average over a finite interval, which may be short compared to the relaxation times of the irreversible process, but is always long compared to the mean collision times of the molecules. The essential point is that we must be free to choose which coordinates to specify, i.e., at what instant to perform an observation.

We shall use the following notation: for p instants $t_1 < t_2 \cdots < t_p$ we write the cumulative distribution function (c.d.f.)

$$F_p\begin{pmatrix} \alpha^{(1)} & \alpha^{(2)} & & \alpha^{(p)} \\ t_1 & t_2 & \cdots & t_p \end{pmatrix} = \text{Prob.}\{\alpha(t_k) \leq \alpha^{(k)}, k = 1\cdots p\}.$$

$$(3\text{-}2)$$

---

[10] Why this particular definition? It gives Eq. (2-5) the desired form: the rate of dissipation is a sum of products of forces and fluxes. Note that the X's are functions of state, i.e., of the α's.

Such functions must be additive set functions satisfying the usual consistency relations for c.d.f.'s.[11]

We define a stationary process as one whose c.d.f. is invariant against an arbitrary shift of the time axis:

$$F_p\begin{pmatrix}\alpha^{(1)} & \cdots & \alpha^{(p)} \\ t_1 & & t_p\end{pmatrix}=F_p\begin{pmatrix}\alpha^{(1)} & & \alpha^{(p)} \\ t_1+\tau & & t_p+\tau\end{pmatrix}, \text{ for any } \tau. \quad (3\text{-}3)$$

Physically, this describes an *aged* system, a system that has been left alone long enough that any initial conditions have "worn off," or "been forgotten." In dealing with fluctuations, we shall always make this assumption of stationarity. This amounts to a strong form of the statistical interpretation of equilibrium. Thus we consider entropy creation as loss of information: a dissipative system forgets its past.

Pictorially, a function F_p tells the probability that the path $\alpha(t)$ lie below the "barriers" $\alpha^{(1)}\cdots\alpha^{(p)}$ at the corresponding instants $t_1\cdots t_p$. Alternatively, one will ask for the probability that the path pass through narrow "gates" of width $\Delta\alpha^{(k)}$:

$$\sum_{\mu_1\cdots\mu_p=0}^{1}(-)^q F_p\begin{pmatrix}\alpha^{(1)}+\mu_1\Delta\alpha^{(1)}\cdots\alpha^{(p)}+\mu_p\Delta\alpha^{(p)} \\ t_1 & \cdots & t^p\end{pmatrix}$$

$$\equiv f_p\begin{pmatrix}\alpha^{(1)}\cdots\alpha^{(p)} \\ t_1 \cdots t_p\end{pmatrix}\Delta\alpha^{(1)}\cdots\Delta\alpha^{(p)}, \quad (3\text{-}4)$$

where

$$q=p-\sum_{k=1}^{p}\mu_k.$$

The function f_p is a probability density function (p.d.f.) and will henceforth be used with the tacit assumption that it exists.

Any finite number of gates may be erected, specifying the path to an arbitrary degree of precision in any finite interval.

The conditional probability function for the $(p+1)$th event given the previous p,

$$F_1\begin{pmatrix}\alpha^{(p+1)} & \Big| & \alpha^{(1)}\cdots\alpha^{(p)} \\ t_{p+1} & \Big| & t_1 \cdots t_p\end{pmatrix}$$

$$=\text{Prob}\{\alpha(t_{p+1})=\alpha^{(p+1)}\,|\,\alpha(t_k)=\alpha^{(k)},\ k=1\cdots p\}, \quad (3\text{-}5)$$

is defined by the relation:

$$F_{p+1}\begin{pmatrix}\alpha^{(1)}\cdots\alpha^{(p+1)} \\ t_1 \cdots t_{p+1}\end{pmatrix}$$

$$=\int_{-\infty}^{\alpha^{(1)}}(p\text{-fold})\int_{-\infty}^{\alpha^{(p)}}F_1\begin{pmatrix}\alpha^{(p+1)} & \Big| & x^{(1)}\cdots x^{(p)} \\ t_{p+1} & \Big| & t_1 \cdots t_p\end{pmatrix}$$

$$\times dF_p\begin{pmatrix}x^{(1)}\cdots x^{(p)} \\ t_1 \cdots t_p\end{pmatrix}. \quad (3\text{-}6)$$

We define a *Markoff process* as one whose conditional probabilities are independent of all but the immediately preceding instant:

$$F_1\begin{pmatrix}\alpha^{(p+1)} & \Big| & \alpha^{(1)} & \cdots & \alpha^{(p)} \\ t_{p+1} & \Big| & t_1 & & t_p\end{pmatrix}=F_1\begin{pmatrix}\alpha^{(p+1)} & \Big| & \alpha^{(p)} \\ t_{p+1} & \Big| & t_p\end{pmatrix}. \quad (3\text{-}7)$$

Intuitively: the system has a short memory.

For a Markoff process, Eqs. (3-6) and (3-7) give

$$f_p\begin{pmatrix}\alpha^{(1)} & \cdots & \alpha^{(p)} \\ t_1 & & t_p\end{pmatrix}=f_1\begin{pmatrix}\alpha^{(p)} & \Big| & \alpha^{(p-1)} \\ t_p & \Big| & t_{p-1}\end{pmatrix}$$

$$\times f_1\begin{pmatrix}\alpha^{(p-1)} & \Big| & \alpha^{(p-2)} \\ t_{p-1} & \Big| & t_{p-2}\end{pmatrix}\cdots f_1\begin{pmatrix}\alpha^{(2)} & \Big| & \alpha^{(1)} \\ t_2 & \Big| & t_1\end{pmatrix}f_1\begin{pmatrix}\alpha^{(1)} \\ t_1\end{pmatrix}. \quad (3\text{-}8)$$

For the physical systems considered, the last factor in this product, the one-gate p.d.f., is known from the entropy function by Boltzmann's principle, Eq. (2-11). Thus, in view of the stationarity, all that is needed for the distribution function for an arbitrary number of gates is to evaluate the conditional p.d.f.,

$$f_1\begin{pmatrix}\alpha^{(2)} & \Big| & \alpha^{(1)} \\ t+\tau & \Big| & t\end{pmatrix}, \quad (3\text{-}9)$$

which is independent of t: The p-gate problem reduces to the two-gate problem.

All of the above can be extended without modification to the case of several variables.

We shall want to know what form Eqs. (3-1) take for a Markoff process. The result is that the random force terms ϵ_i are purely random "functions"; i.e., values of the ϵ_i at different times are wholly uncorrelated.

$$\langle\epsilon(t)\epsilon(t+\tau)\rangle=0, \quad \tau\neq0.$$

Such quantities are, to be sure, mathematical freaks: nowhere defined, not even square-integrable, but the seeming difficulties arise only from our inappropriate choice of differential notation, rather than a more rigorous one using finite differences. Only the latter, of course, are accessible to physical observation, since successive measurements cannot be arbitrarily close in time.

A *Gaussian* stochastic process is one whose p.d.f. is a multivariate Gaussian distribution. We shall make the assumption that fluctuations are Gaussian processes. The physical basis for this lies in Rule 2 (Sec. 2) for choosing the α variables: if molecular variables are "sufficiently" independent (weakly coupled), algebraic sums of them will behave as Gaussian random variables. Mathematically, on the other hand, we know that a Gaussian process will always be a good approximation, in the following sense: There is a theorem (Khinchin-Cramèr theorem[12]) which states that for every stochastic (L^2) process, there exists a Gaussian process

[11] See, for example, Harald Cramèr, *Mathematical Methods of Statistics* (Princeton University Press, Princeton, 1946).

[12] A. Khinchin, Math. Ann. **109**, 604 (1933).

which is unitarily equivalent to it, meaning that all first and second moments are equal for the two processes.

The Regression of Fluctuations

The postulate on which the proof of the reciprocal relations theorem (RRIP II) was based may be stated: the average regression of fluctuations from a given nonequilibrium state will obey the same laws as the corresponding irreversible processes.

By the average regression of, say α_i, from the state $\boldsymbol{\alpha}'$ is meant the conditional mean,

$$\langle \alpha_i, t+\tau \,|\, \boldsymbol{\alpha}', t \rangle_{\text{Av}} = \int \cdots \int \alpha_i f_1 \binom{\boldsymbol{\alpha} \ \Big| \ \boldsymbol{\alpha}'}{t+\tau \ \Big| \ t} d\boldsymbol{\alpha}, \quad (3\text{-}10)$$

of α_i at time $t+\tau$ *given* that the state $\boldsymbol{\alpha}$ (the values of all the α's) at time t was $\boldsymbol{\alpha}'$. The postulate asserts that this quantity obeys the phenomenological laws (2-2.)

Intuitively, this states that, as far as the average behavior is concerned, it does not matter whether a state was the result of a spontaneous fluctuation or of an imposed constraint. It may sound "obviously true." Callen and Greene,[13] for instance, do not even dignify it by the term postulate. It is, however, very strong: for a Gaussian process, it can be shown to be equivalent to the Markoff assumption. We can see this intuitively by stating the postulate: the system does not remember how it got to the given state.

The essence of the proof is given by Doob.[14] It is easily shown that the average regressions, together with the one-gate distribution [known from Boltzmann's principle, Eq. (2-11)], completely determine the statistics of a stationary Gaussian process. We have a stationary Gaussian process whose average regressions obey a system of linear differential equations of first order with constant coefficients (2-13). Then the covariance matrix $\mathbf{A}(\tau) \equiv \langle \boldsymbol{\alpha}_{tr}(t)\boldsymbol{\alpha}(t+\tau) \rangle$ has the "translation operator" property,

$$\mathbf{A}(\tau_1)\mathbf{A}(\tau_2) = \mathbf{A}(\tau_1+\tau_2), \quad (3\text{-}11)$$

which is a sufficient condition for the process to be Markoffian.

We now have an answer to the question "How do you know you have taken enough variables (enough measurements on the system) for it to be Markoffian?" (Example: a two-mesh RC circuit is Markoffian in the two mesh charges, but not in any single variable.) If we accept the Gaussian assumption and the postulate about the average regressions, then the condition that the irreversible behavior of the system be given by equations of the form (2-2) is sufficient for the Markoff property. It is important to realize that this "prediction property" of the phenomenological laws depends on the assumption (Sec. 2) that enough α's have been taken to determine the entropy completely.

[13] H. B. Callen and R. F. Greene, Phys. Rev. **86**, 702 (1952); **88**, 1387 (1952).
[14] Reference 3, V Sec. 8.

4. GAUSSIAN MARKOFF PROCESS

Single Variable

We again consider a single variable α, obeying the equation

$$R\dot{\alpha}+s\alpha = \epsilon \quad (4\text{-}1)$$

(rhs purely random: α is Markoffian). We have for the conditional p.d.f.:

$$f_1\binom{\alpha^{(2)} \ \Big| \ \alpha^{(1)}}{t+\tau \ \Big| \ t} = (2\pi)^{-\frac{1}{2}}\frac{s}{k}(1-e^{-2\gamma\tau})^{-\frac{1}{2}}$$

$$\times \exp\left\{ -\frac{1}{2}\frac{s}{k}\frac{[\alpha^{(2)}-e^{-\gamma\tau}\alpha^{(1)}]^2}{(1-e^{-2\gamma\tau})} \right\}, \quad (4\text{-}2)$$

(where $\gamma=s/R$), as is verified by inspection: it gives the right one-gate distribution (2-11, let $\tau\to\infty$) and the right average regression ($\alpha^{(1)}e^{-\gamma\tau}$). The normalization factor will henceforth be omitted to save writing.

This formula, together with Eq. (3-8), constitutes the solution to the problem of finding the probability of any path. The purpose of this paper is to put the formula into a particularly interesting form.

Let us divide the interval $(t, t+\tau)$ into p small, equal sub-intervals, i.e., introduce $p-1$ new gates at instants $t_1=t$, $t_2=t+\Delta\tau$, $\cdots$, $t_{p+1}=t+\tau=t+p\Delta\tau$. Then we have

$$f_1\binom{\alpha^{(p+1)} \ \Big| \ \alpha^{(1)}}{t_{p+1} \ \Big| \ t_1}$$

$$= \int (p-1)\text{-fold} \int f_1\binom{\alpha^{(p+1)} \ \Big| \ \alpha^{(p)}}{t_{p+1} \ \Big| \ t_p} \cdots$$

$$\times f_1\binom{\alpha^{(2)} \ \Big| \ \alpha^{(1)}}{t_2 \ \Big| \ t_1} d\alpha^{(2)}\cdots d\alpha^{(p)}. \quad (4\text{-}3)$$

This is the Chapman-Kolmogoroff relation,[15] valid for all Markoff processes. Evidently there is no limit to the number $(p-1)$ of gates which are "integrated out." It can be made large enough ($\Delta\tau$ small enough) so that the stochastic differential equation (4-1) is well approximated by the stochastic difference equation

$$\alpha^{(k)}-\lambda\alpha^{(k-1)} = y^{(k)}, \quad (4\text{-}4)$$

where $\lambda=1-\gamma\Delta\tau$ and $y^{(k)}=\epsilon(t_k)\cdot\Delta\tau/R$. Supposing the y's to have Gaussian distributions with mean zero and variance σ_y^2, (4-3) becomes[16]

$$f_1\binom{\alpha^{(p+1)} \ \Big| \ \alpha^{(1)}}{t_{p+1} \ \Big| \ t_1} \propto \int (p-1)\text{-fold}$$

$$\times \int \exp\left\{ -\frac{1}{2\sigma_y^2}[(\alpha^{(p+1)}-\lambda\alpha^{(p)})^2+\cdots \right.$$

$$\left. + (\alpha^{(2)}-\lambda\alpha^{(1)})^2] \right\} d\alpha^{(1)}\cdots d\alpha^{(p)}. \quad (4\text{-}5)$$

[15] A. Kolmogoroff, Math. Ann. **104**, 415 (1931).
[16] This is stronger than assuming α to be a Gaussian process and is not necessary for the proof of (4-7).

If the integrations are performed in succession, we arrive again at (4-2). We now notice that each integration may be replaced by taking the minimum with respect to the corresponding variable in the exponent, obtaining

$$f_1\left(\begin{matrix}\alpha^{(p+1)}\\t_{p+1}\end{matrix}\middle|\begin{matrix}\alpha^{(1)}\\t_1\end{matrix}\right) \propto \exp-\frac{1}{2\sigma_y{}^2}[(\alpha^{(p+1)}-\lambda\alpha^{(p)})^2+\cdots$$

$$+(\alpha^{(2)}-\lambda\alpha^{(1)})^2]_{\min},$$

with respect to $\alpha^{(2)}, \cdots\alpha^{(p)}$;   (4-6)

what has been done is to replace an "average value" by a "most probable value." For Gaussian distributions these are identical.

As p is increased, the sum in the exponent becomes an integral. The value of the integral is specified by the condition that the integral be a minimum subject to given values of α at the end points. Reverting to differential notation, (4-6) becomes

$$f_1\left(\begin{matrix}\alpha^{(p+1)}\\t_{p+1}\end{matrix}\middle|\begin{matrix}\alpha^{(1)}\\t_1\end{matrix}\right)$$

$$\propto \exp\left\{-\frac{1}{4}\frac{1}{k}\left(\int_{t_1}^{t_{p+1}} R[\dot\alpha(t)+\gamma\alpha(t)]^2 dt\right)_{\min}\right.,$$

subject to $\alpha(t_1)=\alpha^{(1)}, \cdots, \alpha(t_{p+1})=\alpha^{(p+1)}$.   (4-7)

The direct verification of this formula—seeing that it gives the correct average regression and one-gate p.d.f.—is simpler and more general than the above "synthetic" proof. To find the regressions we again utilize the identity of means and modes of Gaussian distributions. The regression from the value $\alpha^{(1)}$ will therefore be given by the condition

$$f_1\left(\begin{matrix}\alpha^{(2)}\\t_2\end{matrix}\middle|\begin{matrix}\alpha^{(1)}\\t_1\end{matrix}\right)=\max, \tag{4-8}$$

where the variation is with respect to $\alpha^{(2)}$. As the integrand in (4-7) is always positive or zero, so is the integral. If it can take the value zero, this is its minimum. But this is possible only if the integrand is everywhere zero, i.e., the condition for a minimum is

$$\dot\alpha+\gamma\alpha=0, \tag{4-9}$$

which is recognized as the Eq. (2-13) for the average regression. As the initial condition $\alpha(t_1)=\alpha^{(1)}$ also agrees, (4-7) gives the correct regression.

The one-gate distribution is obtained from the conditional p.d.f. by taking $t_1=-\infty$ and $\alpha^{(1)}=0$. (The aged system certainly was at equilibrium some time long

ago.) We want to prove that

$$f_1\left(\begin{matrix}\alpha^{(2)}\\t_2\end{matrix}\middle|\begin{matrix}0\\-\infty\end{matrix}\right) \propto \exp\left[-\frac{1}{k}\frac{1}{2}s(\alpha^{(2)})^2\right]. \tag{2-11}$$

Letting $\mathcal{L}(\alpha,\dot\alpha)=R[\dot\alpha-\gamma\alpha]^2$, the Euler-Lagrange equation for an extreme value of the integral is

$$\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot\alpha}-\frac{\partial\mathcal{L}}{\partial\alpha}=0, \tag{4-10}$$

or

$$\ddot\alpha-\gamma^2\alpha=0, \tag{4-11}$$

whose solution is a superposition of the functions $e^{-\gamma t}$ and $e^{+\gamma t}$. The initial condition at $t=-\infty$ eliminates the exponentially falling solution, that at $t=t_2$ gives $\alpha(t)=\alpha^{(2)}e^{\gamma t}$, and the integral is

$$\frac{1}{4}\int_{-\infty}^{t_2} R[\dot\alpha(t)+\gamma\alpha(t)]^2 dt_{\min}=\frac{1}{2}s(\alpha^{(2)})^2,$$

subject to $\alpha(-\infty)=0, \alpha(t_2)=\alpha^{(2)}$ (4-12)

as was to be proved.

Several Variables

The possibility of linearly transforming the α's so as to diagonalize the two matrices **s** and **R** simultaneously, implies that the mathematics is no harder for several variables. It can be shown (see dissertation, footnote *) that in the diagonal representation the random impulse functions ϵ_i [Eq. (3-9)] must be uncorrelated, so that we are dealing with n independent Markoff processes. The extension of (4-7) to several variables will point up the physical significance of the integrand $\mathcal{L}$.

For independent processes, the joint p.d.f. is simply the product of the single-variable p.d.f.'s. Thus, in the multivariable extension of (4-7), we shall have a sum in the exponent. If the integrand is multiplied out, and the integration performed on the cross term,

$$R[\dot\alpha(t)+\gamma\alpha(t)]^2=R\dot\alpha^2+(1/R)s^2\alpha^2+(d/dt)(s\alpha^2), \tag{4-13}$$

(recall $\gamma=s/R$), the sum can be written in terms of the entropy and dissipation functions. For in the diagonal representation, these are sums of squares:

$$S=S_0-\frac{1}{2}\Sigma_i s_i\alpha_i{}^2 \text{ [see (2-10)]}, \tag{4-14}$$

$$\Phi(d\alpha/dt, d\alpha/dt)=\frac{1}{2}\Sigma_i R_i\dot\alpha_i{}^2 \text{ [see (2-8)]}, \tag{4-15}$$

$$\Psi(X,X)=\frac{1}{2}\Sigma_i(1/R_i)X_i{}^2=\frac{1}{2}\Sigma_i(1/R_i)s_i{}^2\alpha_i{}^2 \text{ [see (2-9)]}. \tag{4-16}$$

Then

$$\Sigma_i R_i[\dot\alpha_i+\gamma\alpha_i]^2=2\Phi(d\alpha/dt, d\alpha/dt)+2\Psi(X,X)$$
$$-2(d/dt)S(\alpha), \tag{4-17}$$

and the conditional p.d.f. (4-7) for several variables

becomes

$$f_n\left(\begin{array}{c|c} \boldsymbol{\alpha}^{(2)} & \boldsymbol{\alpha}^{(1)} \\ t_2 & t_1 \end{array}\right) = \prod_i f_1\left(\begin{array}{c|c} \alpha_i^{(2)} & \alpha_i^{(1)} \\ t_2 & t_1 \end{array}\right)$$

$$\propto \exp\left\{-\frac{1}{4}\frac{1}{k}\left(\int^{t_2}\left[2\Phi(d\alpha/dt,\, d\alpha/dt)\right.\right.\right.$$

$$\left.\left.\left.+2\Psi(X,X)-2\frac{d}{dt}S(\alpha)\right]dt\right)_{\min}\right\}$$

subject to $\alpha(t_1)=\boldsymbol{\alpha}^{(1)}$, $\alpha(t_2)=\boldsymbol{\alpha}^{(2)}$. (4-18)

In this form, the equation is no longer dependent on the representation chosen, i.e., it is valid also for interacting variables. It must be noticed that Φ is a function of the $\dot{\alpha}$'s, while Ψ, explicitly a function of the X's, is therefore a function of the α's (of the state).

The joint two-gate p.d.f. has a similar appealing form

$$f_{2n}\left(\begin{array}{cc} \boldsymbol{\alpha}^{(1)} & \boldsymbol{\alpha}^{(1)} \\ t_1 & t_2 \end{array}\right) = f_n\left(\begin{array}{c|c} \boldsymbol{\alpha}^{(2)} & \boldsymbol{\alpha}^{(1)} \\ t_2 & t_1 \end{array}\right)f_n(\boldsymbol{\alpha}^{(1)})$$

$$\propto \exp\left[\frac{1}{k}\left\{\frac{1}{2}S(\boldsymbol{\alpha}^{(1)})+\frac{1}{2}S(\boldsymbol{\alpha}^{(2)})\right.\right.$$

$$\left.\left.-\frac{1}{4}\left(\int_{t_1}^{t_2}\left[2\Phi(d\alpha/dt,\, d\alpha/dt)+2\Psi(X,X)\right]dt\right)_{\min}\right\}\right],$$

subject to $\alpha(t_1)=\boldsymbol{\alpha}^{(1)}$, $\alpha(t_2)=\boldsymbol{\alpha}^{(2)}$. (4-19)

The form also lends itself to the expression of the p.d.f. for any number of gates. Using (4-18) in (3-8), we simply add the integrals in the exponents.

$$f_{pn}\left(\begin{array}{ccc} \boldsymbol{\alpha}^{(1)} & \cdots & \boldsymbol{\alpha}^{(p)} \\ t_1 & & t_p \end{array}\right) \propto \exp\left[\frac{1}{k}\left\{\frac{1}{2}S(\boldsymbol{\alpha}^{(1)})+\frac{1}{2}S(\boldsymbol{\alpha}^{(p)})\right.\right.$$

$$\left.\left.-\frac{1}{4}\left(\int_{t_1}^{t_p}\left[2\Phi(d\alpha/dt,\, d\alpha/dt)+2\Psi(X,X)\right]dt\right)_{\min}\right\}\right]$$

subject to $\alpha(t_1)=\boldsymbol{\alpha}^{(1)}$, $\alpha(t_2)=\boldsymbol{\alpha}^{(2)}$, $\cdots$, $\alpha(t_p)=\boldsymbol{\alpha}^{(p)}$. (4-20)

It may be pointed out that the path which makes the integrand a minimum has, in general, discontinuous derivatives at each gate. The separate two-gate paths are just strung end-to-end.

We can make (4-20) particularly simple by permitting integration over an infinite time interval. We assert that

$$\frac{1}{4}\int_{-\infty}^{t_1}\left[2\Phi(d\alpha/dt,\, d\alpha/dt)+2\Psi(X,X)\right]dt_{\min}$$

$$=\frac{1}{2}S(\boldsymbol{\alpha}^{(1)})+\text{const} \quad (4\text{-}21)$$

subject to $\alpha(-\infty)=0=$ equilibrium, $\alpha(t_1)=\boldsymbol{\alpha}^{(1)}$.

This can be checked directly by working it out (say, in

diagonal form). But it can also be seen as follows: the Euler-Lagrange equations (4-10) for an extremum of the integral have double the order of Eqs. (2-13), the phenomenological laws. They are evidently satisfied by the solutions of (2-13), and by the mirror images of these in time ($-t$ replacing t). The former are the "average paths," for which hold Eqs. (2-6) and (2-7). But only the "mirror image" solutions can satisfy the initial conditions of $t=-\infty$ nontrivially. For these

$$2\Phi(d\alpha/dt,\, d\alpha/dt)+2\Psi(X,X)=-2\dot{S}, \quad (4\text{-}22)$$

from which (4-21) follows.[17] From symmetry in past and future a similar equation holds for the range of integration (t_p, ∞). If we substitute these into (4-20) we obtain

$$f_p\left(\begin{array}{ccc} \boldsymbol{\alpha}^{(1)} & & \boldsymbol{\alpha}^{(p)} \\ t_1 & \cdots & t_p \end{array}\right) \propto \exp\left[-\frac{1}{4}\frac{1}{k}\left(\int_{-\infty}^{\infty}\left[2\Phi(d\alpha/dt,\, d\alpha/dt)\right.\right.\right.$$

$$\left.\left.\left.+2\Psi(X,X)\right]dt\right)_{\min}\right], \quad (4\text{-}23)$$

subject to $\alpha(t_1)=\boldsymbol{\alpha}^{(1)}$, $\alpha(t_2)=\boldsymbol{\alpha}^{(2)}$, $\cdots$, $\alpha(t_p)=\boldsymbol{\alpha}^{(p)}$.

A few disconnected remarks may elucidate the physical content of this relation. To form the integrand one must know the form of the dissipation function as well as of the entropy function; the latter is needed in order to give the thermodynamic forces X_i as functions of the state. The path whose probability it is desired to evaluate may be specified as closely as desired, i.e., by any number of gates. Accordingly, the auxiliary functional (4-23) furnishes a complete description of the probability distribution functional. Between gates the path of integration is in a certain sense the "smoothest" path. It is a superposition of decaying and growing exponentials, as required by symmetry in past and future.

The theorem which has been proved is seen to be analogous to the Boltzmann principle. The latter tells the probability of a state in terms of its entropy; this theorem tells the probability of a temporal succession of states in terms of the entropy and dissipation functions.

Incomplete Specification in Some Gates

An important feature of the expression (4-23) is that any variable can be "integrated out" simply by minimizing the exponent with respect to it; i.e., by removing the variable from the set of specifications on the minimum of the integral. For the p.d.f. is Gaussian in all the variables. We can thus invoke the equivalence of means and modes and replace integration over one variable (averaging with respect to it) by maximization of the exponential with respect to it, i.e., taking the minimum of the integral. *Accordingly* (4-23) *is valid as*

---

[17] This is worked out in II, Sec. 3.

it stands whether or not the state is completely specified at each instant. It therefore describes the statistics of sets of variables which are not necessarily Markoffian, but which are sub-sets of Markoffian sets for which the coefficients of the entropy and dissipation functions are known.

Principle of Least Dissipation

The principle of least dissipation (RRIP I), a generalization of a similar principle in hydrodynamics due to Rayleigh,[18] can be read by inspection from (4-18). Given a state α, we ask for the most probable values of the $\dot\alpha$'s (and thus for the most probable path of the α's in time). Taking t_2 close to t_1 ($t_2 - t_1 = \Delta t$ = small) we seek the state $\alpha^{(2)}$ which will maximize the exponent. The integral becomes a simple product:

$$-\frac{1}{4}\frac{1}{k}\left[2\Phi(d\alpha/dt, d\alpha/dt) + 2\Psi(X, X) - 2\frac{d}{dt}S(\alpha)\right]\Delta t = \max.$$

$$(4\text{-}24)$$

If we remember that the variation is with respect to

the $\dot\alpha$'s only, we have

$$\dot S - \Phi(d\alpha/dt, d\alpha/dt) = \max. \qquad (4\text{-}25)$$

This variational principle is formally equivalent to the phenomenological laws [(2-2), including the reciprocal relations (2-4)].

Spectral Description

It may be useful to recall the connection between this work and that of Callen and his co-workers in the same field.[19] The systems treated are of the same type, and the same physical assumptions are made for them. But while this paper uses a *temporal* description of the course of fluctuations, these other papers use a *spectral* description. The two descriptions correspond to two types of experiment: the temporal to taking successive readings at close intervals, obtaining, for instance, a trace on a moving tape; the spectral to recording by means of a frequency analyzer. Where in the temporal description we postulate linear relations between forces and fluxes, defined by the dissipation matrix $\mathbf{R}$, the spectral description has linear relations between their Fourier transforms, defined by an admittance matrix.

---

[18] Lord Rayleigh (J. W. Strutt), Phil. Mag. **26**, 776 (1913).

[19] Footnote 13. See also H. Takasi, J. Phys. Soc. Japan **7**, 439 (1952).

Fluctuations and Irreversible Process. II. Systems with Kinetic Energy*

S. MACHLUP, *Sloane Physics Laboratory, Yale University, New Haven, Connecticut,*

and

Bell Telephone Laboratories, Murray Hill, New Jersey

AND

L. ONSAGER, *Sterling Chemistry Laboratory, Yale University, New Haven, Connecticut*

(Received May 13, 1953).

The results of the previous paper are extended to second-order systems, i.e., systems with inertia. Using a generalized definition of the thermodynamic forces, reciprocal relations for the dissipative coefficients in the equations describing irreversible processes are derived. A dissipation function can again be defined, and it can again be used to express the probability functional for fluctuations.

1. β-TYPE VARIABLES

IN the previous paper of the same principal title[1] the thermodynamic variables treated were limited to those satisfying the hypothesis of microscopic reversibility. The restriction was stated: on a kinetic model, the thermodynamic variables must be algebraic sums of (a large number of) molecular variables, and must be *even* functions of those molecular variables which are odd functions of time (like molecular velocities). Thus

---

* Much of the work reported in this paper appears in more detail in a dissertation by one of us (S.M.), presented to the Faculty of the Graduate School of Yale University (1952) in partial fulfillment of the requirements for the Ph.D. degree.

[1] L. Onsager and S. Machlup, preceding paper [Phys. Rev. **91**, 1505 (1953)].

a reversal of the sense of time would not change the values of these α variables.

It was assumed that all thermodynamic functions (specifically, the entropy S) were given as functions of the α variables, and that the irreversible processes were "caused" by thermodynamic forces $X_i = \partial S/\partial\alpha_i$. If, however, the kinetic energy of the various flows (e.g., magnetic energy of electric currents) contributes appreciably to the entropy, other variables must be taken into consideration: "velocity" variables, i.e., variables which would change their sign if the sense of time were reversed. Casimir[2] calls them β variables. He points

---

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