ten, müssen die Zusatzreflexe ebenen, parallel zu den {311}-Ebenen des Kupfers gelegenen Oberflächenbereichen zugeordnet werden, wobei die Streuzentrenanordnung dieser Bereiche mit der Struktur der {311}-Kupfer-Vizinalflächen übereinstimmt [p(1 × 1)-Struktur].

Im weiteren Verlauf der Reduktion treten in der Nähe der BRAGG-Spannungen für das Kupfer wieder die diffusen Intensitätsflecken auf (Abb. 23 b). Das bedeutet, daß ein Teil der bestrahlten Oberfläche am (100)-Pol die ursprüngliche, vom Sauerstoff unbeeinflußte sphärische Form angenommen hat. Noch sind aber die {410}- und {311}-Vizinalflächen stark ausgeprägt, was durch die nur geringfügig verminderte Intensität der Beugungsreflexe bewiesen wird. Setzt man die Wasserstoffbehandlung fort, so verschwinden schließlich alle Interferenzen der Vizinalflächen, und zwar diejenigen der {410}-Flächen gleichzeitig mit denen, welche von den {311}-Flächen herrühren. Aus dieser Tatsache und der erforderlichen relativ hohen Reaktionstemperatur folgt die Möglichkeit, daß sich die {410}-Oberflächenschichten zunächst in {311}-Bereiche umordnen, von denen der Sauerstoff wahrscheinlich leichter entfernt werden kann.

Übrig bleiben die nunmehr intensiven diffusen Beugungsflecken (Abb. 23 c). Der (100)-Pol ist vollständig reduziert. Sein Beugungsbild entspricht genau dem eines getemperten Würfelpoles einer frisch geschmolzenen Kupferkugel (s. Abb. 5).

Herrn Prof. Dr. K. MOLIÈRE möchte ich meinen besonderen Dank für sein stets förderndes Interesse aussprechen. Für fruchtbare Diskussionen danke ich den Herren Dr. G. LEHMPFUHL und Dr. K. KAMBE ebenso wie Herrn Prof. Dr. E. MENZEL und Herrn Dr. O. SCHOTT. Ich bedanke mich außerdem bei Herrn H.-J. KRAUSS für die Anfertigung der Zeichnungen. — Der Max-Planck-Gesellschaft und der Deutschen Forschungsgemeinschaft bin ich für die Gewährung von Stipendien zu Dank verpflichtet.

Fluctuation Theory of Irreversible Processes

CH. SCHNEEWEISS *

Institut für Physikalische Chemie der Universität Frankfurt/Main

(Z. Naturforschg. **22 a**, 1671—1677 [1967]; received 13 June 1967)

LANGEVIN equations of the type $d^n x(t)/dt^n + \ldots + c\,x(t) = \boldsymbol{K}(t)$ constitute the starting point of a phenomenological fluctuation theory of irreversible processes. These equations are not constructed from transport equations (as in the older theory), but via a generalized MASTER equation from phase space mechanics. The MARKOFF processes of first and higher order defined by the various LANGEVIN equations are studied by the prediction theory of stationary stochastic processes. Instead of the variation principle of the ONSAGER–MACHLUP theory one has the minimization of the prediction error. The mean relaxation path and the entropy of the considered processes are calculated. It is shown that the entropy consists of one part which is given by the relaxation path and another which is determined by the prediction error.

1. Introduction

The fluctuation theory of irreversible processes starts from stochastic differential equations of the type

$$\frac{d^n x(t)}{dt^n} + \ldots + c\,x(t) = \varkappa(t) \qquad (1.1)$$

where $\varkappa(t)$ is a random force, $x(t)$ an external (thermodynamic) parameter and c a constant. Eq. (1.1) is called a LANGEVIN equation of order n.

Originally [1] these equations were derived from the phenomenological equations of irreversible processes. This, however, had as a main disadvantage that the nature of the process $\varkappa(t)$ could only be guessed. Later, VAN KAMPEN [2] was able to remove partly this deficiency in showing the intimate relation between the theory of the phenomenological phase space [3] and fluctuation theory. But since VAN KAMPEN's theory of phenomenological phase space is restricted to MARKOFF processes he was only able

* Present address: Institut für Gesellschafts- und Wirtschaftswissenschaften der Universität Bonn, Statistische Abteilung, 53 Bonn, Lennéstr. 37.

[1] L. ONSAGER and S. MACHLUP, Phys. Rev. **91**, 1505 [1953].
[2] N. G. VAN KAMPEN, Physica **23**, 707 [1957].
[3] M. S. GREEN, J. Chem. Phys. **20**, 1281 [1952].

to derive a first order LANGEVIN equation. Furthermore there is no consequent derivation of the theory of phenomenological phase space from the LIOUVILLE equation, and it is not explicitly seen which approximations have to be introduced in order to arrive at the starting point of fluctuation theory.

The situation now is by far better. The (phenomenological phase space) theory of ZWANZIG [4] provides us with a statistical basis which will prove to be broad enough to derive not only a first order fluctuation theory but also all higher order LANGEVIN equations. By this theory we shall also be able to show explicitly the succession of approximations that have to be introduced to derive the fundamental equations of fluctuation theory from the LIOUVILLE equation.

Having established the validity of the LANGEVIN equations we shall then calculate the mean relaxation path of the irreversible process and will give a new derivation of the fluctuation dissipation theorem [5]. Doing this we shall not use the variation method of ONSAGER and MACHLUP [1]. Instead we will introduce the prediction theory [6] of stationary stochastic processes into statistical mechanics. This theory has not only the advantage of particular mathematical transparency and elegance, but gives also a new suggestive interpretation of entropy.

2. The Theory of Phenomenological Phase Space

Let us first recall some notations and results of ref. [4]. Let $x=\{q_1,\ldots,q_N,p_1,\ldots,p_N\}$ be a vector in phase space Γ, $A(x)$ a set of phase functions A_j $(j=1,\ldots,n)$, and $a_j=A_j(x)$ a special value of the phase function A_j; then the n dimensional linear manyfold of vectors $a=\{a_1,\ldots,a_n\}$ is called n dimensional *phenomenological phase space or a space.* The connection between phase space Γ and a space is given by the projection operator

$$P_a\ldots\equiv\frac{1}{W(a)}\int \mathrm{d}x\,\delta(A(x)-a)\ldots \tag{2.1}$$

where $W(a)=\int \mathrm{d}x\,\delta(A(x)-a)$ is, as usual, the normalization. The probability density $g(a,t)$ in a space is then given by

$$g(a,t)=\int \mathrm{d}x\,(A(x)-a)\,f(x,t) \tag{2.2}$$

with $f(x,t)$ being the phase space density.

If one now introduces (2.1) into the "damping form" of the LIOUVILLE equation [ref. [4], Eq. (18)], one gets [imposing the initial condition $g(a,0)=\delta(a-a_0)$] a kinetic equation in a space [ref. [4], Eq. (24)]

$$\begin{aligned}\dot{g}(a,t)=&-i\int \mathrm{d}a'\int \mathrm{d}x\,\delta(A(x)-a)L\,\delta(A(x)-a')\,\frac{g(a',t)}{W(a)}\\&-\int_0^t \mathrm{d}s\int \mathrm{d}a'\int \mathrm{d}x\,\delta(A(x)-a)\,L\,e^{-is(1-P_a)L}\,(1-P_a)\,\delta(A-a')\,\frac{g(a,t-s)}{W(a')}\end{aligned} \tag{2.3}$$

(L being as usual the LIOUVILLE operator).

Introducing the definitions

$$U(a,a',s)=i\int \mathrm{d}x\,\delta(A(x)-a)\,L\,\delta(A-a')\,\delta(s-s_0)\quad\text{with}\quad 0<s_0<t, \tag{2.4}$$

$$G(a,a',s)=\int \mathrm{d}x\,\delta(A(x)-a)\,L\,e^{-is(1-P_a)L}\,(1-P_a)\,L\,\delta(A-a'), \tag{2.5}$$

and

$$D(a,a',s)=-\{U(a,a',s)+G(a,a',s)\}\,W^{-1}(a'), \tag{2.6}$$

Eq. (2.3) can be written in the form

$$\dot{g}(a,t)=\int_0^t \mathrm{d}s\int \mathrm{d}a'\,D(a,a',s)\,g(a,t-s). \tag{2.7}$$

This important equation is called a generalized MASTER equation in a space. It shows a memory effect in that the change of $g(a,t)$ is not only given by the presence but also by all times of the past.

[4] R. ZWANZIG, Phys. Rev. **124**, 983 [1961].
[5] R. S. DE GROOT and P. MASUR, Non-Equilibrium Thermodynamics, North Holland Publishing Co., Amsterdam 1962.
[6] A. PAPOULIS, Probability, Random Variables and Stochastic Processes, McGraw-Hill Book Co., New York 1965.

Consider now a power series expansion of $D(a, a', s)$ after potentials of $a - a'$. This expansion stops after the second term [7] if one introduces (following ZWANZIG) the assumption

$$O\{(LP)^2\} \gg O\{(LP)^3\} . \tag{2.8}$$

Introducing L and P one can show that (2.8) amounts to the postulate that the considered irreversible process be "sufficiently slow". With (2.8) the expansion yields the generalized FOKKER–PLANCK equation [8]

$$\dot{g}(a,t) = -\sum_r^n \frac{\partial}{\partial a_r}\{v_r(a)\, g(a,t)\} + \int_0^t \mathrm{d}s \sum_r^n \sum_q^n \frac{\partial}{\partial a_r}\left\{K_{qr}(a,s)\, W(a) \frac{\partial}{\partial a_q} \frac{g(a,t-s)}{W(a)}\right\} \tag{2.9}$$

where

$$v_r(a) = \frac{1}{W(a)} \int \mathrm{d}x\, \delta(A(x) - a)\, \dot{A}_r(x) \tag{2.10}$$

and

$$K_{qr}(a,s) = \frac{1}{W(a)} \int \mathrm{d}x\, \delta(A(x) - a)\, \dot{A}_q (1 - P_a)\, e^{-isl}\, \dot{A}_r \tag{2.11}$$

($\dot{A}$ denoting the total time derivative of A). Eq. (2.9) will be the starting point of a generalized fluctuation theory. It still contains a time convolution which we now shall remove by extending the phenomenological phase space.

3. Extension of Phenomenological Phase Space

Let us return to Eq. (2.7). This equation may be interpreted as a differential CHAPMAN–KOLMOGOROFF equation of a non-MARKOFFian process [9]. Such a process, however, can always be described as a MARKOFF process by extending its state space [10]. An increase of the number of dimensions of the state space implies a contraction of the memory interval. It seems to be natural (in view of thermodynamics of irreversible processes) to extend the phenomenological phase space, in which the considered relaxation process is described, by the time derivations of the describing variables A_j

$$\dot{A}_j(x) \equiv B_j(x) . \tag{3.1}$$

Thus we get instead of (2.9), applying (2.8) also to the B variables,

$$\begin{aligned} \dot{g}(a,b,t) = &-\sum_{l=1}^n \frac{\partial}{\partial a_l}\{v_l(a,b)\, g(a,b,t)\} - \sum_{\lambda=1}^n \frac{\partial}{\partial b_\lambda}\{u_\lambda(a,b)\, g(a,b,t)\} \\ &+\int_0^t \mathrm{d}s \sum_{\lambda,\varkappa}^n \frac{\partial}{\partial b_\lambda}\left\{K_{\lambda\varkappa}(a,b,s)\, W(a,b) \frac{\partial}{\partial b_\varkappa} \frac{g(a,b,t-s)}{W(a,b)}\right\} \end{aligned} \tag{3.2}$$

where the notations used here are quite analogue to those after Eq. (2.9). Particularly we have

$$P_{a,b}\ldots = \frac{1}{W(a,b)} \int \mathrm{d}a \int \mathrm{d}b\, \delta(A(x) - a)\, \delta(\dot{A}(x) - b)\ldots , \tag{3.3}$$

$$v_l(a,b) = P_{a,b}\, \dot{A}_l = b_l P_{a,b} , \tag{3.4}$$

$$u_\lambda(a,b) = P_{a,b}\, \ddot{A}_\lambda \tag{3.5}$$

and

$$K_{\lambda\varkappa}(a,b,s) = P_{a,b}\, \ddot{A}_\lambda (1 - P_{a,b})\, e^{-isL}\, \ddot{A}_\varkappa . \tag{3.6}$$

Let us now assume that the relaxation time τ_C of not described internal parameters C are negligibly small compared with the relaxation time τ_B of parameters B. This permits us to take the "PAULI limit" [11] of Eq. (3.2), i. e. we shall (in a well known way) remove the time convolution integral in Eq. (3.2). If the condition $\tau_B \gg \tau_C$ is not satisfied one can take into account further time derivatives of A_j. One will than arrive at LANGEVIN equations of higher than of second order. In this paper, however, we shall restrict our-

[7] CH. SCHNEEWEISS, Thesis, Naturwissenschaftl. Fakultät der Universität Frankfurt a. M.

[8] This equation has also been derived by ZWANZIG [ref. [4], Eq. (27)] in a mathematical easier but less straight forward way.

[9] See e. g. K. KRICKEBERG, Wahrscheinlichkeitstheorie, B. G. Teubner Verlagsgesellschaft, Stuttgart 1963, p. 160.

[10] J. DOOB, Stochastic Processes, John Wiley & Sons, New York 1953.

[11] I. OPPENHEIM and K. E. SHULER, Phys. Rev. **138**, B 1007 [1965].

selves to the PAULI limit of Eq. (3.2), since the higher order case does not show up anything new. Thus we get from (3.2)

$$\dot{g}(a,b,t) = -\sum_{l=1}^{n} \frac{\partial}{\partial a_l}\{b_l\, g(a,b,t)\} - \sum_{\lambda=1}^{n} \frac{\partial}{\partial b_\lambda}\{u_\lambda(a,b)\; g(a,b,t)\} + \sum_{\lambda,\varkappa}^{n} \frac{\partial}{\partial b_\lambda}\left\{K_{\lambda\varkappa}(a,b)\; W(a,b)\; \frac{\partial}{\partial b_\varkappa}\frac{g(a,b,t)}{W(a,b)}\right\} \tag{3.7}$$

where $K_{\lambda\varkappa}(a,b) = \int_0^\infty \mathrm{d}s\, K_{\lambda\varkappa}(a,b,s)$, and (3.4) has been used.

In order to get the ordinary LANGEVIN equations of fluctuation theory, let us "linearize" Eq. (3.7) and assume

$$U_\lambda(a,b) \approx U_\lambda(0,0) + \sum_{j=1}^{n} \eta_{\lambda j}\, a_j\,, \tag{3.8}$$

$$K_{\lambda\varkappa}(a,b) \approx K_{\lambda\varkappa}(0,0) \tag{3.9}$$

and for the equilibrium distribution [12] $W(a,b)$,

$$W(a,b) \sim \exp\left\{-\frac{1}{2k}\left(\sum_{k,l}^{n} g_{kl}\, a_k\, a_l + \sum_{\lambda,\varkappa}^{n} h_{\lambda\varkappa}\, b_\lambda\, b_\varkappa\right)\right. \tag{3.10}$$

k being the BOLTZMANN constant.

Using (3.10) one can easily show [2, 7] that $U_\lambda(0,0) = 0$ and (3.7) takes the form

$$\dot{g}(a,b,t) = -\frac{\partial}{\partial a}\{b\, g(a,b,t)\} - \frac{\partial}{\partial b}\{\eta\, a\, g(a,b,t)\} + \frac{\partial}{\partial b}\left\{K(0,0)\; W(a,b)\; \frac{\partial}{\partial b}\frac{g(a,b,t)}{W(a,b)}\right\} \tag{3.11}$$

where, simply for reasons of convenience, only one A and one B variable are considered.

4. Langevin Equations

We are now in the position to derive the desired LANGEVIN equations of an extended fluctuation theory. The only thing to do is to change the representation of the described stochastic processes, i. e. one has to change from a probability density representation to a representation by stochastic differential equations. This can easily be achieved [7].

Using (3.8), (3.9) and (3.10), Eq. (3.11) becomes

$$\dot{A}(t) = B(t)\,,$$
$$\dot{B}(t) + 2M\, B(t) + \omega_0^2\, A(t) = \varkappa(t) \tag{4.1}$$

where the phase functions $A(x) = A(x(t)) = A(t)$ and $B(t)$ are now considered as a stochastic (vector) process; and

$$2M \equiv k^{-1} K h\,, \qquad \omega_0^2 \equiv g\, h^{-1}.$$

$\varkappa(t)$ is called stationary white noise. One has $\varkappa(t) = \mathrm{d}W(t)$, where $W(t)$ is a WIENER process, and in view of (3.11) $E(\varkappa(t)) = 0$ and

$$E\{\varkappa(t)\, \varkappa(t')\} = c\, \delta(t - t'). \tag{4.2}$$

[12] See e. g. ref. [5].

(E being the expectation and c a constant.) Since (3.11) is a FOKKER–PLANCK equation, $\varkappa(t)$ is of course a stationary GAUSSian–MARKOFF process and $\{A(t), B(t)\}$ a MARKOFF process of first order. This is equivalent to a MARKOFF process of 2. order [6] if one writes instead of (4.1)

$$\ddot{A}(t) + 2M\, \dot{A}(t) + \omega_0^2\, A(t) = \varkappa(t). \tag{4.3}$$

If the transition to the PAULI limit is allowed already in Eq. (2.9) (i. e. if one has for the relaxation times of the A parameters $\tau_A \gg \tau_B$) one gets in an analogous way

$$\dot{A}(t) + \frac{\omega_0^2}{2M}\, A(t) = \varkappa'(t) \tag{4.4}$$

which represents of course the long time behaviour of the process (4.3).

In order to apply the prediction theory to the processes (4.3) and (4.4), let us calculate their correlation functions. Since $\varkappa(t)$ is a stationary process, $A(t)$ [being a linear transformation of $\varkappa(t)$] is also stationary. Therefore (by the WIENER–KINTCHIN-theorem) we can use spectral densities. Considering $\varkappa(t)$ as the *input-process* and $A(t)$ as the *output* of a linear system one gets quite generally [6]

for the spectral densities

$$S(\omega) = |H(i\omega)|^2 S_K(\omega) \tag{4.5}$$

where $H(i\omega)$ is callled the "system function" or "susceptibility". For the process (4.3) one gets

$$S_{AA}(\omega) = \frac{1}{(\omega^2-\omega_0^2)^2+4M^2\omega^2}\,\frac{c}{\pi} \tag{4.6}$$

and in view of the case

$$M^2-\omega_0^2 \equiv \beta^2 > 0\,, \tag{4.7}$$

FOURIER transformation yields the correlation function

$$R_{AA}(\tau) = k\,g^{-1}\,e^{-M|\tau|}\left(\cosh\beta\tau + \frac{M}{\beta}\sinh\beta|\tau|\right). \tag{4.8}$$

For the process (4.4) we get

$$R_A(\tau) = k\,g^{-1}\exp\{(-\omega_0^2/2M)|\tau|\}\,. \tag{4.9}$$

5. Prediction Theory of Irreversible Processes

We shall now calculate the mean relaxation path and the entropy of the processes (4.3) and (4.4) by means of the prediction theory of stochastic processes. For this reason let us first demonstrate the general idea of this theory at a discrete stochastic process.

Let us consider the following situation. We know the n values [13]

$$a_{(1)}(t-1), a_{(1)}(t-2), \ldots, a_{(1)}(t-n) \tag{5.1}$$

of a realization of the stochastic series

$$A(1),\ A(2), \ldots, A(t), \ldots, A(t+\tau)\ldots. \tag{5.2}$$

and we are asked to estimate the value $a_{(1)}(t+\tau)$, [i. e. one is interested in the behaviour of the realization (5.1) τ steps ahead].

The best extrapolation, however, will be a function of the known values (5.1), i. e.

$$\hat{a}_{(1)}(t+\tau) = g(a_{(1)}(t-1), \ldots, a_{(1)}(t-n)) \text{ [13a]}. \tag{5.3}$$

As a criterion of efficiency of the estimation $\hat{a}_{(1)}(t+\tau)$ we may take

$$e_n(\tau) = E\{|A(t+\tau) - g(A(t-1), \ldots, A(t-n))|^2\}. \tag{5.4}$$

$e_n(\tau)$ is called the mean quadratic extrapolation error. In order to get the best estimation we have to minimize the extrapolation error. This is a problem which generally involves great analytical difficulties. Therefore one restricts the class of all possible functions g to the class of linear functions. That is to say we are working with a "wide sense theory" [10], since only covariances are taken into account. For GAUSSian processes which we are concerned with this restriction is of course of no effect.

The minimization of $e_n(\tau)$ is now very simple. Let us define a metric in the space of all random variables by

$$(X, Y) = E(x, y)$$

where (X, Y) is the scalar product of the random variables X and Y. Now construct a perpendicular from $A(t+\tau)$ on to the known random variables $A(t-1), \ldots, A(t-n)$. The square of the length of this perpendicular is then of course $e_n(\tau)$.

In formulas: choose $\eta_1, \ldots, \eta_n$ such that

$$E\{[A(t+\tau) - \{\eta_1 A(t-1) + \ldots + \eta_n A(t-n)\}]\,A(t-i)\} = 0 \qquad (i=1, \ldots, n) \tag{5.5}$$

then the smallest prediction error is

$$e_n(\tau) = E\{[A(t+\tau) - \{\eta_1 A(t-1) + \ldots + \eta_n A(t-n)\}]^2\}\,. \tag{5.6}$$

For the continuous case one has instead of (5.5)

$$E\left\{\left[A(t+\tau) - \int_0^\infty dt'\,A(t-t')\,h(t')\right]A(t'')\right\} = 0\,, \qquad t'' \leqq t \tag{5.7}$$

where a knowledge of the total past is assumed.

Confining to stationary processes Eq. (5.7) becomes

$$R(t+\tau-t'') = \int_0^\infty R(t-t''-t')\,h(t')\,dt' \tag{5.8}$$

[13] Up to Eq. (5.7) t is an integer.

[13a] There should be no confusion, with notations in earlier sections.

or, with $t-t''\equiv\tau'$,

$$R(\tau'+\tau)=\int_0^\infty R(\tau'-t')\,h(t')\,dt'. \tag{5.9}$$

This equation is called the WIENER–HOPF equation. If the stochastic process is known, i. e. if the correlation functions R are known, $h(t)$ can in principle be calculated and one easily gets for the prediction error

$$e(\tau)=R(0)-\int_0^\infty R(-\tau-t')\,h(t')\,dt' \tag{5.10}$$

and for the best prediction

$$\hat{A}(t+\tau)=\int_0^\infty dt'\,A(t-t')\,h(t'). \tag{5.11}$$

Let us now consider the processes (4.4) and (4.3). *The process* (4.4) is of a particular simple structure. Since its correlation function (4.9) is an exponential it belongs to the class of wider sense MARKOFFian processes [10].

These processes have the important property that for linear prediction only the knowledge of the presence is relevant, i. e.

$$g=\eta\,A(t)\ ^{13b}. \tag{5.12}$$

This implies, since

$$\eta=R_A(\tau)/R_A(0)=\exp\{(-\omega_0^2/2M)|\tau|\}, \tag{5.13}$$

$$\hat{a}(t+\tau)=\exp\{(-\omega_0^2/2M)|\tau|\}\,a(t). \tag{5.14}$$

Eq. (5.14) is an interesting and important result. Its importance results from the fact that the best prediction is equal to the expectation α of the conditional density $g(a,t)$ (in the PAULI limit):

$$\hat{a}(\tau)=\alpha(\tau). \tag{5.15}$$

This relation has as a consequence that (5.14) constitutes essentially the statement of the fluctuation dissipation theorem, i. e. the mean relaxation path $[\alpha(\tau)]$ is given by the equilibrium fluctuations (4.9).

The extrapolation error, for which one can show that it is identical with the dispersion of $g(a,t)$, is given by

$$e_A(\tau)=E\{[A(t+\tau)-\eta\,A(t)]\,A(t+\tau)\} \tag{5.16}$$
$$=R_A(0)-\frac{R_A^2(\tau)}{R_A(0)}=k\,g^{-1}(1-\exp\{(-\omega_0^2/M)|\tau|\}).$$

For the second order MARKOFF *process (4.3)*, which describes a memory effect, the prediction problem is not so easy. Here it is necessary to solve the WIENER–HOPF equation (5.9). One gets for the best extrapolation

$$g=\eta\,A(t)+\zeta\,\dot{A}(t) \tag{5.17}$$

where

$$\eta=R_{A\dot{A}}(\tau)/R_{A\dot{A}}(0) \quad\text{and}\quad \zeta=\dot{R}_{A\dot{A}}(\tau)/\ddot{R}_{A\dot{A}}(\tau). \tag{5.18}$$

Using (4.8) Eq. (5.17) becomes

$$\hat{a}(t+\tau)=e^{-M|\tau|}\left(\cosh\tau+\frac{M}{\beta}\sinh|\tau|\right)a(t)+\frac{1}{\beta}e^{-M|\tau|}\sinh\beta|\tau|\,\dot{a}(t). \tag{5.19}$$

This is an interesting result, since it plainly shows the memory character of the process (4.3). The term with $\dot{a}(t)$ improves the prediction.

For the prediction error one gets

$$e_{A\dot{A}}(\tau)=E\{[A(t+\tau)-\eta\,A(t)-\zeta\,\dot{A}(t)]\,A(t+\tau)\}=R_{AA}(0)-\eta\,R_{A\dot{A}}(\tau)+\zeta\,\dot{R}_{A\dot{A}}(\tau)$$
$$=k\,g^{-1}\left\{1-e^{-2M|\tau|}\left[\left(\cosh\beta\tau+\frac{M}{\beta}\sinh\beta|\tau|\right)^2+\frac{\omega_0^2}{\beta^2}\sinh^2\beta|\tau|\right]\right\}. \tag{5.20}$$

In the strong damping case ($M\gg\omega_0$) and for long times $e_{A\dot{A}}(\tau)$ becomes $e_A(\tau)$. Quite generally one has

$$\lim_{\tau\to\infty}e_{A\dot{A}}(\tau)=\lim_{\tau\to\infty}e_A(\tau)=R_A(0)=R_{AA}(0)=k\,g^{-1}. \tag{5.21}$$

6. Prediction Error and Entropy

The GIBBS entropy is given by

$$S=-k\int da\,g(a,t\,|\,a_0,0)\,\ln\frac{g(a,t\,|\,a_0,0)}{W(a)} \tag{6.1}$$

[13b] See footnote [13a].

where $g(a,t\,|\,a_0,0)$ denotes a conditional probability density and $W(a)=\lim_{t\to\infty} g(a,t\,|\,a_0,0)$ is the "equilibrium distribution" (i. e. the stationary probability density). In information theory $H=-(1/k)\,S$ is known as *mean conditional transinformation* and one writes

$$H=I(g(a,t\,|\,a_0,0)\,\|\,W(a)) \tag{6.2}$$

H may be interpreted as the information one possesses over a random variable $A(t)$, if one knows the value a_0 of a random variable $A(0)$.

Besides H there exists the *relative information* I_r which is defined by

$$\begin{aligned} I_r &= \int da_0\, W(a_0)\, I(g(a,t\,|\,a_0,0)\,\|\,W(a)) \\ &= \int da_0 \int da\, g(a,t,a_0,0)\,\ln\frac{g(a,t,a_0,0)}{W(a)\,W(a_0)} \end{aligned} \tag{6.3}$$

where $g(a,t,a_0,0)$ denotes the simultaneous probability density of the random variables $A(0)$ and $A(t)$. I_r may be interpreted as the mean information of one random variable relative to another.

Let us again consider the first order GAUSSIAN–MARKOFF process (4.4). One has

$$g(a,t\,|\,a_0,0)=\frac{1}{\sqrt{2\pi\,e_A(t)}}\exp\left\{-\frac{1}{2\,e_A(t)}\,(a-\alpha(t))^2\right\} \tag{6.4}$$

and

$$W(a)=\frac{1}{\sqrt{2\pi\,k\,g^{-1}}}\exp\{-a^2/2\,k\,g^{-1}\}\,. \tag{6.5}$$

Introducing these expressions into (6.1) one gets

$$\begin{aligned} S_A=&-\tfrac{1}{2}\,g\exp\{-2\,\omega_0^2\,t/M\}\;(\alpha_0^2-k\,g^{-1}) \\ &-\tfrac{1}{2}\,k\ln\frac{e_A(t)}{k\,g^{-1}} \end{aligned} \tag{6.6}$$

where α_0 denotes a macroscopic deflection from equilibrium, i. e. $\alpha_0 \gg k\,g^{-1}$.

Eq. (6.6) consists essentially of two terms. The first is determined by the best extrapolation and the second by the prediction error $e_A(t)$. It is interesting to notice that this last term is except for the factor $-k$ the relative information [7]

$$S_{rA}\equiv -k\,I_{rA}=-\tfrac{1}{2}\,k\ln\frac{e_A(t)}{k\,g^{-1}}\,. \tag{6.7}$$

Introducing Eq. (5.16) one has

$$S_{rA}=-\tfrac{1}{2}\,k\ln(1-\exp\{-\omega_0^2\,t/M\}) \tag{6.8}$$

and it is well known [5] that this term can be neglected in (6.6). Thus the entropy of the process is entirely given by the best extrapolation.

$$\begin{aligned} S_A&=-\tfrac{1}{2}\,g\,\alpha_0^2\exp\{-2\,\omega_0^2\,t/M\} \\ &=-\tfrac{1}{2}\,k\,\frac{k\,g^{-1}}{\alpha_0^2}\exp\{-2\,\omega_0^2\,t/M\} \end{aligned} \tag{6.9}$$

or, using (5.14) with (5.15)

$$S_A=-\tfrac{1}{2}\,g\,\alpha^2(t)\,. \tag{6.10}$$

7. Summary and Discussion

The starting point of the present investigations was the LIOUVILLE equation of ordinary phase space [which lead to (2.3)]. The ZWANZIG formalism then allowed the derivation of a generalized MASTER equation in a space. This equation can be interpreted as a differential CHAPMAN–KOLMOGOROFF equation of a non-MARKOFFian process. In order to get the fundamental equations of a generalized fluctuation theory from phenomenological phase space theory three assumptions have to be introduced

1. $O\{(LP)^2\} \gg O\{(LP)^3\}$
   [Eq. (2.8) implying (2.9)].
2. Extension of a space
   [Eq. (3.1), in order to remove the time convolution of (2.9)].
3. Linearization
   [Eqs. (3.8) and (3.9) implying (3.11)].

Physically these conditions should be fulfilled for slow relaxation processes near equilibrium.

The basic equations of fluctuation theory have then been studied using the prediction theory of stationary stochastic processes. In this theory the minimization of the prediction error replaces ONSAGER's variation principle. It should be mentioned, however, that only the "extrapolation aspect" of ONSAGER's theory has been investigated. But, of course, also the other problems studied by ONSAGER and MACHLUP [1] and TISZA and MANNING [14] can be investigated by the help of interpolation and filtering theory and other well established methods of system theory.

Acknowledgement

The author wishes to express his indebtedness to Professor Dr. H. HARTMANN for valuable support of this paper.

[14] L. TISZA and I. MANNING, Phys. Rev. **105**, 1695 [1957].