Van Kampen, N. G.
1957

Physica XXIII
707–719

DERIVATION OF THE PHENOMENOLOGICAL EQUATIONS FROM THE MASTER EQUATION

I. EVEN VARIABLES ONLY.

by N. G. VAN KAMPEN

Instituut voor theoretische fysica, Rijksuniversiteit te Utrecht, Nederland

Synopsis

The 'master equation' describes the behaviour of a macroscopic system in terms of a time dependent probability distribution. It is here shown that, if the initial distribution is concentrated in a small region, it moves toward equilibrium without spreading. Thus the stochastic process described by the master equation is observed as a deterministic process by an observer whose observations are too coarse to observe the fluctuations. This is the process to which the usual phenomenological equations refer. With the aid of appropriate approximations one finds in this way the well-known linear regression equations, including Onsager's reciprocal relations.

1. *Introduction.* Thermodynamics of irreversible processes is concerned with the phenomenological laws that govern the behaviour of macroscopic systems in non-equilibrium states. Statistical mechanics of irreversible processes attempts to explain these laws as a consequence of the mechanics of a very large number of microscopic particles. In classical mechanics the $3N$ coordinates and $3N$ momenta of these particles are fully determined by the equations of motion, as soon as their initial values are known. This microscopic determinism, however, does not show up in experiments, because the $6N$ microscopic variables cannot be observed separately. In quantum mechanics, the motion of the particles is described by the Schrödinger equation for the wave function ψ; this equation determines ψ at any time, as soon as it is known at an initial time. In this case there are two reasons why this determinism does not appear in experiments. The first reason is again that the large number of particles prevents anything like a complete knowledge of the initial state. The second reason is that, even for small systems, ψ is not directly connected with observable quantities, but only in terms of the standard probability interpretation of quantum mechanics.

Yet the phenomenological laws concerning large systems are fully deterministic: it is empirically known, that a set of macroscopically observable variables can be selected in such a way that, if their values are known at

$t = 0$, they are uniquely determined at later times. Indeed, this fact is the basis of all phenomenological theories of heat conduction, diffusion, etc. The purpose of this paper is to study the connection between this deterministic behaviour of macroscopic phenomena and the deterministic character of the mechanical laws for the microscopic particles. That this connection is by no means trivial is of course well-known, but may be further emphasized by two remarks. Firstly the macroscopic determinism only applies to the future, not to the past, whereas the mechanical laws are deterministic in both time directions. Secondly, the macroscopic determinism breaks down for phenomena on a very small scale, where the separation between macroscopic and microscopic tends to disappear, as for instance the movement of Brownian particles. (In these cases it can be restored by considering averages over a large collection of similar phenomena, because these averages are again truly macroscopic quantities.)

The passage from the equations of motion for the microscopic particles, to the phenomenological laws for macroscopic systems, breaks up into two steps. The first step [1] [2] leads from the microscopic equations to the so-called master equation

$$\mathrm{d}P_J/\mathrm{d}t = \sum_{J'}(W_{JJ'}P_{J'} - W_{J'J}P_J). \tag{1}$$

The subscript J labels the phase cells, which are defined as those regions of phase space in which the macroscopic variables $A^{(r)}$ have given values $A_J^{(r)}$ with a possible error $\Delta A^{(r)}$. The (non-negative) numbers P_J are the probabilities of finding the system in the corresponding phase cells. Of course, one must have

$$\sum_J P_J = 1, \tag{2}$$

which is indeed easily seen to be compatible with (1). The average or expectation value of each $A^{(r)}$ is

$$\langle A^{(r)} \rangle = \sum_J A_J^{(r)} P_J. \tag{3}$$

The transition probabilities $W_{JJ'}$ (from cell J' into J) are also non-negative and satisfy the symmetry relation

$$W_{JJ'}G_{J'} = W_{J'J}\, G_J, \tag{4}$$

where G_J denotes the phase volume of cell J. Actually, the symmetry relation only takes the simple form (4) if all the macroscopic variables $A^{(r)}$ are even functions of the particle velocities, and if the same is true of the Hamiltonian of the total system; that will be assumed throughout this paper. The symmetry property (4) is also called 'principle of detailed balancing' *).

*) Originally it was called 'microscopic reversibility' [3], but this name is somewhat confusing, because (4) does not involve microscopic quantities, nor is it identical with – although a consequence of – the time reversibility of the microscopic equations of motion.

Although (1) fully determines the P_J from their initial values, it does not furnish a deterministic description of the system, because the P_J are only probabilities. The second step consists therefore in deriving phenomenologically deterministic equations from (1), *i.e.*, equations which determine the future values of actually observed quantities from their initial values. This is the object of the present paper. It clearly applies to both classical and quantum statistics, because these have already merged in the master equation (1).

In section 2 the rather elementary step is taken of replacing the discrete label J by a set of continuous coordinates a_r. This transforms (1) into the same integral equation that was recently used by de Groot and Mazur [4]). Next, in section 3, this integral equation is turned into a Fokker-Planck differential equation. This equation was first used to describe the 'diffusion in a-space' by M. S. Green [5]). The application in the present paper, however, is much simpler, since we can start from the master equation rather than from the microscopic equations themselves. In section 3 it is shown that by inserting for the unspecified functions in the Fokker-Planck equation plausible approximations, one finds immediately the usual linear form of the phenomenological equations.

In the case of a Brownian particle the stochastic character of the motion is caused by a randomly fluctuating external force on the particle. Some authors [6]) have argued that even for large systems some external force is responsible for the stochastic behaviour. This, of course, is incorrect: the stochastic nature comes in through the coarseness of the macroscopic observations. Yet it is possible to employ, like Onsager and Machlup [7]), a random force term, as an artifice to obtain a stochastic description of the behaviour of the system without going into the details of the microscopic processes that are responsible for the randomness. It is shown in section 4 that this artifice leads indeed to correct results, provided the necessary stochastic properties of the random force term are assumed.

The symmetry property (4) of the transition probabilities gives rise to certain relations among the coefficients in the Fokker-Planck equation. These relations are shown in section 5 to be generalisations of the well-known Einstein relation for Brownian motion. Finally, in section 6 the various assumptions and approximations in the preceding work are discussed.

2. *Definition of a-space.* In each phase cell J the values of the whole set of variables $A^{(r)}$ are fixed. These values may be represented by a point in 'a-space', the coordinates a_r corresponding to the values $A_J^{(r)}$. Actually this a-space consists of discrete points, each of which corresponds to one phase cell. The differences between the coordinates of neighbouring points are Δa_1, Δa_2, ...; they are identical to the sizes $\Delta A^{(1)}$, $\Delta A^{(2)}$, ... of the phase cells. However, provided only smoothly varying functions of the a's are taken

into consideration, the a's may be regarded as continuously varying. One then gets the following transcription of the quantities mentioned in the introduction.

$$P_J \to P(a_1, a_2, \ldots)\, \Delta a_1\, \Delta a_2 \ldots = P(a)\, \Delta a;$$
$$W_{JJ'} \to W(a,a')\, \Delta a;$$
$$G_J \to G(a)\, \Delta a.$$

Equations (2) and (3) take the form

$$\int P(a)\, \mathrm{d}a = 1$$
$$\langle A^{(r)} \rangle = \int a_r\, P(a)\, \mathrm{d}a.$$

More generally one has, for an arbitrary smooth function $f(a_1, a_2, \ldots) \equiv f(a)$,

$$\langle f(A^{(1)}, A^{(2)}, \ldots) \rangle = \int f(a_1, a_2, \ldots)\, P(a)\, \mathrm{d}a. \tag{5}$$

Equations (1) and (4) become

$$\dot{P}(a) = \int W(a, a')\, P(a')\, \mathrm{d}a' - \int W(a', a)\, \mathrm{d}a' \,.\, P(a), \tag{6}$$
$$W(a, a')\, G(a') = W(a', a)\, G(a). \tag{7}$$

In order to fix the origin in a-space, it should first be noted that the total energy need not be included among the a_r, although it is certainly a macroscopic variable. For, as no transitions between energy shells are possible, one may confine the whole consideration to one shell, so that the energy value only appears as a constant in the equations. Thus there are a number of disconnected a-spaces, one for each energy shell. Now, for each fixed energy value the function $G(a)$ has a maximum. This will be taken as origin in the a-space belonging to this energy value.

This substitution of a set of variables a_r for the single label J is not just a matter of convenience. In doing so, one arranges the phase cells according to the values of the macroscopic variables. As a result, it is not unreasonable to take for $P(a)$ a more or less smoothly varying function, whereas smoothness of P_J with respect to J would be meaningless (since it depends on the arbitrary choice of the label J). In fact, if at some instant $P(a)$ were not smooth, then it would only need a short transient period to smooth out through the transitions between cells. The question why $W(a, a')$ may be expected to be smooth (in an appropriate sense, see the first paragraph of the next section) is much harder, and will not be discussed here. That $G(a)$ varies smoothly from one phase cell in the next – although steeply varying on a larger scale – is of course known from ordinary statistical mechanics of equilibrium states.

This a-space or 'phenomenological space' or 'macrophase space' has been used by M. S. Green [5]), Hashitsume [8]), Yamamoto [9]) and De Groot and Mazur [4]).

3. *Differential form of the master equation.* Let it be supposed that the transition probability $W(a, a')$ is only appreciable for small values of $y_r = a_r - a_r'$; that is, if the distance between the starting point and the end point of the transition is small. It is convenient to write $W(a'; y)$, so that this function is peaked with respect to y, but may still be smooth with respect to the first argument a'. One may then write for (1) the corresponding multivariate Fokker-Planck equation [10])

$$\dot{P} = \partial_r \partial_s (\xi_{rs} P) + \partial_r (\eta_r P), \tag{8}$$

where

$$\eta_r(a) = \int W(a; y)\, y_r \, \mathrm{d}y,$$

$$\xi_{rs}(a) = \tfrac{1}{2} \int W(a; y) y_r y_s \, \mathrm{d}y. \tag{9}$$

(∂_r stands for $\partial/\partial a_r$ and summation over double indices is implied). It will be suitable to write (8) in the slightly different form

$$\dot{P} = \partial_r \{\xi_{rs} G \, \partial_s (P/G)\} + \partial_r (\tilde{\eta}_r P), \tag{10}$$

where $\tilde{\eta}_r$ is connected with η_r, ξ_{rs} and G by

$$\tilde{\eta}_r = \eta_r + G^{-1} \, \partial_s \, \xi_{rs} G. \tag{11}$$

It should be noted that ξ_{rs} in (10) is the same as in (8), so that it is still a symmetric tensor:

$$\xi_{rs}(a) = \xi_{sr}(a). \tag{12}$$

In order to find the properties of ξ_{rs} and $\tilde{\eta}_r$ that correspond to the symmetry relation (7), we employ an artifice, which was previously used in a slightly different connection [11]). It is easily checked that a necessary and sufficient condition for (7) is that for any two solutions $P(a, t)$ and $P^*(a, t)$ of (6) the identity

$$\int P^* \dot{P} \, \mathrm{d}a/G - \int \dot{P}^* P \, \mathrm{d}a/G = 0 \tag{13}$$

is satisfied. This formulation of the symmetry property can readily be applied to (10). Substitute for $\dot{P}$ and $\dot{P}^*$ in (13) their values according to (10). By partial integration both terms involving ξ_{rs} cancel, so that one is left with

$$\int \{(P^*/G) \, \partial_r (\tilde{\eta} P) - (P/G) \, \partial_r (\tilde{\eta} P^*)\} \, \mathrm{d}a = 0.$$

As this must hold for every function P^*,

$$(1/G) \, \partial_r (\tilde{\eta} P) + \tilde{\eta}_r \, \partial_r (P/G) = 0,$$

or

$$(2\tilde{\eta}_r/G) \, \partial_r P + (\ldots) \, P = 0.$$

As this must hold for every function P, the coefficient of each derivative $\partial_r P$ must vanish. Hence $\tilde{\eta}_r = 0$, so that (10) reduces to

$$\dot{P} = \partial_r \, \xi_{rs} \, G \, \partial_s (P/G). \tag{14}$$

From this one finds for the expectation value of a function $f(A^{(1)}, A^{(2)}, \ldots)$ of the macroscopic observables, using partial integration,

$$\begin{aligned}(\mathrm{d}/\mathrm{d}t)\,\langle f(A)\rangle &= \int f(a)\,\dot{P}(a,t)\,\mathrm{d}a \\ &= \int (P/G)\,\partial_s\,\xi_{rs}\,G\,\partial_r\,f(a)\,\mathrm{d}a \\ &= \langle G^{-1}\,\partial_s\,\xi_{rs}\,G\,\partial_r\,f\rangle. \end{aligned} \tag{15}$$

This result is simpler than that of M. S. Green [5]) and Hashitsume [8]) owing to the application of the symmetry property. In the next section (15) will be reduced by the use of certain approximations.

Equation (14) may be visualised as a 'diffusion in a-space'. $P(a)$ is then to be regarded as a density of a-points, satisfying the generalised diffusion equation (14). The quantities one is physically interested in are expectation values of the form $\langle f(A^{(1)}, A^{(2)}, \ldots)\rangle$; according to (5) this is just the average of $f(a_1, a_2, \ldots)$ with the density $P(a)$. It is therefore convenient to use as an alternative notation $\langle f(a_1, a_2, \ldots)\rangle$. We shall be particularly interested in averages $\langle a_r\rangle$, and the second moments $\langle a_r a_s\rangle$.

Alternatively, (14) may be visualised as an equation for 'heat conduction in a-space', which has the advantage that the weight factor G can be more easily be brought into the picture. Indeed, if P is the heat content per unit volume in a-space and G the specific heat, then P/G is the temperature and $\xi_{rs}G$ the conduction tensor.

In the case of two variables a_1, a_2, a third, even more concrete, picture can be made. Take a thin sheet in the (a_1, a_2)-plane whose thickness is $G(a_1, a_2)$. If the specific heat is now constant, and if one takes for the conduction tensor ξ_{rs}, the flow of heat in the sheet is again described by (14). This picture can, of course, also be formulated in the language of diffusion.

3. *Linear approximation.* According to the usual approximation [12]) one may put in some neighbourhood of the origin in a-space

$$G(a) = G(0)\,e^{-\frac{1}{2}g_{rs}a_r a_s}. \tag{16}$$

(The ultimate ground for the validity of this formula is the the law of great numbers: the value of each $A^{(r)}$ is made up by a great number of microscopic variables.) Moreover, it is reasonable to replace $\xi_{rs}(a)$ by $\xi_{rs}(0)$, if it is a slowly varying function of a. We call this the linear approximation, because it entails linearity of the phenomenological equations.

In this approximation (15) becomes

$$\partial_t\,\langle f(a)\rangle = \xi_{rs}\,\langle \partial_r\partial_s f\rangle - \xi_{rs}\,g_{sp}\,\langle a_p \partial_r f\rangle. \tag{17}$$

First take $f(a) \equiv a_m$:

$$\partial_t\,\langle a_m\rangle = -\,\xi_{ms}g_{sp}\,\langle a_p\rangle. \tag{18}$$

These equations have the apperance of the well-known phenomenological regression equations, which are usually written in the form [12])

$$\dot{\alpha}_m = L_{ms}X_s = -L_{ms}g_{sp}\alpha_p. \quad . \tag{19}$$

Thus the coefficients $\xi_{ms} = \xi_{ms}(0)$ are nothing but the usual L_{ms}. Onsager's reciprocal relations [13]) [12]) are automatically satisfied owing to (12).

Yet the equations (18) *are not identical with* (19), because they involve averages, whereas phenomenological equations – like the laws of Fourier and Fick – are formulated in terms of measurable quantities with sharply defined unique values. Of course, their values are only sharply defined as long as no precision is required of the order of magnitude of the fluctuations. The phenomenological equations (19) tacitly presuppose that the fluctuations remain small during the approach toward equilibrium and do not grow at the same rate as the α's decrease. This is not implied in (18), because (18) does not exclude the possibility that the averages satisfy these equations, but that at the same time $P(a)$ spreads out over such a broad region in a-space, that the averages occurring in (18) have no practical significance. We shall now show that this contingency is *not* materialised.

Taking $f(a) = a_m a_n$ in (17),

$$\partial_t \langle a_m a_n \rangle = 2\xi_{mn} - \xi_{ms}\, g_{sp} \langle a_p a_n \rangle - \xi_{ns}\, g_{sp} \langle a_p a_m \rangle. \tag{20}$$

This set of linear differential equations for the $\langle a_m a_n \rangle$ can readily be solved. Here only the case of one variable is treated explicitly, because that suffices to demonstrate why $P(a)$ does not spread out. Besides, the case of more variables can be reduced to it by transforming the a's linearly in such a way that both g_{rs} and ξ_{rs} reduce to diagonal matrices.

For one variable a, (18) reads

$$\partial_t \langle a \rangle = -\xi g \langle a \rangle,$$

with the solution

$$\langle a \rangle_t = \langle a \rangle_0\, e^{-\xi g t}. \tag{21}$$

This result exhibits the usual exponential decay, with relaxation time $1/\xi g$. Similarly (20) reduces to

$$\partial_t \langle a^2 \rangle = 2\xi - 2\xi g \langle a^2 \rangle,$$

with the solution

$$\langle a^2 \rangle_t = \langle a^2 \rangle_0\, e^{-2\xi g t} + g^{-1}\,(1 - e^{-2\xi g t}). \tag{22}$$

For the mean square deviation $\sigma^2(t) \equiv \langle a^2 \rangle_t - \langle a \rangle_t^2$ one finds

$$\sigma^2(t) = \sigma^2(0)\, e^{-2\xi g t} + g^{-1}\,(1 - e^{-2\xi g t}).$$

The first term on the right shows that the mean square deviation present in the initial situation decreases exponentially at the same rate as $\langle a \rangle$ itself.

The second term on the right shows an exponential approach to the final value g^{-1}, which is the equilibrium value known from ordinary statistical mechanics.

Now *for an observer whose measurements are too coarse to observe these fluctuations, the a_r will appear as sharply defined quantities*, each of which may be denoted by a single symbol, say α_r. These α_r are defined with an indeterminacy or 'margin' of the size of the fluctuations, and satisfy with this same margin such equations as $\alpha_r = \langle a_r \rangle$, $\alpha_r^2 = \langle a_r^2 \rangle$, $\alpha_r \alpha_s = \langle a_r a_s \rangle$. Also, with the same margin they obey the deterministic equation (19).

4. *Comparison with Brownian motion.* The motion of a Brownian particle satisfies of course Newton's equation of motion

$$\dot{v} = K, \tag{23}$$

where the mass is taken unity and K is the force exerted by the surrounding medium. One then *assumes* [14])

$$K = -\beta v + \kappa(t), \tag{24}$$

where $\kappa(t)$ is a rapidly fluctuating force with zero time average. In order to deal with such a fluctuating force it is convenient to introduce an ensemble of Brownian particles, and to study only ensemble averages (to be denoted by a bar). The statistical properties of $\kappa(t)$ are futher specified by assuming

$$\overline{\kappa(t)\,\kappa(t')} = C\,\delta(t - t'), \tag{25}$$

with a certain constant C. On solving the Langevin equation

$$\dot{v} = -\beta v + \kappa(t), \tag{26}$$

one then obtains without difficulty [10])

$$\overline{v(t)} = \overline{v(0)}\, e^{-\beta t},$$

$$\overline{v(t)^2} = \overline{v(0)^2}\, e^{-2\beta t} + (C/2\beta)(1 - e^{-2\beta t}).$$

These equations are identical with (21) and (22) if one puts

$$\beta = \xi g, \quad C = 2\xi. \tag{27}$$

The fact that these results are identical suggests the following alternative procedure for obtaining the results of section 3. One imagines a fictitious system that mimics the real one inasmuch as its state is described by the same set of variables a_r. However, it must be borne in mind that for the real system the a's only served as coordinates in a space, in which a diffusion process took place governed by (14); for the fictitious system the a_r are to be regarded as rigorously defined variables satisfying an equation of motion.

This equation of motion is chosen to be identical to (19) but for a set of additional fluctuating 'random forces' $\kappa_m(t)$:

$$\dot{a}_m = - L_{ms}\, g_{sp}\, a_p + \kappa_m(t).$$

Subsequently an averaging process is introduced to deal with the random character of the κ_m, and appropriate statistical properties are prescribed for them. It is readily seen that in order to obtain (21) and (22) one must prescribe

$$\overline{\kappa_m(t)} = 0, \tag{28}$$

$$\overline{\kappa_m(t)\, \kappa_n(t')} = L_{mn}\, \delta(t - t'). \tag{29}$$

The higher moments of the $\kappa_r(t)$ can likewise be obtained from a comparison with (17); for the case of one variable one finds for instance

$$\overline{\kappa(t)\, \kappa(t')\, \kappa(t'')\, \kappa(t''')} = L^2\, [\delta(t - t')\, \delta(t'' - t''') + \delta(t - t'')\, \delta(t' - t''') + \delta(t - t''')\, \delta(t' - t'')]. \tag{30}$$

It should be borne in mind that in this case the κ_m are no actual physical forces.

This was the procedure employed by Hashitsume [8]) and by Onsager and Machlup [7]) in order to obtain a stochastic process for which (18) holds. Apparently this description is equivalent with the general equation (17); hence it is valid even for quantummechanical systems, provided it is regarded only as a formal way to describe a Markov process. The limitations of this procedure are firstly, that the phenomenological coefficients L_{rs} have to be introduced *ad hoc*, without any connection with the underlying microscopic mechanism; and secondly that the statistical properties of the $\kappa_m(t)$ have to be guessed.

Actually the possibilities for the $\kappa_m(t)$ are severely restricted by the general features that are postulated. In the first place one wants the resulting phenomenological equations to be linear; this entails (28). Secondly the δ-function in (29) is necessitated by the requirement that the stochastic process be Markovian. Thirdly, the precise form of the coefficient (L_{mn}) in front of the delta function is determined by the condition that the fluctuations must approach their equilibrium values, which are known from the ordinary statistical mechanics of equilibrium states. (This amounts to an Einstein relation, as explained in the next section.) The higher moments of the $\kappa_m(t)$, for example (30), are also determined by the Markovian character. This is the way in which one usually arrives at a characterisation of the stochastic properties of the force $\kappa(t)$ in the theory of Brownian motion [10]).

5. *The Einstein relation.* Langevin's treatment of the Brownian motion involves two constants, β and C, which are connected by (27) with the con-

stants ξ and g occurring in the microscopic treatment. Now ξ is actually connected with the irreversible process, whereas g belongs to the statistical mechanics of equilibrium processes, and may therefore be considered known. Consequently there is a relation between β and C, *viz.*,

$$C/\beta = 2/g = 2\langle v^2\rangle_{\text{eq}} = 2kT. \tag{31}$$

Thus our knowledge of the equilibrium yields a connection between the two constants β and C, which occur in the first and second moments of the fluctuating force K. This connection is closely related to Einstein's equation; we shall therefore refer to (31), and other equations of the same type, as 'Einstein relations' *).

It is of interest to note that this connection between first and second moments is essentially due to the symmetry property (7). Indeed, from (8) would follow

$$\partial_t \langle a_m\rangle = -\langle \eta_m(a)\rangle,$$

$$\partial_t \langle a_m a_n\rangle = \langle \xi_{mn}(a)\rangle - \langle a_m \eta_n\rangle - \langle a_n \eta_m\rangle.$$

Clearly, if the η's and the ξ's were unrelated there would be no relationship between these two lines. However, the symmetry property tells that (11) must vanish, so that, in linear approximation,

$$\eta_r(a) = -G^{-1}\,\partial_s\,\xi_{rs}\,G = \xi_{rs}(0)\,g_{sp}\,a_p.$$

Thus the symmetry of the transition probabilities permits to express the η_r in terms of the ξ_{rs}, by means of relations involving the g_{rs}.

Equation (32) also shows that the symmetry property is responsible for the identity of the L_{rs} with the ξ_{rs}. This identity was used in section 3 to prove the symmetry of the matrix L_{rs}. It can also be used to show that the matrix L_{rs} is positive definite, *i.e.*, that

$$L_{rs}\lambda_r\lambda_s \geqslant 0$$

for any set of real numbers λ_r. (This must be true in order that the regression equations (19) cannot give rise to a decrease in entropy.) Indeed, one has with the aid of (9)

$$\begin{aligned} L_{rs}\lambda_r\lambda_s &= \int W(0;y)\,\lambda_r y_r \lambda_s y_s\,\mathrm{d}y \\ &= \int W(0;y)\,\{\lambda_r y_r\}^2\,\mathrm{d}y \geqslant 0. \end{aligned}$$

*) The actual equation of Einstein [15]) expresses the mean square *displacement* of the particle in terms of C. However, at the bottom of his equation lies (31); indeed, his equation can readily be derived from (26) when (31) is used. It may also be noted that every derivation of Einstein's equation [10]) is based on a comparison between the stochastic process and the known properties of equilibrium states. Nyquist's equation [16]) is another example of what we generally call 'Einstein relations'.

6. *Discussion.* In the previous sections it has been shown that the application of the Fokker-Planck approximation to the integral equations (6) leads to the conclusion that the macroscopic observables obey a deterministic set of equations; at least for an observer whose measurements are sufficiently inaccurate to disregard the fluctuations. In terms of the diffusion picture this means that the cloud of a-points does not spread out, but remains condensed while moving toward the origin. This somewhat unfamiliar behaviour for a diffusion process is caused by the weight factor $G(a)$. This factor increases so strongly in the direction toward the origin that each a-point has a much greater probability to move in this direction than in any other.

In order to make this argument stick, however, it is necessary to make plausible that the gradient of the $\xi_{rs}(a)$ in a-space is negligible compared to that of $G(a)$. For convenience the following discussion is based on classical mechanics. It has to be remembered that $W_{JJ'}$ is the average probability per unit time for a transition *from one particular* unit volume in cell J' *to anywhere* inside cell J (see section 7 of reference [2]). Now the definition (9) of $\xi_{rs}(a)$ may also be written as

$$\xi_{rs}(J) = \tfrac{1}{2} \sum_{J'} W_{J'J}(A_{J'}^{(r)} - A_{J}^{(r)})\,(A_{J'}^{(s)} - A_{J}^{(s)}). \tag{33}$$

Hence $\xi_{rs}(J)$ describes how an ensemble that is initially concentrated in one unit volume expands into an ellipsoidal cloud in a-space. The size and shape of this cloud depend, of course, somewhat on the position of the initial unit volume, but (33) does not involve the steeply increasing function G. It is therefore not unreasonable to assume that the $\xi_{rs}(a)$ vary much more slowly than $G(a)$. Of course the final justification can only be found in an actual computation of $\xi_{rs}(a)$ and $G(a)$ for each specific system one wants to study *). The replacement of $\xi_{rs}(a)$ by the constant $\xi_{rs}(0)$ is based on the same argument.

Next the use of the Fokker-Planck approximation must be justified. In order that the integral equation (6) can be replaced by the differential equation (8), it is necessary that $P(a)$ varies smoothly. More precisely, it must be possible to approximate $P(a') = P(a + y)$ by the first three terms of its expansion in y, in that range of y in which $W(a; y)$ differs appreciably from zero. On the other hand, in order that the fluctuations shall not be excessive, $P(a)$ must be confined to a narrow interval (of the order $g^{-\frac{1}{2}}$). Hence it is necessary that the peak of $W(a; y)$ is so narrow that $W(a; y) \approx 0$ unless $y \ll g^{-\frac{1}{2}}$.

On the other hand, this peak must be broad enough to cover several phase

*) A familiar application of this assumption occurs in the theory of beta-decay. The shape of the energy distribution of the emitted electron is determined from the known weight function G, while the unknown transition probability is taken constant.

cells. This is not only required for the use of continuously varying coordinates (section 2), but is already implied in the assumptions that had to be made to arrive at the master equation (1) [1] [2]). There a short time τ was introduced, such that the number of transitions from cell J' to J equals $\tau W_{JJ'}$. Moreover, the phase points arriving in J from J' must be more ot less equally distributed through J, in order that, in calculating the transitions during the next time interval τ, they may be regarded as a homogeneous distribution. This could not be the case if $W(a; y)$ would vary steeply within one cell.

Summarising the various conditions mentioned, one arrives at the following hierarchy of orders of magnitude *).

(i) The smallest quantity in a-space is the size of the phase cells, *i.e.* the grain of a-space.

(ii) Next there is the width of the peak of $W(a; y)$ as a function of y. This must be large compared to the size of the grain in order that the master equation shall hold.

(iii) The range covered by $P(a)$, that is, the size of the fluctuations. This must be large compared to the previous one in order that the Fokker-Planck approximation shall hold. Unless the fluctuations were excessive in the initial state, they will be of the order $g^{-\frac{1}{2}}$, that is the width of the peak of $G(a)$.

(iv) The inaccuracy of the observations. This must be larger than the size of the fluctuations, in order that all observed state variables have unique values satisfying deterministic equations of motion.

(v) The range over which the $\xi_{rs}(a)$ are almost constant and the approximation (16) for $G(a)$ holds. This must be large compared to the experimental inaccuracy, in order that the linear phenomenological equations (19) are valid. Of course, this is not a *sine qua non*, but the study of non-linear phenomenological equations will be left for a future paper.

Received 9-4-57.

*) In my original publication [1]) dealing with the derivation of (1), the size of the phase cells was supposed to be determined by the inaccuracy of the experimental observations. For the greater part of the work this was immaterial, but in the last section it was essential. (There it was supposed that the matrix $A_J^{(r)}$, by which according to (3) the expectation values of the $A^{(r)}$ are linearly expressed in the P_J, could be inverted. That would mean that the P_J themselves are macroscopically observable, which can only be true if the phase cells are not smaller than corresponds to experimental inaccuracy.) This point of view has to be abandoned, even if the validity of the Fokker-Planck approximation is not demanded. For the P_J must extend smoothly over several phase cells and yet be confined to a region in a-space of the order of the fluctuations, which is usually small compared to the experimental inaccuracy. The phase cells must still be sufficiently large to ensure the necessary randomness, but they must be small compared to the size of the fluctuations. Groenewold calls them 'sub-macro cells'. The derivation of the Onsager relations in the last section of reference [1]) should therefore be replaced by the work in the present paper.

REFERENCES

1) Van Kampen, N. G., Physica **20** (1954) 603.
2) Van Kampen, N. G., Fortschritte der Physik **4** (1956) 405.
3) Tolman, R. D., Proc. Nat. Acad. Sci. **11** (1925) 436.
4) De Groot, S. R., and Mazur, P., Physica **22** (1956).
5) Green, M. S., J. chem. Phys. **20** (1952) 1281.
6) E.g., Fowler, R. H., and Guggenheim, E. A., *Statistical thermodynamics* (Cambridge 1939).
7) Onsager, L., and Machlup, S., Phys. Rev. **91** (1953) 1505.
8) Hashitsume, H., Prog. theor. Phys. **8** (1952) 461.
9) Yamamoto, T., Prog. theor. Phys. **10** (1953) 11.
10) See e.g. Uhlenbeck, G. E., and Ornstein, L. S., Phys. Rev. **36** (1930) 823;
Wang, M. C. and Uhlenbeck, G. E., Rev. mod. Phys. **17** (1945) 323.
11) De Groot, S. R. and Van Kampen, N. G., Physica **21** (1955) 39; see also: Jaglom, A. M., Abhandlungen aus der Sowjetischen Physik **3** (Berlin 1953) 7.
12) De Groot, S. R. *Thermodynamics of irreversible processes* (Amsterdam 1951).
13) Onsager. L., Phys. Rev. **37** (1931) 405; **38** (1931) 2265;
Casimir, H. B. G., Rev. mod. Phys. **17** (1945) 343.
14) Kramers, H. A., Physica **7** (1940) 284.
15) Einstein, A., Ann. Physik (4) **17** (1905) 549.
16) Nyquist, H., Phys. Rev. **32** (1928) 110.