Contents

1. Introduction

The usual phenomenological laws of matter, like equations of state and transport equations, are *average* laws. They deal in macroscopic variables like pressure, temperature, heat flux and electrical current, which represent the aggregate effect of myriad molecular interactions. Likewise, the physical parameters appearing in the laws (compressibility, friction constant) characterize the behaviour of materials on a macroscopic scale, and represent molecular quantities only implicitly, as specialized indices of their overall, collective effect.

The equations themselves derive from the underlying laws of molecular motion by averaging procedures applied to the Liouville equation, which governs the dynamical evolution of the complete set of molecular coordinates (though few such derivations have actually been carried out, and none of the laws were found in this way). The idea of fluctuation theory is to refine the description such phenomenological laws provide by treating them literally as average equations, and seeking probabilistic laws for the macroscopic variables. From this point of view the usual macroscopic law only describes how the mean values behave, and one seeks the probability distribution of actual values around the mean.

The motion of a free particle of mass M and (small) velocity v in a fluid, for example, is well described by the linear friction law

$$M\frac{\mathrm{d}v}{\mathrm{d}t} + \gamma v = 0. \tag{1}$$

The particle slows down to rest exponentially, with a damping constant γ/M. Since the mass, and friction constant γ, are measurable macroscopic parameters, this provides a complete, autonomous description of the situation, without reference to any molecular quantities. Equation (1) can be obtained from hydrodynamics (the Navier – Stokes equations), which also gives the friction constant in terms of properties of the fluid and particle: $\gamma = 6\pi\eta a$ for a sphere of radius a in a medium of ordinary fluid viscosity η (with eq. (1) this is "Stokes' formula").

Hydrodynamics, in turn, is derivable from a molecular theory in the case

of dilute gases, our main subject. This theory (the Chapman - Enskog procedure applied to the Boltzmann equation) also gives the fluid viscosity as a function of state variables, in terms of atomic masses and scattering cross sections. The Boltzmann equation is derivable from the Liouville equation, so the connection with microscopic dynamics appears complete:

Liouville equation
↓
Boltzmann equation
↓
Navier – Stokes equations
↓
Stokes' formula.

But eq. (1) describes irreversible behavior (the particle slows down, its kinetic energy is transferred to the fluid), whereas the Liouville equation is completely time-symmetric (the time-reversible laws of mechanics govern the individual molecules). The H-theorem establishes that the irreversible character is possessed by the Boltzmann equation, so that some "non-mechanical" element must have already entered here. This was of course the content of Loschmidt's (1876) objection, which eventually led to the understanding that Boltzmann's equation is, in effect, an *average* law in just the sense hydrodynamics or Ohm's law is. The molecular distribution function $f(\boldsymbol{r}\boldsymbol{v};t)\mathrm{d}\boldsymbol{r}\mathrm{d}\boldsymbol{v}$ appearing in it just represents the average number of particles near $\boldsymbol{r}$ with velocity near $\boldsymbol{v}$; the *actual* number must vary around this value. So the Boltzmann equation itself has to be seen as a kind of phenomenological law, a proper subject in its own right for fluctuation theory.

One argues for Stokes' formula, eq. (1), as follows. Even in a time interval small on the macroscopic scale, enormously many collisions will occur between the fluid molecules and the particle. Because the molecular motions are capricious their effects would cancel out completely if the particle were at rest. In motion, it should suffer more change of momentum in its (momentary) direction of travel since there will be more collisions from the front. For small velocities the net effect is the linear friction law, eq. (1).

Langevin's (1908) theory of Brownian motion begins by reinterpreting eq. (1) explicitly as an *average* equation, and extends it by adding a random "fluctuating force" to represent that portion of the effect of molecular collisions not already accounted for in the linear friction term. The fluctuating term appearing in the Langevin equation

$$M\frac{\mathrm{d}v}{\mathrm{d}t} + \gamma v = \tilde{e}(t) \tag{2}$$

is not a function at all but rather represents an ensemble of possible instantaneous values, according to a definite statistical law (see sect. 5), and has mean zero by definition. Having established or guessed the properties of the fluctuating force, one can then calculate all the properties of the particle motion, which is now conceived of as a stochastic process. In particular the ensemble-average velocities follow Stokes' formula.

So one separates, conceptually, a "systematic" effect of molecular interactions and the remaining "random" part, and represents the latter explicitly by a stochastic process with certain properties. It is remarkable that such a scheme is possible, since the validity of phenomenological laws like eq. (1) – and even the possibility of description in one or a few macroscopic variables - depends on quite special relationships of magnitude, and makes heavy use of the limit $N \to \infty$.

Fluctuation phenomena, by contrast, arise precisely on account of the noncontinuous, molecular constitution of matter, and depend on the fact that the number of molecules N in a sample of bulk matter is actually finite. Fluctuations mediate between the macroscopic and microscopic worlds, yet it remains possible to describe them in completely macroscopic variables, merely by augmenting the original macroscopic laws.

At least that is the case for systems close to equilibrium, where linear laws apply. Cast in appropriate thermodynamical variables, such systems are seen to be merely one or multi-dimensional variants of Stokes' formula, so that the phenomenological theory of fluctuations is just Langevin's theory of Brownian motion. This idea, due to Onsager and Machlup (1953), Green (1952) and others (sect. 7) has been applied to the linearized Navier – Stokes equations of hydrodynamics by Landau and Lifshitz (1957) and to the linearized Boltzmann equation governing the molecular distribution function $f(\boldsymbol{r}\boldsymbol{v};t)$ close to equilibrium by Bixon and Zwanzig (1969), Fox and Uhlenbeck (1970), and others (sect. 8).

These phenomenological fluctuation theories can be connected in a hierarchical way. As the Stokes' formula can be obtained from linearized hydrodynamics, so can the Langevin – Uhlenbeck – Ornstein theory of Brownian motion be obtained from the fluctuating hydrodynamical equations of Landau and Lifshitz. In the case of dilute gases, the latter can be deduced from the Boltzmann equation with fluctuations. If this could in turn be derived from the Liouville equation, one would finally have connected the generalized hierarchy

Liouville equation
↓
Boltzmann equation with fluctuations
↓

Landau – Lifshitz equations

↓

Langevin equation

with microscopic dynamics and so provided a proper justification for it. This program remains to be completed. The extended Boltzmann equation has at least been derived from a master equation for the spatially homogeneous gas (sect. 8), to which the master equation approach is intrinsically limited.

For systems not close to equilibrium, governed by nonlinear laws, there is yet no general phenomenological theory of fluctuations like the Onsager – Machlup generalization of Langevin's theory of Brownian motion. A few such problems have been solved on a microscopic basis starting from a master equation, including the extension of the nonlinear Boltzmann equation for fluctuations arbitrarily far from equilibrium for a homogeneous dilute gas (sects. 9, 10).

We restrict ourselves entirely to fluctuation problems involving simple dilute gases, but begin with a still simpler model system, due to the Ehrenfests.

2. The Ehrenfest dog-flea model

Consider $2R$ balls numbered from 1 to $2R$, distributed in two urns I and II. We choose at random an integer between 1 and $2R$ (say by picking a lottery ticket from a bag containing tickets numbered 1, 2, . . . , $2R$, reading the ticket and replacing it in the bag), and move the ball whose number has been drawn from the urn it is in to the other. This procedure can be continued as long as you like. One has $N_{\mathrm{I}} + N_{\mathrm{II}} = 2R$, and we set $N_{\mathrm{I}} - N_{\mathrm{II}} = 2n$. Supposing one performs a trial every ten seconds, the results of a twenty-three day experiment might be like those in fig. 1.

This is the model analyzed by the Ehrenfests (1906, 1907, 1912)*. If initially there are more balls in I, we expect an approach to an equilibrium. The "naive equilibrium" is the situation in which there are R balls in each urn; i.e. $n = 0$. Yet this is not realistic, because there are fluctuations. The concept of equilibrium is more subtle, and is connected with the probabilities of deviations from "naive equilibrium". This leads to two problems:

(i) Find the equilibrium distribution (the "static" problem). In actual

* The "dog-flea" model served a crucial role in resolving the host of puzzles surrounding the question of irreversibility and the Boltzmann equation. See especially Ehrenfest (1912), Klein (1970), ch. 6 sect. 8, and Kac (1959) ch. 3 sect. 7.

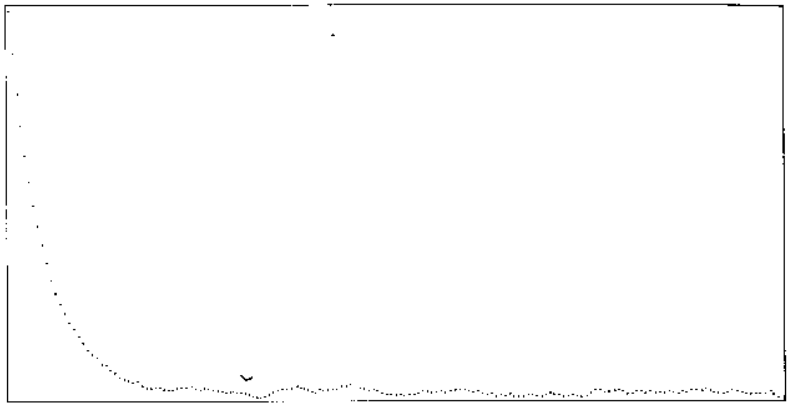

Fig. 1. Computer simulation of 200,000 drawings of the urn model with 16,384 balls. Starting with all the balls in urn I, N_{I} is plotted for every thousandth drawing.

physical situations, the answer to this problem is known and is given by the canonical or microcanonical distribution.

(ii) Determine how deviations from the "naive equilibrium" decay in time (the "dynamic" problem). For actual physical systems, this problem is unsolved in general. For small deviations the answer is provided by the theory of linear irreversible thermodynamics.

For this case, both problems can be solved. The Ehrenfest model is an example of a *Markov process*; i.e. the probability of finding after s steps a particular value of n depends only on the value of n found after $s-1$ steps (given the present, the future is independent of the past history).

Let us introduce the conditional probability $P(m_0|m;s)$, the probability of finding $N_{\mathrm{I}} = m$ after s steps, starting from m_0. Generally, we have for a Markov process

$$P(m_0|m_1 s_1; m_2 s_2; \ldots m_r s_r) = \prod_{i=0}^{r-1} P(m_i|m_i + 1; s_{i+1} - s_i). \tag{3}$$

One can easily see that for the problem at hand

$$P(m_0|m;s+1) = \frac{2R-(m-1)}{2R} P(m_0|m-1;s) + \frac{m+1}{2R} P(m_0|m+1;s).$$

Writing $n = m - R$, this difference equation becomes

$$P(n_0|n;s+1) = \frac{R-(n-1)}{2R} P(n_0|n-1;s) + \frac{R+(n+1)}{2R} P(n_0|n+1;s) \tag{4}$$

with the initial condition

$$P(n_0|n;0)=\delta(n_0,n)$$

where $\delta(n_0,n)$ is the Kronecker delta.

This formulation is seen to be equivalent to that of a random walk with attraction if one considers that the particle can move either Δ to the right or Δ to the left (see Kac 1947). That is, the probability of moving in either direction depends on the position of the particle. More precisely, if the particle is at position $\alpha\Delta$, the probabilities of moving right and of moving left are

$$P=\tfrac{1}{2}(1-\alpha/R), \qquad \text{or } =\tfrac{1}{2}(1+\alpha/R)$$

respectively. Possible positions of the particle are limited by the condition $-R\le\alpha\le R$. Let us call $W(n)$ the *stationary* probability distribution; that is, the one satisfying

$$\sum_{n_0} W(n_0)P(n_0|n;1)=W(n).$$

In our case, one can verify that

$$W(n)=\frac{(2R)!}{N_{\mathrm{I}}!\,N_{\mathrm{II}}!}\left(\frac{1}{2}\right)^{2R}$$
$$=\frac{(2R)!}{(R+n)!(R-n)!}\left(\frac{1}{2}\right)^{2R}. \tag{5}$$

This expression has a very sharp maximum for large R if $N_{\mathrm{I}}=N_{\mathrm{II}}=R$, the "naive equilibrium". In fact for large R one has

$$W(n)\sim(1/\sqrt{\pi R})\mathrm{e}^{-n^2/R}. \tag{6}$$

The average number of balls in say the first urn

$$\langle n(s)\rangle=\sum_n nP(n_0|n;s)$$

decreases to its equilibrium value exponentially

$$\langle n(s+1)\rangle=\sum nP(n_0|n;s+1)$$
$$=\sum n\left\{P(n_0|n-1;s)\left(\frac{1}{2}-\frac{n-1}{2R}\right)\right.$$
$$\left.+P(n_0|n+1;s)\left(\frac{1}{2}+\frac{n+1}{2R}\right)\right\}$$
$$=\sum(n-1)P(n_0|n-1;s)\left(\frac{1}{2}-\frac{n-1}{2R}\right)$$

$$+\sum(n+1)P(n_0|n+1;s)\left(\frac{1}{2}+\frac{n+1}{2R}\right)$$
$$+\sum P(n_0|n-1;s)\left(\frac{1}{2}-\frac{n-1}{2R}\right)$$
$$-\sum P(n_0|n+1;s)\left(\frac{1}{2}+\frac{n+1}{2R}\right)$$
$$=\langle n(s)\rangle-(1/R)\langle n(s)\rangle$$

hence

$$\langle n(s)\rangle=n_0(1-(1/R))^s$$

with $n_0=\langle n(0)\rangle$. In the limit $R\to\infty$, $\tau\to0$, $1/R\tau\to\gamma$, $s\tau=t$, this becomes

$$\langle n(t)\rangle=n_0\mathrm{e}^{-\gamma t} \tag{7}$$

just the monotonic approach to equilibrium characteristic of the linear laws like eq. (1) in sect. 1.

Coming back to the random walk formulation of the model, we are led to consider $P(n_0\Delta|n\Delta;s\tau)$. In the limit above, and for $\Delta\to0$, $\Delta^2/2\tau=D$; $n\Delta\to x$, $n_0\Delta\to x_0$, the difference equation, (4), becomes formally the differential equation

$$\frac{\partial}{\partial t}P(x_0|x;t)=\gamma\frac{\partial}{\partial x}[xP(x_0|x;t)]+D\frac{\partial^2}{\partial x^2}P(x_0|x;t). \tag{8}$$

This is the diffusion equation for the probability density $P(x_0|x;t)$ that a Brownian particle acted upon by the harmonic force $-\kappa x$ will be found at x at time t, having started from x_0 at $t=0$. D is the diffusion constant and $\gamma=\kappa/f$, where f is the usual friction coefficient (Smoluchowski 1916, or see Uhlenbeck and Ornstein 1930, and Wang and Uhlenbeck 1945).

In general, one can approach the problem of extending an average description [like eq. (7)] by explicitly adding a stochastic term, so that one is dealing with stochastic equations like the Langevin equation (2), or else by defining probability density functions like that above, seeking the equations which determine them, and finally solving for the explicit probability distribution. In the case of Brownian motion these two approaches – the Langevin equation and the Fokker–Planck equation – are equivalent.

The solution of Smoluchowski's Fokker–Planck equation (8) is

$$P(x_0|x;t)=(\gamma/(2\pi D))^{1/2}\Big/(1-\mathrm{e}^{-2\gamma t})^{1/2}\exp\left[-\frac{\gamma}{D}\frac{(x-x_0\mathrm{e}^{-\gamma t})^2}{2(1-\mathrm{e}^{-2\gamma t})}\right] \tag{9}$$

with

$$P(x_0|x;0) = \delta(x - x_0).$$

One can easily verify the semi-group property

$$\int_{-\infty}^{\infty} P(x_0|\xi;t)P(\xi|x;\tau)\mathrm{d}\xi = P(x_0|x;t+\tau). \tag{10}$$

In the limit $t \to \infty$, we obtain the solution

$$W(x) = (1/\sqrt{2\pi})\sqrt{\gamma/D}\,\mathrm{e}^{-(\gamma/D)x^2} \tag{11}$$

which has the property of stationarity

$$\int_{-\infty}^{\infty} W(x_0)P(x_0|x;t)\mathrm{d}x_0 = W(x). \tag{12}$$

Notice that we analyzed the Ehrenfest model in terms of the single "macroscopic" variable $n = N_1 - R$ which characterizes completely how far from "naive equilibrium" the system is.

Obviously the probabilistic law governing the change in time of the model depends only on this variable, and it is possible to treat it as a simple Markov process in n.

However the complete specification of the state of the system requires the location of every single ball, given by a $2R$-dimensional vector, each component of which can take one of two values indicating the location of a particular ball. There are 2^{2R} distinct "microstates" in this sense. It is obvious that in this microscopically detailed picture, the system can also be described as a Markov process. In fact this has been done by Siegert (1949) and Hess (1954), whose solution to the problem (surprisingly easy) is equivalent to what one gets in the one-dimensional version (Kac 1947).

In general, this is not so. The "contraction" of description (throwing away of information) involved may or may not simplify the problem, but generally does not preserve the Markov property: lumping states together for analysis generally destroys the Markovian character of the process. Little information is yet available on the conditions under which the Markov property is inherited on contraction. The character of relations between detailed and contracted descriptions, always a subtle issue, arises continually in fluctuation theory.

3. Density fluctuations in dilute gases

Suppose a gas of N particles (without internal degrees of freedom), interacting via short-range central forces, is contained in a volume V. The

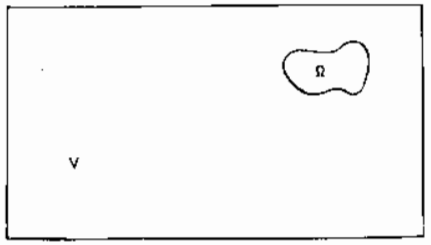

Fig. 2.

system is characterized by the Hamiltonian

$$H(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_n; \boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_N) = \sum_{i<j} \phi(|\boldsymbol{r}_i - \boldsymbol{r}_j|) + \sum_i p_i^2/2m$$

where the $\boldsymbol{r}_i$ and $\boldsymbol{p}_i$ are individual molecular spatial coordinates and momenta, respectively.

The probability of finding the system in the configuration $(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_n)$ at the equilibrium temperature T is proportional to

$$\mathrm{e}^{-\beta H}\mathrm{d}\boldsymbol{r}_1 \ldots \mathrm{d}\boldsymbol{r}_N, \qquad \beta = 1/kT.$$

Consider an arbitrary region Ω in the container and define the characteristic function ("indicator function")

$$g_\Omega(\boldsymbol{r}) = \begin{cases} 1 & \text{if } \boldsymbol{r} \in \Omega \\ 0 & \text{if not.} \end{cases}$$

The number of particles in Ω is then given by

$$N_\Omega = \sum_{k=1}^{N} g_\Omega(\boldsymbol{r}_k).$$

Now define

$$X_N = \frac{\sum_{k=1}^{N} g_\Omega(\boldsymbol{r}_k) - \rho|\Omega|}{\sqrt{\rho|\Omega|}} \tag{13}$$

where ρ is the average particle density $\rho = N/V$.

We want to evaluate the average $\langle \exp(\mathrm{i}\xi X_N) \rangle$. We have

$$\langle \mathrm{e}^{\mathrm{i}\xi X_N} \rangle = \frac{\int \mathrm{e}^{\mathrm{i}\xi X_N}\mathrm{e}^{-\beta H}\mathrm{d}\boldsymbol{r}_1 \ldots \mathrm{d}\boldsymbol{r}_N}{\int \mathrm{e}^{-\beta H}\mathrm{d}\boldsymbol{r}_1 \ldots \mathrm{d}\boldsymbol{r}_N}$$

$$= e^{-i\xi\sqrt{\rho|\Omega|}} \frac{\int \exp(i\xi \sum_k g_\Omega(r_k)/\sqrt{\rho|\Omega|} - \beta H) dr_1 \dots dr_N}{\int e^{-\beta H} dr_1 \dots dr_N}$$

We use the fact that, since $g = 0$ or 1 so that

$$e^{ag} = 1 + (e^a - 1)g.$$

and let

$$\zeta = e^{i\xi/\sqrt{\rho|\Omega|}} - 1$$

then

$$\langle e^{i\xi X_N} \rangle = e^{-i\xi\sqrt{\rho|\Omega|}} \frac{\int e^{-\beta H} \prod_{i=1}^{N} [1 + \zeta g_\Omega(r_i)] dr_1 \dots dr_N}{\int e^{-\beta H} dr_1 \dots dr_n}$$

$$= e^{-i\xi\sqrt{\rho|\Omega|}} \frac{\int e^{-\beta H} \left[1 + \zeta \sum_{i=1}^{N} g_\Omega(r_i) + \zeta^2 \sum_{i \le j}^{N} g_\Omega(r_i) g_\Omega(r_j) + \dots \right] dr_1 \dots dr_N}{\int e^{-\beta H} dr_1 \dots dr_N}.$$

The k-particle distribution function for the canonical ensemble is defined as

$$n_k(r_1, r_2, \dots, r_N; N, V) = \frac{N!}{(N-k)!} \frac{\int e^{-\beta H} dr_{k+1} \dots dr_N}{\int e^{-\beta H} dr_1 \dots dr_N} \tag{14}$$

so that, for example, $n_2(r_1, r_2; N, V)$ is the usual pair correlation function (cf. Uhlenbeck et al. 1963).

Thus

$$\langle e^{i\xi X_N} \rangle = e^{-i\xi\sqrt{\rho|\Omega|}} \sum_{k=0}^{\infty} \frac{\zeta^k}{k!} \int n_k(r_1, \dots, r_k; N, V) dr_1 \dots dr_k$$

which in the bulk limit $N \to \infty$, $V \to \infty$, $N/V = \rho$, becomes (if we are dealing with a single-fluid phase)

$$e^{-i\xi\sqrt{\rho|\Omega|}} \sum_{k=0}^{\infty} \frac{\zeta^k}{k!} \int \bar{n}_k(r_1, \dots, r_k; \rho) dr_1 \dots dr_k.$$

Following Mayer (1947, 1948), we introduce the usual cluster functions

$$\bar{\chi}_1(r_1) = \bar{n}_1(r_1)$$
$$\bar{\chi}_2(r_1, r_2) = \bar{n}(r_1, r_2) - \bar{n}_1(r_1)\bar{n}_1(r_2) \tag{15}$$
etc.

Then

$$\langle \exp(i\xi \sum_k [g_\Omega(r_k) - \rho|\Omega|]/\sqrt{\rho|\Omega|}) \rangle$$
$$= e^{-i\xi\sqrt{\rho|\Omega|}} \exp\left[\sum_{k=1}^{\infty} \frac{\zeta^k}{k!} \int_\Omega \dots \int_\Omega \bar{\chi}_k(r_1 \dots r_k) dr_1 \dots dr_k. \right.$$

In the limit $|\Omega| \to \infty$, $\zeta^k \sim (i\xi/\sqrt{\rho|\Omega|})^k$, so that since $\bar{\chi}_2(r_1, r_2) = \bar{\chi}_2(|r_2 - r_1|)$

$$\lim_{|\Omega| \to \infty} \left\langle \exp\left(i\xi \sum_k [g_\Omega(r_k) - \rho|\Omega|]/\sqrt{\rho|\Omega|} \right) \right\rangle$$
$$= \exp\{-\tfrac{1}{2}\xi^2 [1 + \int dr \bar{\chi}_2(|r|)]\}$$
$$= \lim_{|\Omega| \to \infty} \langle e^{i\xi X_N} \rangle \tag{16}$$

from the definition, eq. (13).

Thus the density fluctuations are Gaussian and are characterized by

$$\sigma^2 = 1 + \frac{1}{\rho} \int dr \bar{\chi}_2(|r|) = kT \left(\frac{\partial \rho}{\partial p} \right)_T$$

(the second equality is a so-called fluctuation theorem). As a random process, the fluctuations are clearly stationary. As we have just seen, the first density function W_1 is Gaussian. It has not been proved that the higher density functions are also Gaussian, but it is generally assumed to be so.

4. Stochastic processes

This is a brief review of standard notions. Many good treatments are available, including Wang and Uhlenbeck (1945) and Stratonovich (1963). Further information can be found in Feller (1950, 1966) and Doob (1953).

A stochastic process is simply a probability process: that is, any process whose evolution we can analyze in terms of probability. More precisely, we define a *stochastic process* as a family of random variables $x(t)$, $t \in \mathbb{R}$ such that the joint distributions of

$$x(t_1), x(t_2), \dots, x(t_n); \qquad t_1 < t_2 < \dots < t_n$$

are prescribed. The behavior of the random variables for various values of t can be thought of as a physical process with some random element in its structure. The classical example is the displacement of a Brownian particle in time, $x(t)$.

The above joint probabilities are assumed to be of the form

$$P\{x_1 < x(t_1) < x_1 + dx_1; \ldots x_n < x(t_n) < x_n + dx_n\} = W_n(x_1 t_1; \ldots; x_n t_n) dx_1 \ldots dx_n \tag{17}$$

i.e. $W_n dx_1 \ldots dx_n$ is the probability for the path to pass through the specified "gates" $x_i < x(t_i) < x_i + dx_i$. These *absolute probability densities* W_n must satisfy the general conditions of

(i) positivity: $W_n \geq 0$

(ii) normalization: $\int W_n dx_1 \ldots dx_n = 1$

(iii) compatibility:

$$\int W_n dx_j = W_{n-1}(x_1 t_1; \ldots x_{j-1} t_{j-1}; x_{j+1} t_{j+1}; \ldots x_n t_n). \tag{18}$$

For physicists, the W_n for $n = 1, 2, \ldots$ define the process. For mathematicians, it is necessary to prove that there is a measure space Ω such that $x(t)$ can be considered as $x(\omega; t)$, $\omega \in \Omega$, $t \in \mathbb{R}$, and the measure in Ω of the set of ω's such that

$$\begin{aligned} &x_1 < x(\omega; t_1) < x_1 + dx_1 \\ &\quad\vdots \\ &x_n < x(\omega; t_n) < x_n + dx_n \end{aligned} \tag{19}$$

is equal to $W_n(x_1 t_1; \ldots x_n t_n) dx_1 \ldots dx_n$. The space Ω can be taken to be the space of *all* "paths" $x(t)$, so that what is involved is the introduction of a measure in a space of paths consistent with (19). A theorem of Kolmogorov (1933) assures that such an underlying measure space always exists.

Stationarity is the property that all the probability distributions W_n are invariant under time translation

$$W_n(x_1 t_1, \ldots, x_n t_n) = W_n(x_1 t_1 + \tau, \ldots, x_n t_n + \tau). \tag{20}$$

(This is sometimes called "strong stationarity"; in the weak form, only the two-time correlations are invariant.)

The process is *time reversible* if

$$W_2(x_1 t_1; x_2 t_2) = W_2(x_2 t_1; x_1 t_2). \tag{21}$$

The process is called *Gaussian* if all its distributions are of the form

$$W_n \sim \exp\left[-\sum_{ij} a_{ij}(t_1, \ldots, t_n) x_i x_j \right]. \tag{22}$$

A Gaussian process need not be stationary; but when this is the case, the functions a_{ij} depend only upon the time differences $t_j - t_i$.

Instead of defining the process by the absolute probabilities $W_1, \ldots, W_n$ we can also do it by the *conditional probabilities* to find $x(t_{k+1})$ in the interval $(x_{k+1}, x_{k+1} + dx_{k+1}), \ldots,$ and $x(t_n)$ in the interval $(x_n, x_n + dx_n)$ *given* that $x(t_1), x(t_2), \ldots, x(t_k)$ have preassigned values $x_1, x_2, \ldots, x_k$. By definition these probabilities are given by the formula

$$P(x_1 t_1; \ldots x_k t_k | x_{k+1} t_{k+1}; \ldots x_n t_n) = \frac{W_n(x_1 t_1; \ldots x_n t_n)}{W_k(x_1 t_1; \ldots x_k t_k)}. \tag{23}$$

The process is Markovian (is a *Markov process*) if

$$P(x_1 t_1; \ldots x_{n-1} t_{n-1} | x_n t_n) = P(x_{n-1} t_{n-1} | x_n t_n) \tag{24}$$

i.e. the future is independent of the past, given the present. As a corollary we have

$$\begin{aligned} P(x_1 t_1 | x_2 t_2; x_3 t_3) &= \frac{W(x_1 t_1; x_2 t_2; x_3 t_3)}{W(x_1 t_1)} \\ &= \frac{W(x_1 t_1; x_2 t_2)}{W(x_1 t_1)} P(x_1 t_1; x_2 t_2 | x_3 t_3) \\ &= \frac{W(x_1 t_1; x_2 t_2)}{W(x_1 t_1)} P(x_2 t_2 | x_3 t_3) \\ &= P(x_1 t_1 | x_2 t_2) P(x_2 t_2 | x_3 t_3) \end{aligned}$$

and more generally

$$\begin{aligned} P(x_1 t_1 | x_2 t_2; \ldots x_n t_n) = P(x_1 t_1 | x_2 t_2) P(x_2 t_2 | x_3 t_3) \\ \ldots P(x_{n-1} t_{n-1} | x_n t_n). \end{aligned} \tag{25}$$

This is often taken as the definition of a Markov process instead of eq. (24), to which it is easily seen to be equivalent.

The generalization to multi-dimensional processes is completely straightforward – the x's are then vectors.

We state a few simple facts about Gaussian processes. Let $x(t)$ be a Gaussian process with $\langle x(t) \rangle = 0$, i.e.

$$W(x_1 t_1; \ldots x_n t_n) = \exp\left(-\tfrac{1}{2} \sum_{ij}^{n} a_{ij} x_i x_j \right) \sqrt{\det A} / (\sqrt{2\pi})^n \tag{26}$$

where $a_{ij} = a_{ij}(t)$ and $A = ((a_{ij}))$. We can write

$$\left\langle \exp\left(i \sum_{k=1}^{n} \xi_k x_k(t) \right) \right\rangle = \int_{-\infty}^{\infty} \ldots \int \exp\left(i \sum_{k=1}^{n} \xi_k x_k \right)$$

$$\times W(x_1 t_1; \ldots x_n t_n)\mathrm{d}x_1 \ldots \mathrm{d}x_n$$

$$= \int_{-\infty}^{\infty} \ldots \int \exp\left(\mathrm{i}\sum_k \xi_k x_k\right)\sqrt{\det A}/(2\pi)^{n/2}$$

$$\times \exp\left(-\tfrac{1}{2}\sum_{ij} a_{ij}x_i x_j\right)\mathrm{d}x_1 \ldots \mathrm{d}x_n$$

$$= \exp\left(-\tfrac{1}{2}\sum_{ij}\rho_{ij}\xi_i\xi_j\right) \tag{27}$$

where $R = ((\rho_{ij}))$ is the covariance matrix, $A = ((a_{ij}))$, and we have the identity $RA = I$. It follows at once that

$$\langle x^2(t_j)\rangle = \rho_{jj}, \qquad \langle x(t_i)x(t_j)\rangle = \rho_{ij}.$$

If the process is stationary, ρ_{jj} is independent of j and is a constant

$$\rho_{jj} = \langle x^2(t_j)\rangle = \sigma^2$$

and ρ_{ij} depends only on the time difference $t_i - t_j$

$$\rho_{ij} = \rho(|t_i - t_j|).$$

The function ρ is non-negative definite, i.e.

$$\sum_{ij}\rho_{ij}\xi_i\xi_j \geq 0$$

since

$$\left(\sum_{i=1}^{n}\xi_i x(t_i)\right)^2 \geq 0.$$

The process is completely determined by $\rho(t)$.

Doob's theorem (Doob 1942, 1953; also e.g. Wang and Uhlenbeck 1945): A stationary Gaussian process is Markovian iff the covariance has the form

$$\rho(t) = \mathrm{e}^{-\gamma|t|}. \tag{28}$$

Proof: From the assumption that the process is Markovian, it follows that (semigroup property)

$$P(x_0|x;t+\tau) = \int_{-\infty}^{\infty} P(x_0|y;t)P(y|x;\tau)\mathrm{d}y.$$

However, for every Gaussian process

$$P(x_0|x;t) = \frac{\exp[-(x - x_0\rho(t))^2/2\sigma^2(1-\rho^2(t))]}{\sigma\sqrt{2\pi}(1-\rho^2(t))^{1/2}} \tag{29}$$

we verify immediately (taking $\sigma = 1$) that

$$\int_{-\infty}^{\infty} P(x_0|x;t)x\,\mathrm{d}x = \langle x(t)\rangle_{x_0} = x_0\rho(t).$$

Then

$$\int P(x_0|x;t+\tau)x\,\mathrm{d}x = \int P(x_0|y;t)P(y|x;\tau)x\,\mathrm{d}x\mathrm{d}y \tag{30}$$

that is

$$x_0\rho(t+\tau) = x_0\rho(t)\rho(\tau).$$

The only continuous solution to (30) is

$$\rho(t) = \mathrm{e}^{-\gamma|t|}$$

which is actually the answer even for the normalization $\langle x^2(t)\rangle = \sigma^2 \neq 1$. The proof of the converse is very simple by verification.

Now, given an equation like the Langevin equation (2)

$$\frac{\mathrm{d}}{\mathrm{d}t}x + \gamma x = \tilde{e}(t) \tag{31}$$

what should be the properties of the fluctuating force $\tilde{e}(t)$ so that the solution of eq. (31) will be Gaussian, stationary and Markovian? The formal solution of this equation is

$$x(t) = x_0\mathrm{e}^{-\gamma t} + \int_0^t \mathrm{e}^{-\gamma(t-\tau)}\tilde{e}(\tau)\mathrm{d}\tau.$$

Then

$$\langle x(t)\rangle = x_0\mathrm{e}^{-\gamma t} \tag{32}$$

given that $x(o) = x_0$, so that $\langle\tilde{e}(t)\rangle = 0$. From the assumption that $x(t)$ is stationary, we deduce the expression for the variance

$$\langle(x(t) - x_0\mathrm{e}^{-\gamma t})^2\rangle = \sigma^2(1 - \mathrm{e}^{-2\gamma t}) \tag{33}$$

(where now we are averaging *twice* the second time over an ensemble of initial values x_0 distributed according to $W(x_0) = (1/\sigma\sqrt{2\pi})\exp(-x_0/2\sigma^2)$

$$\langle(x(t) - x_0\mathrm{e}^{-\gamma t})^2 = \int_0^t\int_0^t \mathrm{e}^{-\gamma(t-\tau_1)}\mathrm{e}^{-\gamma(t-\tau_2)}\langle\tilde{e}(\tau_1)\tilde{e}(\tau_2)\rangle\mathrm{d}\tau_1\mathrm{d}\tau_2$$

$$= \mathrm{e}^{-2\gamma t}\int_0^t\int_0^t \mathrm{e}^{-\gamma(\tau_1+\tau_2)}S(|\tau_2 - \tau_1|)\mathrm{d}\tau_1\mathrm{d}\tau_2$$

where we use the the easily verifiable fact that for $x(t)$ to be stationary, it is

necessary for $\tilde{e}(t)$ to be stationary. The only solution is

$$S(|\tau_2 - \tau_1|) = 2\gamma\sigma^2\delta(\tau_2 - \tau_1).$$

Thus $\tilde{e}(t)$ is a stationary process with a singular covariance.

The requirement that $x(t)$ is Gaussian implies a definite form for all the higher-order correlations of $\tilde{e}(t)$. In fact, it is easy to show that we must have

$$\langle \tilde{e}(\tau_1) \ldots \tilde{e}(\tau_{2k+1}) \rangle = 0$$

and

$$\langle \tilde{e}(\tau_1) \ldots \tilde{e}(\tau_{2k}) \rangle = \sum \langle \tilde{e}(\tau_i)\tilde{e}(\tau_j) \rangle \langle \tilde{e}(\tau_k)\tilde{e}(\tau_l) \rangle \ldots \tag{34}$$

where the sum is to be taken over all the different ways in which the $2k$ labels can be divided into k pairs.

Thus $\tilde{e}(t)$ is a stationary, Gaussian process with singular covariance proportional to $\delta(t - \tau)$. Such a random process is called "white noise" because its power spectrum is flat.

5. Brownian motion of a free particle: the Langevin equation*

For a free particle of mass M moving with velocity v in a fluid, the equation of motion is taken to be

$$M\frac{dv}{dt} + \gamma v = \tilde{e}(t) \tag{35a}$$

where γ is again the friction constant and $\tilde{e}(t)$ is the "fluctuating force", representing the nonsystematic portion of the interaction between the Brownian particle and the medium, not accounted by the linear friction term.

Following Langevin, one raises this [eq. (2)] from the status of a tautology [defining $\tilde{e}(t)$] by assigning definite properties to the fluctuating force:

$$\langle \tilde{e}(t) \rangle = 0, \qquad \langle \tilde{e}(\tau_1)\tilde{e}(\tau_2) \rangle = 2D\delta(\tau_2 - \tau_1) \tag{35b, c}$$

where D is a constant, and by assuming the process is Gaussian, which is equivalent to the specification of all the higher correlations according to eq. (34).

* The "classical" theory of Brownian motion of Langevin (1908), developed by Uhlenbeck and Ornstein (1930), is reviewed in Wang and Uhlenbeck (1945), Doob (1942), and in the other references given at the beginning of sect. 4. Chandrasekhar (1943) and Nelson (1967) contain historical notes on this and the earlier Einstein (1905, 1906) and Smoluchowski (1906) theories.

The idealization applies on an intermediate time scale τ, $M/\gamma \ll \tau \ll \tau_{\text{coll}}$, much larger than the time scale of molecular collisions, but shorter than the characteristic damping time, so that local molecular equilibrium is at each instant approximately attained. The random force is thus taken to be uncorrelated at different times (35c), and the local Maxwell – Boltzmann velocity distribution applies, supporting the Gaussian assumption.

Once this guess has been made it is straightforward to work out the consequences. In particular, one can derive the conditional distribution $P(v_0|v;t)$ for the probability that $v(t) = v$ given that $v(0) = v_0$ (Wang and Uhlenbeck 1945). The limit $t \to \infty$ gives the distribution function $W(v)$:

$$W(v) = \sqrt{\gamma/(2\pi D)}\, e^{-\gamma v^2/2D} \tag{36}$$

(the initial value disappears in the limit $t \to \infty$; the system eventually "forgets").

But on general grounds (the equipartition theorem) the stationary distribution of the Brownian particle should be Maxwellian

$$W(v) = \sqrt{M/(2\pi kT)}\, e^{-Mv^2/2kT}. \tag{37}$$

These are consistent if $\gamma/D = M/kT$, or

$$D = kT\gamma/M \tag{38}$$

which is the *Einstein relation*, the simplest (and first) "fluctuation-dissipation theorem"*.

To derive the "old" Einstein – Smoluchowski theory for a free particle, write

$$x(t) = x_0 + \int_0^t v(\tau)\,d\tau.$$

Because $v(t)$ is a Gaussian process with exponential covariance, $x - x_0$ is also a Gaussian process, since (39) is a linear relation. Moreover

$$\begin{aligned}\langle (x - x_0)^2 \rangle &= \int_0^t \int_0^t \langle v(\tau_1)v(\tau_2) \rangle \, d\tau_1 \, d\tau_2 \\ &= \sigma^2 \int_0^t \int_0^t \exp[-(\gamma/M)|\tau_2 - \tau_1|]\,d\tau_1\,d\tau_2 \\ &= \frac{2kT}{M} \int_0^t \int_0^\tau \exp[-(\gamma/M)(\tau - \tau_1)]\,d\tau\,d\tau_1\end{aligned}$$

* Since the universal gas constant R was measurable and since $k = R/N$, this relation made it possible for Perrin to determine the Avogadro number N from Brownian motion experiments. Fluctuation-dissipation theorems have a thermodynamical meaning and also a practical value which gives them a large and important role in physics.

$$= \frac{2kT}{\gamma}\left[t - \int_0^t \exp\left[-(\gamma/M)\tau\right]d\tau\right]$$

thus for $t \gg M/\gamma$

$$\langle (x - x_0)^2 \rangle \sim (2kT/\gamma)t = 2\bar{D}t \tag{39}$$

where the diffusion constant

$$\bar{D} = kT/\gamma.$$

This is the Einstein diffusion law, valid for times large compared to the damping time, as indicated (in fact the sample paths are everywhere nondifferentiable, indicating that the theory is not relevant at all for very small times). The Einstein – Smoluchowski theory is derived assuming $x(t)$ is a Markov process. It is clear however that $x(t)$ cannot be Markovian if $v(t)$ is Markovian. It is only for large times that we have a sufficient loss of "memory" and that $x(t)$ is approximately Markovian.

If the Langevin theory is right in assuming that the force exerted by the medium on a particle is

$$-\gamma\frac{dx}{dt} + \tilde{e}(t)$$

where $\tilde{e}(t)$ is the Gaussian process defined by (35), one should be able to analyze by the same method the Brownian motion of the harmonic oscillator*. For this case the equation of motion is

$$M\frac{d^2x}{dt^2} = -\kappa x(t) + \left[-\gamma\frac{dx}{dt} + \tilde{e}(t)\right] \tag{40}$$

where κ is the spring constant in the restoring force. It follows that $x(t)$ is a Gaussian stationary process, and the first task is to determine its covariance function $\rho(\tau)$. We shall treat this in more detail further on, but sketch here a way of deriving the formula.

Introduce the function $W(x_0p_0; xp; t)$, $p = M\mathrm{d}x/\mathrm{d}t$, and compute ($\tau > 0$)

$$\begin{aligned}\langle x(t)x(t+\tau)\rangle &= \int\int\int\int x_0 x W(x_0p_0; xp; \tau)\mathrm{d}x_0\mathrm{d}x\mathrm{d}p_0\mathrm{d}p \\ &= \int_{-\infty}^{\infty}\int \mathrm{d}p_0\mathrm{d}p \int_{-\infty}^{\infty} W(x_0p_0)P(x_0p_0|xp;\tau)xx_0\mathrm{d}x_0\mathrm{d}x \\ &= \int_{-\infty}^{\infty}\int \mathrm{d}p_0\mathrm{d}p \int_{-\infty}^{\infty} \mathrm{d}x_0 x_0 W(x_0p_0)\int_{-\infty}^{\infty} P(x_0p_0|xp;\tau)x\mathrm{d}x \\ &= \int_{-\infty}^{\infty}\mathrm{d}p_0 \int_{-\infty}^{\infty} \mathrm{d}x_0 x_0 W(x_0p_0)\langle x(\tau)\rangle_{x_0p_0}.\end{aligned}$$

*The rotational motion of a small mirror suspended on a fine wire in a gas (a harmonic oscillator for small angular displacements) was studied experimentally by Gerlach (1927) and analyzed by Uhlenbeck and Goudsmit (1929).

Now it remains to determine the conditional average $\langle x(\tau)\rangle$; $x(0) = x_0$, $p(0) = p_0$, and here one is tempted to merely average the equation of motion

$$M\frac{d^2}{dt^2}\langle x(t)\rangle + \gamma\frac{d}{dt}\langle x(t)\rangle + \kappa\langle x(t)\rangle = 0 \tag{41}$$

and solve it with the initial conditions

$$\langle x(0)\rangle = x_0, \qquad \langle p(0)\rangle = p_0.$$

Actually one gets the correct answer this way, but for a deeper reason than one might at first think. The deeper reason is that $(x(t), p(t))$ is jointly Markovian, a *vector* Markov process. It is not enough, as it appears at first glance, to merely use the fact that $\langle\tilde{e}(t)\rangle = 0$. *All* the stochastic properties of $\tilde{e}(t)$ are required; i.e. that it is a stationary, purely random Gaussian process.

Only if $\tilde{e}(t)$ is purely random will the process $(x(t), p(t))$ be stationary. To see this, consider the simpler case

$$\frac{d\xi}{dt} + \gamma\xi = \tilde{e}(t).$$

As before, we get

$$\xi(t) = \xi_0 e^{-\gamma t} + \int_0^t e^{-\gamma(t-\tau)}\tilde{e}(\tau)d\tau$$

and

$$\xi(t) = \xi_0 e^{-\gamma t}$$

if $\langle\tilde{e}(t)\rangle = 0$. However if $\xi(t)$ is to be stationary, we must have

$$\begin{aligned}\langle\xi^2\rangle &= \int_{-\infty}^{\infty} W(\xi_0)P(\xi_0|\xi;t)\xi^2 d\xi \\ &= \langle\xi_0^2\rangle = \int_{-\infty}^{\infty} W(\xi_0)\xi_0^2 d\xi_0.\end{aligned}$$

On the other hand

$$\langle\xi_0^2\rangle = \langle\xi_0^2\rangle e^{-2\gamma t} + \int_0^t\int e^{-\gamma(t-\tau_1)}e^{-\gamma(t-\tau_2)}\sigma(\tau_2-\tau_1)d\tau_1 d\tau_2$$

or

$$\langle\xi_0^2\rangle(1 - e^{-2\gamma t}) = e^{-2\gamma t}\int_0^t\int e^{\gamma(\tau_1+\tau_2)}\sigma(\tau_2-\tau_1)d\tau_1 d\tau_2$$

or

$$\langle\xi_0^2\rangle(e^{2\gamma t} - 1) = 2\int_0^t e^{\gamma\tau_2}d\tau_2\int_0^{\tau_2} e^{\gamma\tau_1}\sigma(\tau_2-\tau_1)d\tau_1.$$

By differentiation

$$\langle \xi_0^2 \rangle \gamma e^{2\gamma t} = e^{\gamma t} \int_0^t e^{\gamma \tau} \sigma(t-\tau) d\tau$$

so that

$$\sigma(t-\tau) = 2\gamma \langle \xi_0^2 \rangle \delta(t-\tau).$$

In calculating

$$\langle \xi \rangle_{\xi_0} = \xi_0 e^{-\gamma t}$$

from the equation

$$\frac{d\xi}{dt} + \gamma \xi = \tilde{e}(t)$$

we have in mind a prepared ensemble of trajectories $\xi(t)$ originating from ξ_0 and averaging (over the ensemble) the values at time t.

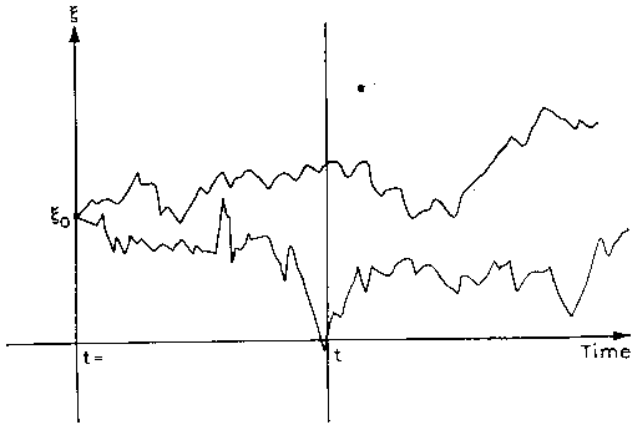

Fig. 3.

In calculating $\langle \xi \rangle_{\xi_0}$ by the formula

$$\langle \xi \rangle_{\xi_0} = \int_{-\infty}^{\infty} P(\xi_0 | \xi; t) \xi d\xi$$

we actually have in mind one *time* record of $\xi(t)$, taking the times $t_1, t_2, \ldots$, at which $\xi(t) = \xi_0$, looking at the values $\xi(t_1 + t), \xi(t_2 + t), \ldots$, and tracing their time averages. In order for the two ways of calculating $\langle \xi \rangle_{\xi_0}$ to give

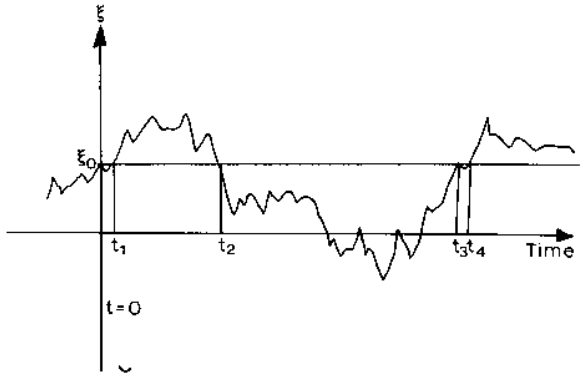

Fig. 4.

the same result an ergodic theorem must hold, and this can happen only if the process is stationary.

Returning to eq. (40), the equation of motion has the following solution in the overdamped case ($\bar{\gamma} = \gamma/M \gg \omega_0$, where $\omega_0^2 = \kappa/M$)

$$\langle x(t) \rangle = \frac{\alpha_2 x_0 + p_0}{\alpha_2 - \alpha_1} \left\{ \exp - \frac{\bar{\gamma} - \omega}{2} \right\} t + \frac{\alpha_1 x_0 + p_0}{\alpha_1 - \alpha_2} \exp \left\{ - \frac{\bar{\gamma} + \omega}{2} t \right\} \quad (42)$$

where $\omega^2 = \gamma^2 - 4\omega_0^2$. Integrating on p_0, we obtain

$$\int_{-\infty}^{\infty} W(p_0) \langle x(t) \rangle_{x_0 p_0} dp_0 = \frac{\alpha_2 x_0}{\alpha_2 - \alpha_1} \exp - \left\{ \frac{\bar{\gamma} - \omega}{2} t \right\} - \frac{\alpha_1 x_0}{\alpha_2 - \alpha_1} \exp \left\{ - \frac{\bar{\gamma} + \omega}{2} t \right\}$$

and then

$$\rho(t) = \sigma^2 \left[\frac{\alpha_2}{\alpha_2 - \alpha_1} e^{-\alpha_1 |t|} - \frac{\alpha_1}{\alpha_2 - \alpha_1} e^{-\alpha_2 |t|} \right] \quad (43)$$

where $\alpha_1 = \frac{1}{2}(\bar{\gamma} - \omega)$, $\alpha_2 = \frac{1}{2}(\bar{\gamma} + \omega)$. In the limit $\bar{\gamma}/\omega_0 \to \infty$, we have

$$\rho(t) \sim \sigma^2 e^{-(\omega_0^2/\bar{\gamma})t}. \quad (44)$$

$x(t)$ is "approximately" Markovian and its probability distribution satisfies the Fokker – Planck equation

$$\frac{\partial}{\partial t} P(x;t) = \bar{D} \frac{\partial^2 P}{\partial x^2} + \frac{\partial}{\partial x} [(\omega_0^2/\bar{\gamma}) x P(x;t)]. \quad (45)$$

In this limit, we find

$$P(x) \sim e^{-(\kappa/2\bar{D}\gamma)x^2} = e^{-(\kappa/kT)x^2}$$

that is

$$P \sim e^{-E/kT}.$$

6. The Fokker–Planck equation

The Markov processes dealt with in the preceding sections are amenable to treatment by mathematical methods which give a large insight into actual systems encountered in practice. The essential starting point in describing a continuous (in a definite sense made clear below) Markov process $X(t)$ is the Fokker – Planck equation obeyed by the function $P(X_0|X;t,t_0)$, $P(X_0|X;t,t_0)\mathrm{d}X$ being the probability that X assumes a value between X and $X+\mathrm{d}X$ at time t, starting from $X=X_0$ at $t=t_0$.

Assuming homogeneity in time and setting $t_0=0$, we observe that

$$P(X_0|X;t+\Delta t) = \int P(X_0|Y;t)P(Y|X;\Delta t)\mathrm{d}Y \tag{46}$$

and we assume that the moments of the changes in coordinate in time interval Δt (the "jump moments") are such that

$$\int_{-\infty}^{\infty}(X-X_0)P(X_0|X;\Delta t)\mathrm{d}X = \langle (X-X_0)\rangle_{\Delta t} \sim A(X_0)\Delta t + \mathrm{O}((\Delta t)^2) \tag{47a}$$

$$\int_{-\infty}^{\infty}(X-X_0)^2 P(X_0|X;\Delta t)\mathrm{d}X = \langle (X-X_0)^2\rangle_{\Delta t} \sim B(X_0)\Delta t + \mathrm{O}((\Delta t)^2) \tag{47b}$$

$$\int_{-\infty}^{\infty}(X-X_0)^k P(X_0|X;\Delta t)\mathrm{d}X = \langle (X-X_0)^k\rangle_{\Delta t} \sim \mathrm{O}((\Delta t)^2), \qquad k\geq 3. \tag{47c}$$

There are various ways to obtain the Fokker – Planck equation (see for example Stratonovich 1963, ch. 4; Wang and Uhlenbeck 1945, sects. 8–10; Feller 1950); we describe the method of Kolmogorov (1933).

Let $\phi(X)$ be a test function, as nice as it can be. Then, from eq. (46)

$$\int_{-\infty}^{\infty} P(X_0|X;t+\Delta t)\phi(X)\mathrm{d}X = \int_{-\infty}^{\infty}\int P(X_0|Y;t)P(Y|X;\Delta t)\phi(X)\mathrm{d}X\mathrm{d}Y.$$

Expanding $\phi(X)$ about Y we have

$$\phi(X) = \phi(Y) + (X-Y)\phi'(Y) + (1/2!)(X-Y)^2\phi''(Y) + \ldots.$$

Hence

$$\int_{-\infty}^{\infty} P(X_0|X;\Delta t)\phi(X)\mathrm{d}X = \phi(y) + \int_{-\infty}^{\infty} P(X_0|Y;t)[\phi'(Y)\{A(Y)\Delta t + \mathrm{O}((\Delta t)^2)\} + (1/2!)\phi''(Y)\{B(Y)\Delta t + \mathrm{O}((\Delta t)^2)\} + \mathrm{O}((\Delta t)^2)]\mathrm{d}y.$$

This yields

$$\int_{-\infty}^{\infty}\frac{\partial}{\partial t}P(X_0|X;t)\phi(X)\mathrm{d}X = \int_{-\infty}^{\infty} P(X_0|X;t)[A(X)\phi'(X) + \tfrac{1}{2}B(X)\phi''(X)]\mathrm{d}X = \int_{-\infty}^{\infty}\left\{-\frac{\partial}{\partial X}[A(X)P(X_0|X;t)] + \frac{1}{2}\frac{\partial^2}{\partial X^2}[B(X)P(X_0|X;t)]\right\}\phi(X)\mathrm{d}X$$

hence, the Fokker – Planck equation

$$\frac{\partial}{\partial t}P(X_0|X;t) = -\frac{\partial}{\partial X}[A(X)P(X_0|X;t)] + \frac{1}{2}\frac{\partial^2}{\partial X^2}[B(X)P(X_0|X;t)] \tag{48}$$

with the initial condition

$$P(X_0|X;0) = \delta(X-X_0).$$

The existence and uniqueness of solutions of the Fokker – Planck equation have been studied, using the theory of semigroups (see for example Feller 1966).

Several well-known diffusion processes are particular cases of (48). If we assume $B(X) = \text{constant} = 2D$ and $A=0$, we obtain the equation for the Wiener process

$$\frac{\partial}{\partial t}P(X;t) = D\frac{\partial^2}{\partial X^2}P(X;t) \tag{49}$$

where D is the coefficient of diffusion. If the coefficients A and B are such

that the Fokker – Planck equation takes the form

$$\frac{\partial}{\partial t}P(X;t)=\gamma\frac{\partial}{\partial X}[XP(x;t)]+\frac{\partial^2}{\partial X^2}P(X;t) \tag{50}$$

we have the equation for the one-dimensional Uhlenbeck – Ornstein process.

With X replaced by v, this process can be derived from the Langevin equation (35)

$$M\frac{\mathrm{d}}{\mathrm{d}t}v(t)+\gamma v(t)=\tilde{e}(t).$$

Thus

$$M[v(t+\Delta t)-v(t)]+\gamma\int_t^{t+\Delta t}v(\tau)\mathrm{d}\tau=\int_t^{t+\Delta t}\tilde{e}(\tau)\mathrm{d}\tau$$

hence

$$M\Delta v+\gamma v(t)\Delta t\int_t^{t+\Delta t}\tilde{e}(\tau)\mathrm{d}\tau+\mathrm{O}((\Delta t)^2).$$

Thus

$$M\langle\Delta v\rangle=-\gamma v\Delta t+\mathrm{O}((\Delta t)^2)$$

which corresponds to $A(v)=-\gamma v$, a linear function of v, in eq. (47a). Similarly

$$M^2\langle(\Delta v)^2\rangle=\left\langle\left[\int_t^{t+\Delta t}\tilde{e}(\tau)\mathrm{d}\tau\right]^2\right\rangle$$
$$=\int_t^{t+\Delta t}\int\langle\tilde{e}(\tau_1)\tilde{e}(\tau_2)\mathrm{d}\tau_1\mathrm{d}\tau_2\rangle=2D\Delta t$$

using eq. (35c). Thus the coefficient for eq. (47b) is $B(v)=2D/M^2$, a constant. Using the Gaussian property assumed for eq. (35), one can then show that all the higher moments (47c) vanish.

This analysis can be easily extended to a vector Markov process consisting of several independent random functions $\alpha_N(t)=\{\alpha_1(t,\ldots,\alpha_N(t)\}$. The probability density satisfies then the multi-dimensional Fokker – Planck equation

$$\frac{\partial P}{\partial t}=-\sum_{i=1}^{N}\frac{\partial}{\partial\alpha_i}[A_i(\alpha_N)P]+\frac{1}{2}\sum_{i,j=1}^{N}\frac{\partial^2}{\partial\alpha_i\partial\alpha_j}[B_{ij}(\alpha_N)P]. \tag{51}$$

For the Brownian motion of a particle subject to an outside force $\kappa(x)$, described by the Langevin equation

$$\frac{\partial p}{\partial t}+\gamma p+\kappa(x)=\tilde{e}(t),\qquad p(t)=M\frac{\mathrm{d}x}{\mathrm{d}t}$$

one gets the two-dimensional Fokker – Planck equation

$$\frac{\partial P}{\partial t}=-p\frac{\mathrm{d}P}{\mathrm{d}x}+\frac{\partial}{\partial p}[(\gamma p+\kappa(x))P]+D\frac{\partial^2}{\partial p^2}$$

called the Kramers equation (Kramers 1940), which has to be solved with the initial condition

$$P(x,p;0)=\delta(x-\breve{x}_0)\delta(p-p_0).$$

This is a diffusion equation in the momentum but *not* in position.

The stochastic process describing the coordinates of a Brownian particle in a force field is in general *not* Markovian by itself, only the joint process $(x(t), p(t))$ governed by the Kramers equation is. But in the special case of the overdamped harmonic oscillator $[\kappa(x)=\kappa x]$, eq. (40), the Smoluchowski equation is an adequate approximation.

7. An example of Brownian motion *

A little mirror is rigidly attached to a stiff wire, which can glide without friction along another wire. This arrangement is immersed in a large container of volume V containing N particles of mass m, kept at temperature T. The gas is supposed to be so dilute that the particles are completely independent (Knudsen limit), and the motion of the mirror does not disturb the Maxwellian distribution of velocities. That is, the mean-free path is comparable to the size of the vessel, and large compared with the size of the mirror.

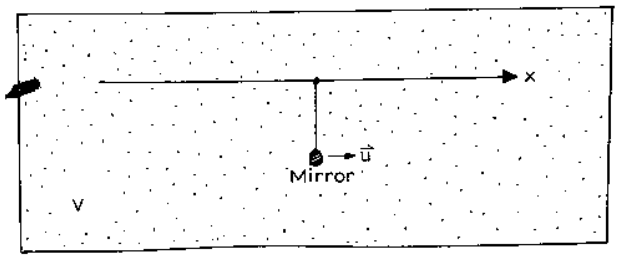

Fig. 5.

* This analysis is taken from Uhlenbeck and Goudsmit (1929) with minor changes.

The mirror, of mass M and surface area A (surface density $\rho = M/A$) moves along the x-axis with velocity $u(t)$. The mean velocity $\bar{c}$ is (ξ is the component of molecular velocity along x, η and ζ the components perpendicular)

$$\bar{c} = (m/2\pi kT)^{3/2} \int_{-\infty}^{\infty}\int\int (\xi^2+\eta^2+\xi^2)^{1/2} \times \exp - \left\{\frac{m}{2kt}(\xi^2+\eta^2+\xi^2)\right\} d\xi\, d\eta\, d\zeta = \left(\frac{8kT}{\pi m}\right)^{1/2}. \tag{52}$$

In the time interval Δt the amount of momentum imparted to the mirror from the left is (writing $b^* = b, 0$ according as $b > 0$, $b \le 0$)

$$M_L = \int_0^\infty 2m(\xi-u)^* \frac{NA}{V}(\xi-u)^* e^{-m\xi^2/2kT}\sqrt{m/(2\pi kT)}\, d\xi\, \Delta t$$

$$= \frac{2Nm}{V}\int_u^\infty A(\xi-u)^2\sqrt{m/(2\pi kT)}\, e^{-m\xi^2/2kT} d\xi\, \Delta t$$

in the same way, that from the right is

$$M_R = -\frac{2Nm}{V}\int_{-u}^{\infty} A(\xi+u)^2\sqrt{m/(2\pi kT)}\, e^{-m\xi^2/2kT} d\xi\, \Delta t.$$

The assumption $u \ll \bar{c}$ *allows the neglect of the integrals* $\int_0^u \ldots d\xi$ and $\int_{-u}^0 \ldots d\xi$, since

$$\int_0^u (\xi-u)^2 e^{-m\xi^2/2kT}\sqrt{m/(2\pi kT)}\, d\xi \sim (2\pi)^{-1/2}(\sqrt{m/kT})u^3\int_0^1 (1-v)^2 dv$$

which is to be compared with

$$u\int_0^\infty \xi e^{-m\xi^2/2kT}\sqrt{m/(2\pi kT)}\, d\xi = (2\pi)^{-1/2}u\sqrt{kT}/\sqrt{m}.$$

But

$$(2\pi)^{-1/2}(\sqrt{m}/\sqrt{kT})u^3 = (2\pi)^{-1/2}(\sqrt{kT}/\sqrt{m})u^2/(kT/m) = (2\pi)^{-1/2}(\sqrt{kT}/\sqrt{m})u \times O(u^2/\bar{c}^2)$$

so that the corrections are of order $(u/\bar{c})^2 \ll 1$.

Thus the total momentum transferred to the mirror per unit time is

$$-(4Nm/V)2u\int_0^\infty A\xi e^{-m\xi^2/2kT}\sqrt{m/(2\pi kT)}\, d\xi$$
$$= -8Nm/V)\sqrt{kT/(2\pi m)}\, Au = -2(N/V)m\bar{c}Au$$
$$= -2m\bar{c}Aup/kT$$

using $pV = NkT$, for the ideal gas. So the friction constant is given by the formula

$$\gamma = -2m\bar{c}(p/kT)M/\rho. \tag{53}$$

Now let us calculate the probability that during Δt a molecule of the gas will impart to the surface element ΔA, held stationary, a momentum normal to it lying between G and $G + \Delta G$. One begins by calculating the closely related probability that the momentum imparted normally to ΔA in time Δt should lie between G and ΔG and at the same time the other two components of velocity are η and ξ within $\Delta\eta$ and $\Delta\xi$, respectively. This probability is

$$\left(\frac{m}{2\pi kT}\right)^{3/2}\exp -\frac{m}{2kT}\left[\left(\frac{G}{2m}\right)^2+\eta^2+\zeta^2\right] \times \frac{\Delta G}{2m}\Delta\eta\,\Delta\zeta[(G/2m)\Delta t\,\Delta A]/V, \qquad (G/2m) \le \xi \le (G+\Delta G)/2m).$$

Observe that $(G/2m)\Delta A\,\Delta t$ is the volume (*independent* of η and ζ) of the "collision cylinder", i.e., the cylinder containing particles which will impart the desired normal momentum and have their y and z components of velocity equal to η and ζ within $\Delta\eta$ and $\Delta\zeta$. Integrating this expression over η and ζ, we obtain

$$\left(\frac{m}{2\pi kT}\right)^{3/2}\frac{2\pi kT}{m}\frac{1}{(2m)^2}\exp -(G^2/8mkT)\frac{G\Delta G\Delta T\Delta A}{V} = \frac{\pi kT}{2V}(2\pi mkT)^{-3/2}\exp -(G^2/8mkT)G\Delta G\Delta t\Delta A$$

which is the desired probability.

Let n_j be the number of molecules which during the interval Δt impart a momentum G_j within ΔG_j to ΔA, normal to it. Then the total momentum imparted to ΔA normal to it is

$$\sum_{j=0}^{\infty} n_j G_j.$$

The average square of this quantity centered about its mean is

$$\left\langle \left(\sum_j n_j G_j - \sum_j \langle n_j G_j \rangle \right)^2 \right\rangle = \left\langle \left[\sum_j G_j (n_j - \langle n_j \rangle) \right]^2 \right\rangle$$
$$= \sum_j G_j^2 (\langle n_j^2 \rangle - \langle n_j \rangle^2)$$
$$= \sum_j G_j^2 \langle n_j \rangle \tag{54}$$

where we have used the independence of particles in setting cross terms equal to zero, and the relation

$$\langle n^2 \rangle - \langle n \rangle^2 = \langle n \rangle$$

which follows from the fact that, for small Δt, n has the Poisson distribution. Replacing the sum by an integral, we find for the mean-square average of the total normal momentum imparted to the mirror the quantity

$$\frac{\pi N}{2V} kT \left(\frac{1}{2\pi mkT} \right)^{3/2} \int_{-\infty}^{\infty} \mathrm{d}G e^{-G^2/(8mkT)} G^3 \Delta t \Delta A = 2m\bar{c}pA\Delta t.$$

The force is this momentum divided by Δt. Thus the square of the fluctuations in time Δt of the force on the mirror is

$$\langle [F(t) - \langle F(t) \rangle]^2 \rangle = 2m\bar{c}pA/\Delta t. \tag{55}$$

Since the mirror is not stationary and is bombarded from both sides, the effective average force is zero and its fluctuation squared twice the above; i.e.

$$\frac{4m\bar{c}pA}{\Delta t} = \frac{4m\bar{c}p}{\Delta t} \frac{M}{\rho}.$$

But

$$F(t) = \frac{1}{\Delta t} \int_t^{t+\Delta t} \tilde{e}(\tau) \mathrm{d}\tau, \qquad (\langle \tilde{e}(\tau) \rangle = 0)$$

according to the Langevin equation; in order that $\langle F^2(t) \rangle$ be proportional to $1/\Delta t$ the covariance of $\tilde{e}(\tau)$ must be singular, i.e.

$$\langle \tilde{e}(\tau_1) \tilde{e}(\tau_2) \rangle = 2D\delta(\tau_2 - \tau_1).$$

Thus

$$D = 2m\bar{c}pM/\rho \tag{56}$$

and we find again the Einstein relation

$$D = \gamma kT \tag{57}$$

with

$$\gamma = 2m\bar{c} \frac{p}{kT} \frac{M}{\rho}.$$

8. Brownian motion far from equilibrium*

The little mirror interacting with a Knudsen gas of the last section can be studied in the regime $u \sim \bar{c}$, where the mirror's velocity is comparable to the mean velocity of a gas molecule, and very much larger than its own equilibrium velocity. Since the molecules of a Knudsen gas interact with each other negligibly compared to their interaction with the mirror, and since the mirror is small compared to the mean free path, the molecular distribution is insignificantly affected by the mirror and so does not deviate from its initial form, the Maxwell distribution for temperature T.

Thus the system can be described by the *linear*** Boltzmann equation, which can be written in the form (Wang Chang and Uhlenbeck 1970)

$$\frac{\partial}{\partial t} P(v_x, t) = \int_{-\infty}^{\infty} \{R(v_x, v_x') P(v_x', t) - R(v_x', v_x) P(v_x, t)\} \mathrm{d}v_x' \tag{58}$$

where

$$R(v_x', v_x) = \left(2\pi \frac{kT}{m} \right)^{-1/2} A \frac{N}{V} \left(\frac{M+m}{2m} \right)^2$$
$$\times \exp\left\{ -\frac{(\frac{1}{2}(M+m)v_x' - \frac{1}{2}(M-m)v_x)^2}{2mkT} \right\} |v_x - v_x'| \tag{59}$$

in which the conservation of energy and momentum have been utilized. Equation (58) has the form of a linear master equation of the type which Van Kampen has taken as the basis of a general treatment of fluctuations far from equilibrium (Van Kampen 1961, 1976a, b; see also Kubo et al. 1973). In the sense of the discussion in sect. 1, the Boltzmann equation is to the (average and fluctuating) particle equation of motion as the N-particle

*This section is taken from Fox and Kac (1977).

**This is the Boltzmann equation for a test particle influenced by, but not reacting back upon, the distribution of field particles, familiar from the Rayleigh problem. It should not be confused with the *linearized* Boltzmann equation, describing a fully interacting collection of particles close to equilibrium. See Uhlenbeck and Ford (1963).

master equation is to the (average and fluctuating) Boltzmann equation.

Near equilibrium, when the velocity of the mirror is small compared to the mean molecular velocity and for a relatively heavy mirror

$$\mu = m/M \ll 1$$

it is known (Wang Chang and Uhlenbeck 1970) that to lowest order in μ the linear Boltzmann equation can be approximated by a Fokker – Planck equation equivalent to eq. (35), with approximately defined constants γ and D.

Again, near equilibrium, we have two equivalent ways of treating Brownian motion: by the Langevin equation, and by the Fokker – Planck equation. Attempts to generalize away from equilibrium have proceeded on both lines (e.g. Mori 1973, Van Kampen 1973, 1976a, b, c, for reviews: Keizer 1975, 1976). Here we proceed by the Fokker - Planck method, starting from the "master equation" (58).

The problem is to solve the initial value problem for the Boltzmann equation, to find $P(v_x; t)$ given $P(v_x; 0)$. For the near-equilibrium case an appropriate solution, to lowest order in μ, is obtained by solving the Fokker – Planck equation mentioned above. What remains to be seen is whether any similar simplification can be obtained far from equilibrium.

First of all, from (58, 59) we obtain, by averaging with respect to $P(v_x; t)$

$$M\frac{\mathrm{d}}{\mathrm{d}t}\langle v_x(t)\rangle = A\frac{N}{V(1+\mu)m}\frac{2}{\sqrt{}}(2\pi mkT)^{-1/2} \times \int_{-\infty}^{\infty} \mathrm{d}\xi \left\langle \exp\left\{-\frac{(\xi + mv_x(t))^2}{2mkT}\right\}\right\rangle \xi|\xi|.$$

Hence *neglecting fluctuations*

$$M\frac{\mathrm{d}}{\mathrm{d}t}\langle v_x(t)\rangle = A\frac{N}{V(1+\mu)m}\frac{2}{\sqrt{}}(2\pi mkT)^{-1/2} \times \int_{-\infty}^{\infty} \mathrm{d}\xi \exp\left\{-\frac{(\xi + m\langle v_x(t)\rangle)^2}{2mkT}\right\}\xi|\xi|. \tag{60}$$

It is not difficult to show that this transcendental dependence upon $\langle v_x(t)\rangle$ leads to the Rayleigh form

$$M\frac{\mathrm{d}}{\mathrm{d}t}\langle v_x(t)\rangle = -\frac{8}{\sqrt{2\pi}}(mkT)^{1/2}A\frac{N}{V}\frac{1}{1+\mu}\langle v_x(t)\rangle \tag{61}$$

when $\langle v_x(0)\rangle$ is small compared with $(kT/m)^{1/2}$, and to the Newtonian form

$$M\frac{\mathrm{d}}{\mathrm{d}t}\langle v_x(t)\rangle = -2A\frac{N}{V}\frac{m}{1+\mu}|\langle v_x(t)\rangle|\langle v_x(t)\rangle \tag{62}$$

when $\langle v_x(0)\rangle$ is of the order of $(kT/m)^{1/2}$ or larger.

To insure that at $t = 0$ the fluctuations are negligible it is convenient and natural to take

$$P(v_x, 0) = \delta(v_x - v_x(0)) \tag{63}$$

i.e., the initial velocity is sharply prescribed so that $\langle v_x(0)\rangle = v_x(0)$.

However, this initial distribution will spread in time (for $t \to \infty$ it must approach the Maxwellian distribution) and the validity of eq. (60) depends on fluctuations about the average remaining small. As long as one is far away from equilibrium, i.e., $v_x(0)$ is of the order of $(kT/m)^{1/2}$, "small" means that deviations from the average are small compared to the average itself. Near equilibrium (i.e., $v_x(0)$ of the order $(kT/M)^{1/2}$) the above statement has to be modified.

Before proceeding with the discussion it will be convenient to introduce dimensionless variables v, u and τ.

$$v_x' = \sqrt{\frac{kT}{M}}v, \qquad v_x = \sqrt{\frac{kT}{M}}u, \qquad t = \frac{V}{AN}\sqrt{\frac{M}{kT}}\tau.$$

In these variables the Boltzmann equation (58) becomes

$$\frac{\partial}{\partial\tau}p(u;\tau) = \left(\frac{1+\mu}{2\mu}\right)^2 \int_{-\infty}^{\infty} \left\{p(v;\tau)\phi_0\left(\frac{1+\mu}{2\sqrt{\mu}}u - \frac{1-\mu}{2\sqrt{\mu}}v\right) - p(u;\tau)\phi_0\left(\frac{1+\mu}{2\sqrt{\mu}}v - \frac{1-\mu}{2\sqrt{\mu}}u\right)\right\}|u-v|\,\mathrm{d}v \tag{64}$$

where

$$p(u;\tau) = \sqrt{\frac{kT}{M}}p\left(\sqrt{\frac{kT}{M}}u; \frac{V}{AN}\sqrt{\frac{M}{kT}}\tau\right)$$

and

$$\phi_0(u) = (1/\sqrt{2\pi})\exp(-\tfrac{1}{2}u^2).$$

Away from equilibrium where $u\sqrt{\mu}$ is of order 1 we set

$$p(u;\tau) = p((w(\tau)/\sqrt{\mu}) + \xi; \tau) = \bar{p}(\xi;\tau) \tag{65}$$

where $w(\tau)$ is to be determined a little later.

The motivation behind this substitution is the following: one expects the average (mean) to be of the order $1/\sqrt{\mu}$ so that the distribution of u would be centered about $w(\tau)/\sqrt{\mu}$, and one hopes that the deviation from the mean will be of order 1.

Such a hope can only be justified if at $\tau = 0$ the (initial) distribution about $u(0) = w(0)/\sqrt{\mu}$ has width of order 1. To simplify matters we shall first assume that

$$p(u;0) = \delta(u - (w(0)/\sqrt{\mu}))$$

or, equivalently

$$\bar{p}(\xi;0) = \delta(\xi). \tag{66}$$

This is merely a translation of (63) into dimensionless terms.

Strictly speaking one must now prove that the width of $\bar{p}(\xi;\tau)$ will be of order 1 for all $\tau > 0$ [in an analogous problem concerning the birth – death model of chemical reactions such a proof has in fact been given by Kurtz (1972)] but we shall be content here to show that the assumption that the width of $\bar{p}(\xi;\tau)$ remains of order 1 leads to a consistent theory.

In terms of $\bar{p}(\xi;\tau)$, and upon setting

$$v = \zeta + w(\tau)/\sqrt{\mu}$$

the Boltzmann equation assumes the form

$$\frac{\partial \bar{p}(\xi;\tau)}{\partial \tau} - \frac{\partial \bar{p} w'(\tau)}{\partial \xi \sqrt{\mu}}$$
$$= \left(\frac{1+\mu}{2\mu}\right)^2 \int_{-\infty}^{\infty} \left\{ \bar{p}(\zeta;\tau)\phi_0\left(\frac{1+\mu}{2\sqrt{\mu}}\xi - \frac{1-\mu}{2\sqrt{\mu}}\zeta + w(\tau)\right)\right.$$
$$\left. - \bar{p}(\xi;\tau)\phi_0\left(\frac{1+\mu}{2\sqrt{\mu}}\zeta - \frac{1-\mu}{2\sqrt{\mu}}\xi + w(\tau)\right)\right\} |\xi - \zeta| \, d\zeta. \tag{67}$$

We introduce now an arbitrary smooth function $g(\xi)$ which vanishes sufficiently fast at $\pm\infty$ and after multiplying both sides of (67) by g we integrate on ξ from $-\infty$ to $+\infty$.

Introducing on the right hand side the new variable η, by the linear substitution

$$\xi = \zeta + \eta(2\sqrt{\mu}/(1+\mu))$$

we recast (67) in the form

$$\int_{-\infty}^{\infty} g(\xi)\frac{\partial \bar{p}}{\partial \tau} d\xi - \frac{w'(\tau)}{\sqrt{\mu}} \int_{-\infty}^{\infty} g(\xi)\frac{\partial \bar{p}}{\partial \xi} d\xi$$

$$= \frac{1}{\mu} \int_{-\infty}^{\infty} \bar{p}(\zeta;\tau) \int_{-\infty}^{\infty} \left[g\left(\zeta + \eta\frac{2\sqrt{\mu}}{1+\mu}\right) - g(\zeta)\right] |\eta|$$
$$\times \phi_0(\eta + w(\tau) + \zeta\sqrt{\mu}) d\eta \, d\zeta.$$

Expanding in powers of $\sqrt{\mu}$, up to and including $(\sqrt{\mu})^2$, we obtain after a few transformations

$$\int_{-\infty}^{\infty} g(\xi)\frac{\partial \bar{p}}{\partial \tau} d\xi - \frac{w'(\tau)}{\sqrt{\mu}} \int_{-\infty}^{\infty} g(\xi)\frac{\partial \bar{p}}{\partial \xi} d\xi$$
$$= \frac{2}{\sqrt{\mu}} \int_{-\infty}^{\infty} \bar{p}(\zeta;\tau) g'(\zeta) d\zeta K_1(w(\tau))$$
$$+ 2 \int_{-\infty}^{\infty} \bar{p}(\zeta;\tau) g'(\zeta)\zeta d\zeta \int_{-\infty}^{\infty} \phi_0'(\eta + w(\tau))\eta|\eta| d\eta$$
$$+ 2 \int_{-\infty}^{\infty} \bar{p}(\zeta;\tau) g''(\zeta) d\zeta K_2(w(\tau))$$
$$+ \text{terms of order } (\sqrt{\mu})^3 \text{ and higher}$$

where

$$K_n(w) = \int_{-\infty}^{\infty} \phi_0(\eta + w)\eta^n |\eta| d\eta.$$

It is clear that in order for the two terms of $1/\sqrt{\mu}$ to cancel we must have

$$w'(\tau) = 2K_1(w(\tau)). \tag{68}$$

This is just the dimensionless version of eq. (60). Thus $w(\tau)/\sqrt{\mu}$ with $w(\tau)$ obeying (68) is the *deterministic* law of decay of velocity in dimensionless terms.

Finally

$$\int_{-\infty}^{\infty} \phi_0'(\eta + w(t))\eta|\eta| d\eta = K_1'(w(t))$$

and we are led, upon integration by parts, using the arbitrariness of g to the Fokker – Planck equation

$$\frac{\partial \bar{p}}{\partial \tau} = -2K_1'(w(\tau))\frac{\partial}{\partial \xi}(\xi\bar{p}) + 2K_2(w(\tau))\frac{\partial^2 \bar{p}}{\partial \xi^2} \tag{69}$$

with $w(\tau)$ being a solution of (68). The initial condition for $\bar{p}$ is (66), i.e.

$$\bar{p}(\xi;0) = \delta(\xi) \tag{70}$$

and we must also prescribe the initial value of $w(\tau)$

$$w(0) = w_0. \tag{71}$$

The solution of eq. (69) with the initial conditions (70) and (71) is easily seen to be

$$\bar{p}(\xi;\tau) = \frac{1}{\sigma(\tau)\sqrt{2\pi}} \exp\left\{-\frac{\xi^2}{2\sigma^2(\tau)}\right\} \tag{72}$$

where $\sigma^2(\tau)$ satisfies the ordinary differential equation

$$\frac{d\sigma^2}{d\tau} = 4K_1'(w(\tau))\sigma^2 + 4K_2(w(\tau)) \tag{73}$$

and the initial condition

$$\sigma^2(0) = 0.$$

The expression for $\sigma^2(\tau)$ will depend (functionally) on $w(\tau)$ and hence on w_0. The dependence on w_0 is displayed by writing $\sigma^2(\tau; w_0)$.

If the initial distribution $p(u;0)$ is arbitrary one can write

$$p(u;0) = \frac{1}{\sqrt{\mu}} \int_{-\infty}^{\infty} p\left(\frac{w_0}{\sqrt{\mu}};0\right)\delta\left(u - \frac{w_0}{\sqrt{\mu}}\right) dw_0$$

which easily leads to the solution

$$p(u;\tau) = \frac{1}{\sqrt{\mu}} \int_{-\infty}^{\infty} p\left(\frac{w_0}{\sqrt{\mu}};0\right) \frac{1}{\sigma(\tau;w_0)\sqrt{2\pi}} \times \exp\left\{-\frac{(u - w(\tau))/\sqrt{\mu}}{2\sigma^2(\tau;w_0)}\right\} dw_0. \tag{74}$$

This form is deceptively simple; the highly complex form of $\sigma^2(\tau;w_0)$ is determined by eq. (73).

The fluctuation-dissipation theorem eq. (57) relating the friction and diffusion constants γ and D in the near-equilibrium theory is here replaced by eq. (73), relating their respective analogues $2K_1'(w(t))$ and $2K_2(w(t))$, time-dependent coefficients of the corresponding Fokker - Planck equation. The comparative weakness of the relation (73) reflects the lack of a simple universal thermodynamical theorem governing fluctuations far from equilibrium. Yet a description of fluctuations far from equilibrium in terms of nonstationary random processes, governed by Fokker - Planck

equations with time-dependent coefficients, may be possible for a large class of phenomena.

9. Linear nonequilibrium thermodynamics

The linear equation of motion of a particle in a fluid is one example from a class including Ohm's law, Fourier's law of heat conduction, the linearized Navier – Stokes equations, and a host of other linear laws of irreversible behavior near equilibrium. In appropriate thermodynamical variables these all turn out to be variations on eq. (1) – linear friction – in one or several variables. The thermodynamical setting was of course provided by Onsager (1931a, b).

In the sense that Langevin's theory of Brownian motion is the fluctuation theory for eq. (1), one gets a fluctuation theory for all these cases by rewriting Langevin's equation (or the equivalent Fokker – Planck equation) in appropriate variables. This (phenomenological) theory of linear nonequilibrium thermodynamics we outline below, in the version due to Onsager and Machlup (1953)*.

It is supposed that the thermodynamic state of a system close to equilibrium is described by a complete set of n extensive variables $(\alpha_1(t), \alpha_2(t), \ldots, \alpha_n(t)) = \boldsymbol{\alpha}(t)$, defined for convenience to vanish at equilibrium, $\boldsymbol{\alpha}^0 = 0$. The entropy is determined by the macroscopic state $\boldsymbol{\alpha}$; close to equilibrium, the entropy should be given by the first two terms in the Taylor expansion about equilibrium

$$S = S_0 - \tfrac{1}{2}k \sum_{ij}^{n} \alpha_i E_{ij} \alpha_j$$

where S_0 is the entropy of the macroscopic equilibrium state, and the entropy matrix E_{ij} is a symmetric positive-definite time independent matrix characteristic of the system. This is equivalent to the assumption that the fluctuations are Gaussian, for inverting the Boltzmann – Planck formula

* Equivalent theories were published around the same time by Hashitsume (1952), Takahasi (1952), and Yamamoto (1952, 1953). Callen and Greene (1952), Callen et al. (1952), and Green (1952) invented equivalent theories in frequency language (spectral densities instead of time evolution), developing a line of reasoning due to Nyquist (1928). Kirkwood (1946) seems to have first formulated the problem in the sufficiently general mathematical language of stationary Markov processes.

Here we follow the exposition and notation in Fox and Uhlenbeck (1970); the theory is outlined in textbooks on nonequilibrium thermodynamics such as that of de Groot and Mazur (1962), where detailed references can be found.

for the entropy gives

$$P(\alpha) = (\text{const}) \times e^{-(1/2)\sum \alpha_i E_{ij} \alpha_j} \tag{75}$$

for the instantaneous probability of occurrence of state α near equilibrium.

Generalized thermodynamical forces are defined:

$$X_i = \frac{\partial S}{\partial \alpha_i} = -k \sum_j E_{ij} \alpha_j$$

conjugate to the "fluxes" $J_i = \dot{\alpha}_i$. We consider systems satisfying linear phenomenological laws connecting the forces and fluxes

$$X_i = \sum_j R_{ij} J_j = \sum_j R_{ij} \dot{\alpha}_j = -k \sum E_{ij} \alpha_j \tag{76}$$

where R_{ij} is a real, nonsingular matrix, the real parts of whose eigenvalues are positive (corresponding to dissipative, rather than divergent systems). This generalizes eq. (1). If X is the voltage across a resistor R and J the current through it, eq. (76) is Ohm's law.

Writing L_{ij} for the inverse of R_{ij}, (76) becomes

$$\frac{d\alpha_i}{dt} + k \sum_{jk} L_{ij} E_{jk} \alpha_k = 0$$

or

$$\frac{d\alpha_i}{dt} + \sum_j G_{ij} \alpha_j = 0 \tag{77}$$

where

$$G_{ij} = k \sum L_{ik} E_{kj}.$$

For example we might have in mind linear electrical and thermal conduction in matter: in forces and fluxes language we would write

$$X_1 = R_1 J_1, \qquad X_2 = R_2 J_2 \tag{78, 79}$$

where in the first equation (Ohm's law) X_1 is the voltage across, and J_1 the current through, a resistance R_1, and in the second (Fourier's law of heat conduction), $X_2 = -(1/T)/(dT/dx)$ (one-dimensional temperature gradient): J_2 is the heat flux and R_2 is the heat resistance (reciprocal of the thermal conductivity).

In circumstances where both processes occur together, one has to expect reciprocal interactions

$$X_1 = R_{11} J_1 + R_{12} J_2, \qquad X_2 = R_{21} J_1 + R_{22} J_2$$

instead of (78, 79). Hence one provides for cross-terms R_{ij}, $i \neq j$ in (76), or G_{ij}, $i \neq j$ in (77). In this particular case Kelvin conjectured in 1854 on the basis of

experiment that $R_{12} = R_{21}$, which does not follow from any geometrical symmetry of the system, and offered something towards a proof. This was of course the first of the "reciprocal relations" established in Onsager (1931a, b).

The phenomenological variables must have essentially the properties discussed in connection with the Brownian particle. They must represent aggregates of a great many molecular variables, in such fashion that a central limit theorem applies. That they satisfy the phenomenological laws on the average already guarantees they possess most of the needed features, but is not quite sufficient. Precisely what constitutes an appropriate and "complete" set of variables for the theory is a subtle physical question, which is not solved. Onsager and Machlup (1953, p. 1506) list three guiding rules for choosing variables, which are a bit more restrictive than necessary. In particular, as we shall see in the next section on the Boltzmann equation, they need not be macroscopic. Their third rule, that the variables must be even functions of those molecular variables which are odd functions of time (like velocity) is needed for the proof of the reciprocal relations, but is too restrictive here. They develop the case of all odd functions (Machlup and Onsager 1953), but the Navier – Stokes equations and the Boltzmann equation have mixed odd and even character. Fox and Uhlenbeck (1970) provide the analysis for this completely general case.

The generalization of the Langevin equation of ordinary Brownian motion to the class of systems described, on the average, by eq. (77) is

$$\frac{d\alpha_i}{dt} + \sum_j G_{ij} \alpha_j = \tilde{F}_i(t) \tag{80a}$$

where $\tilde{F}_i(t)$ is the stationary purely random Gaussian process determined by

$$\langle \tilde{F}_i(t) \tilde{F}_j(t') \rangle = 2 Q_{ij} \delta(t' - t) \tag{80b}$$

where Q_{ij} is a symmetric, positive-definite, time-independent matrix.

Formally

$$\langle \alpha_i(t) \rangle^{\alpha(0)} = \sum_j [e^{-Gt}]_{ij} \alpha_j(0)$$

where again the conditional average is indicated, for $\alpha = \alpha(0)$ at $t = 0$, and where $[e^{-Gt}]_{ij}$ stands for

$$\sum_{n=0}^{\infty} \frac{(-t)^n}{n!} (G^n)_{ij}$$

the formal solution to the initial value problem for eq. (77), and

$$\langle \alpha_i(t)\alpha_j(t)\rangle^{\alpha(0)} = \sum_{kl} \{[e^{-Gt}]_{ik}[e^{-Gt}]_{jl}\alpha_k(0)\alpha_l(0) + M^{-1}_{ij} - [e^{-Gt}]_{ik}M^{-1}_{kl}[e^{-G^{T}t}]_{lj}\}. \tag{81}$$

G^{T} stands for the transpose of G, and M_{ij} is defined by

$$2Q_{ij} = \sum \left\{ G_{ik}M^{-1}_{kj} + M^{-1}_{ik}G^{T}_{kj} \right\}. \tag{82}$$

For the conditional distribution one gets

$$P(\alpha_0|\alpha;t) = \left(\frac{||v||}{(2\pi)^n}\right)^{1/2} \exp\left\{ -\frac{1}{2}\sum_{ijkl} V_{ij}(\alpha_i - [e^{-Gt}]_{ik}\alpha^0_k) \times (\alpha_j - [e^{-Gt}]_{jl}\alpha^0_l) \right\} \tag{83}$$

the probability that $\alpha(t) = \alpha$ given that $\alpha(0) = \alpha_0$, where V_{ij} is defined

$$V^{-1}_{ij} = M^{-1}_{ij} - \sum_{kl} [e^{-Gt}]_{ik}M^{-1}_{kl}[e^{-G^{T}t}]_{lj}. \tag{84}$$

The limit $t \to \infty$ gives the distribution function $W(\alpha)$; according to (84), $V_{ij} = M_{ij}$ in this limit, so one gets from (83)

$$W(\alpha) = \left(\frac{||M||}{(2\pi)^n}\right)^{1/2} e^{-(1/2)\sum \alpha_i M_{ij}\alpha_j}. \tag{85}$$

As we did before, we require that this be the Maxwell - Boltzmann distribution. Equation (75), properly normalized, is

$$W(\alpha) = \left(\frac{||E||}{(2\pi)^n}\right)^{1/2} e^{-(1/2)\sum \alpha_i E_{ij}\alpha_j} \tag{86}$$

so that the relation

$$E_{ij} = M_{ij} \tag{87}$$

is imposed, the generalization of the Einstein relation (37). Substituting (87) and (82) gives for (80b)

$$\langle \tilde{F}_i(t)\tilde{F}_j(t')\rangle = \delta(t'-t)\sum (G_{ik}E^{-1}_{kj} + E^{-1}_{ik}G^{T}_{kj}) \tag{88}$$

the fluctuation-dissipation theorem.

If we assume that the phenomenological variables are even functions of those molecular variables which are odd functions of time, i.e., that the phenomenological variables are time-reversal invariant (of course we assume the microscopic Hamiltonian is time-independent – we are studying stationary processes), then the joint distribution must satisfy

$$W(\alpha_1, \alpha_2; t) = W(\alpha_2, \alpha_1; t). \tag{89}$$

If these two quantities are calculated from (83) and (85) with (87) [recall that $W(\alpha, \alpha'; t) = W(\alpha)P(\alpha|\alpha'; t)$] and set equal, there results the identity

$$\sum_k E^{-1}_{ik}[e^{-G^{T}t}]_{kj} = \sum_k E^{-1}_{jk}[e^{-G^{T}t}]_{ki}$$

equating coefficients of t (the second term in the formal series for the exponentials) gives

$$\sum_k E^{-1}_{ik}G^{T}_{kj} = \sum_k E^{-1}_{jk}G^{T}_{ki}. \tag{90}$$

Using this and the symmetry of E_{ij} in the definition of G following eq. (77) gives

$$L_{ij} = L_{ji} \tag{91}$$

– the Onsager reciprocal relations.

Using this, the fluctuation-dissipation relation (88) can be written

$$\langle \tilde{F}_i(t)\tilde{F}_j(t')\rangle = 2L_{ij}\delta(t'-t). \tag{92}$$

The symmetry of L_{ij} implies that of R_{ij}; since it was assumed that the real parts of its eigenvalues are positive definite, the positivity of R_{ij} implies then that the eigenvalues are positive and real. Then the matrix $R^{1/2}_{ij}$ can be defined, which is necessarily symmetric, and we can define

$$\alpha'_i = \sum R^{1/2}_{ij}\alpha_j \tag{93}$$

whence (80a) becomes

$$\frac{d\alpha'_i}{dt} + k\sum_{jk} R^{-1/2}_{ij}E_{jk}R^{-1/2}_{kl}\alpha'_l = \sum_j R^{1/2}_{ij}\tilde{F}_j \tag{94}$$

and (92) can be written

$$\left\langle \sum_{kl} R^{1/2}_{il}\tilde{F}_l(t)R^{1/2}_{jk}F_k(t) \right\rangle = 2\delta_{ij}\delta(t'-t). \tag{95}$$

The matrix $\sum_{jk} R^{-1/2}_{ij}E_{jk}R^{-1/2}_{kl}$ is real, symmetric and positive definite, so that it can be diagonalized by an orthogonal matrix O_{ij}. Introducing the

transformed variables

$$\alpha_i'' = \sum_j O_{ij}\alpha_j' = \sum_{jk} O_{ij}R_{jk}^{1/2}\alpha_k \tag{96}$$

and denoting the eigenvalues λ_i, one has then (by definition of O_{ij})

$$\sum_{iklq} O_{ij}R_{jk}^{-1/2}E_{kl}R_{lq}^{-1/2}O_{qr}^{\mathrm{T}} = \lambda_i\delta_{ir}$$

so that (94) is transformed into

$$\frac{\mathrm{d}\alpha_i''}{\mathrm{d}t} + k\lambda_i\alpha_i'' = \sum_{jk} O_{ij}R_{jk}^{1/2}\tilde{F}_k. \tag{97}$$

Defining

$$\tilde{\theta}_i = \sum_{jk} O_{ij}R_{jk}^{1/2}\tilde{F}_k \tag{98}$$

transforms (95) into

$$\begin{aligned}\langle\tilde{\theta}_i(t)\tilde{\theta}_j(t')\rangle &= 2\sum_{kl} O_{ik}O_{jl}\delta_{kl}\delta(t'-t)\\ &= 2\delta_{ij}\delta(t'-t)\end{aligned} \tag{99}$$

by the definition of orthogonality. Thus the coupled equations (80) can be transformed into n independent Langevin equations

$$\frac{\mathrm{d}\alpha_i''}{\mathrm{d}t} + k\lambda_i\alpha_i'' = \tilde{\theta}_i \tag{100a}$$

$$\langle\tilde{\theta}_i(t)\tilde{\theta}_j(t')\rangle = 2\delta_{ij}\delta(t'-t). \tag{100b}$$

Since the transformations (93)–(98) are linear, the Gaussian character of the $\alpha_i''(t)$ process still follows from that assumed for $\tilde{F}(t)$, and so the considerations of the last section (namely Doob's theorem) again provide that it must be a Markov process.

What we have so far, then, is a recipe: Given linear phenomenological equations written in the form

$$\frac{\mathrm{d}a_i}{\mathrm{d}t} + \sum_j G_{ij}a_j = 0 \tag{101}$$

and the expression for entropy close to equilibrium in terms of the phenomenological variables $a_i(t)$, which is to say the entropy matrix E_{ij}, we reinterpret (101) as regression equations for the *average* behavior

$$\frac{\mathrm{d}\langle a_i\rangle}{\mathrm{d}t} + \sum G_{ij}\langle a_j\rangle = 0$$

and write instead of (101) the stochastic equations

$$\frac{\mathrm{d}a_i}{\mathrm{d}t} + \sum G_{ij}a_j = \tilde{F}_i(t) \tag{102}$$

where the stationary, purely random Gaussian process $\tilde{F}_i(t)$ is determined by G_{ij} and E_{ij}. Then it follows that the $a_i(t)$ can be written as superpositions of Uhlenbeck – Ornstein processes.

10. Fluctuations and the Boltzmann equation: phenomenological theory

The phenomenological theory of the preceding section, and in fact the theory of stationary Gaussian Markov processes, was developed for a finite, or at most denumerably infinite set of macroscopic variables. However it has been applied to situations where the indices run through a continuum; in particular to the electromagnetic field (Rytov 1953, and see Landau and Lifshitz 1960) and to the Boltzmann equation for a dilute gas* (Bixon and Zwanzig 1969, Montgomery 1969, Hinton 1970, Fox and Uhlenbeck 1970).

The Boltzmann equation

$$\frac{\partial f}{\partial t} = -\boldsymbol{v}\cdot\frac{\partial f}{\partial \boldsymbol{r}} + \int \mathrm{d}\boldsymbol{v}_1 \int \mathrm{d}\Omega g I(g,\theta)(f'f_1' - ff_1) \tag{103}$$

describes the evolution of the molecular distribution function $f(\boldsymbol{r}\boldsymbol{v};t)$, so defined that $f(\boldsymbol{r}\boldsymbol{v};t)\mathrm{d}\boldsymbol{r}\mathrm{d}\boldsymbol{v}$ represents the number of molecules with velocity $\boldsymbol{v}$ within $\mathrm{d}\boldsymbol{v}$, near point $\boldsymbol{r}$, within $\mathrm{d}\boldsymbol{r}$**. It provides a quite satisfactory description within a useful domain of validity (dilute gases not so violently perturbed that large variations on the scale of r_0, the range of molecular forces, occur), and it serves as a microscopic basis for the hydrodynamics of dilute gases. From it the Navier – Stokes equations can be derived, complete with formulas for the transport coefficients (viscosity, thermal conductivity, diffusion constant) in terms of molecular quantities. Macroscopic quantities like the hydrodynamical fluid velocity are obtained by averaging over the molecular distribution,

$$\boldsymbol{u}(\boldsymbol{r};t) = (1/n)\int f(\boldsymbol{r}\boldsymbol{v};t)\boldsymbol{v}\,\mathrm{d}\boldsymbol{v}$$

where n represents the local density.

* The idea of treating the Boltzmann equation as a phenomenological average equation and adding fluctuations seems to have been investigated first by Kadomtsev (1957), and was utilized by Abrikosov and Khalatnikov (1958).

** f_1 means $f(\boldsymbol{r}\boldsymbol{v};t)$, f' means $f(\boldsymbol{r}\boldsymbol{v}';t)$ and so forth, with reference to the four velocities $(\boldsymbol{v},\boldsymbol{v}_1)\leftrightarrow(\boldsymbol{v}_1',\boldsymbol{v}_1')$ characterizing a binary collision; $g = |\boldsymbol{v}-\boldsymbol{v}_1|$ is the relative velocity which in the collision turns over the angle θ. $I(g,\theta)$ is the scattering cross section and $\mathrm{d}\Omega$ the element of solid angle in the standard collision analysis. See e.g. Uhlenbeck and Ford (1963).

Yet the molecular distribution function itself represents an $(N-1)$-fold integral of the complete microscopic distribution so defined that

$$D_N(r_1 v_1, r_2 v_2, \ldots, r_N v_N; t)\mathrm{d}r_1 \mathrm{d}v_1 \ldots \mathrm{d}r_N \mathrm{d}v_N \tag{104}$$

is the density of states in which the N particles are found near points $r_1, \ldots, r_N$ with velocities $v_1, \ldots, v_N$, respectively, and so represents an average in its own right. The Boltzmann equation, then, has to be regarded as a kind of phenomenological equation in molecular phase space, and attention be given to the fluctuations in the molecular distribution about the average form predicted by the Boltzmann equation*.

Near equilibrium, where the molecular distribution differs only slightly from the Maxwell – Boltzmann distribution

$$f^{(0)}(v) = n_0 \left(\frac{m}{2\pi kT}\right)^{3/2} \mathrm{e}^{-mv^2/2kT} \tag{105}$$

(where n_0 and T represent equilibrium values of the density and temperature), one can set

$$f(rv;t) = f^{(0)}[1 + h(rv;t)]$$

in the Boltzmann equation and neglect terms quadratic in h compared to the linear terms. This gives the linearized Boltzmann equation

$$\frac{\partial h}{\partial t} + v \cdot \frac{\partial h}{\partial r} = \int \mathrm{d}v_1 \int \mathrm{d}\Omega g I(g,\theta) f_1^{(0)}[h' + h_1' - h - h_1]. \tag{106}$$

To determine the form of the fluctuating term according to phenomenological theory, we must write this in the canonical form of the last section. This is done in Fox and Uhlenbeck (1970), and we follow their notation in the following, which also appears in Logan and Kac (1976).

One defines the new function

$$a(rv;t) = f_{(0)}^{1/2}(v) h(rv;t)$$

and operators

$$A(rv, r'v') = f_{(0)}^{1/2}(v) v \cdot \operatorname{grad} \delta(r - r')(v - v')$$

$$S(rv, r'v') = f_{(0)}^{1/2}(v) f_{(0)}^{1/2}(v') K(v, v') \delta(r - r')$$

* Boltzmann himself initially thought eq. (103) was exact, and made it the basis of a "purely analytical, mechanical proof" of the second law of thermodynamics (the large object of his program in kinetic theory). As his H-theorem demonstrated, eq. (103) describes irreversible behavior. Yet the underlying distribution (104) is governed by the Liouville equation, which is completely time-reversible. This leads to contradictions, notably Loschmidt's paradox, which eventually led to the understanding that the function given by the Boltzmann equation describes the *average* behavior. See Klein (1973).

where $K(v, v')$ is the symmetric kernel in terms of which the collision operator in (106) is written in the Hilbert – Enskog fashion

$$\int \mathrm{d}v' f^{(0)}(v) K(v, v') h(rv; t)$$

so that the linearized Boltzmann equation is now written

$$\frac{\partial a(rv;t)}{\partial t} = -\int\int A(rv, r'v') a(r'v';t) \mathrm{d}r' \mathrm{d}v' - \int\int S(rv, r'v') a(r'v';t) \mathrm{d}r' \mathrm{d}v'. \tag{107}$$

If r and v may be conceived as continuous indices, this appears in the standard form of the average regression equations for a linear process. A and S possess odd and even time symmetry, respectively, through their respective dependence on odd and even functions of the molecular velocity v. Then one needs the entropy matrix. Boltzmann's expression

$$S(t) = -k \int\int f(rv;t) \ln f(rv;t) \mathrm{d}r \mathrm{d}v \tag{108}$$

gives for the entropy of a gas close to equilibrium

$$S(t) = S_0 - \tfrac{1}{2} k \int\int a^2(rv;t) \mathrm{d}r \mathrm{d}v \tag{109}$$

where S_0 characterizes the equilibrium state; terms beyond quadratic in h are neglected, and the fact that 1 and v^2 are eigenfunctions of the collision operator with vanishing eigenvalues is used. This gives by inspection

$$E(rv, r'v') = \delta(r - r')\delta(v - v').$$

Then the "Boltzmann – Langevin equation" turns out to be, in the original variables

$$\frac{\partial h}{\partial t} = -v \cdot \frac{\partial h}{\partial t} + \mathrm{L}h + \tilde{C}(rv;t) \tag{110a}$$

where L stands for the linearized collision operator in (106) and the analog of the random force is the stationary Gaussian process determined by

$$\tilde{C}(rv;t)\tilde{C}(r'v';t;) = 2K(v, v')\delta(r - r')\delta(t - t'). \tag{110b}$$

Fox and Uhlenbeck have shown, as we remarked in the first section, that the fluctuating hydrodynamical equations which follow from (110) by the Chapman – Enskog procedure are identical to those obtained by Landau and Lifshitz (1957), applying the method of irreversible thermodynamics to the linearized Navier – Stokes equations. The method is thus consistent on three levels. A really satisfactory justification, though, must come from microscopic theory.

11. Fluctuations and the Boltzmann equation: coarse-grain master equation

In the simple spatially homogeneous case, one can study the Boltzmann equation starting from the N-dimensional master equation for the evolution of an N-particle dilute gas treated as a random process or, as we do in the following, from the K-dimensional coarse-grain master equation of the system*.

One imagines the set of velocity states available to the N particles of a homogeneous dilute gas to be divided into K cells $i = 1, \ldots, K$, each of which corresponds to an interval $(v, v + \Delta v)$ small macroscopically but large enough to be populated by many particles; then the coarse-grained state of the system is specified by the complete set of occupation numbers $(n_1(t), n_2(t), \ldots, n_K(t)) = \boldsymbol{n}(t)$ where

$$\sum_{i=1}^{K} n_i(t) = N. \tag{111}$$

The idea, apparently first used by Nordsieck et al. (1940), is to treat the evolution of the system as a random process, to suppose that transitions from cell to cell occur from time to time according to a probabilistic law, so determining a vector stochastic process $\boldsymbol{n}(t)$.

For the dilute gas it is reasonable – but not patent – that in the limit $K \to \infty$, $N \to \infty$, $N/K \sim \bar{n}$ finite, $\boldsymbol{n}(t)$ is a K-dimensional vector Markov process with stationary transition probabilities. Thus the conditional probability of finding the system in state $\boldsymbol{n}$, given initial state $\boldsymbol{m}$

$$P(\boldsymbol{m}; t|\boldsymbol{n}; t + \tau) = P(\boldsymbol{m}|\boldsymbol{n}; \tau)$$

which describes the process, depends only on the difference in times, and the master distribution satisfies

$$P(\boldsymbol{n}; \tau) = \sum_{\boldsymbol{m}} P(\boldsymbol{m}; 0) P(\boldsymbol{m}|\boldsymbol{n}; \tau) \tag{112}$$

where the sum is taken over all sets of occupation numbers $(m_1, \ldots, m_K)$ consistent with (111). By the nature of the process $\boldsymbol{n}$ can change by only a small amount in an infinitesimal time interval, and so one can write

$$P(\boldsymbol{m}|\boldsymbol{n}; \Delta t) = \delta(\boldsymbol{m}|\boldsymbol{n}) + Q(\boldsymbol{m}|\boldsymbol{n})\Delta t + \mathrm{O}((\Delta t)^2) \tag{113}$$

* This approach is due to Siegert (1949). This kind of description was used by Tolman (1938) and Moyal (1949); and also more recently by Van Vliet (1964) to study fluctuations in electron-hole levels in solids. Quite recently Skorobogatov (1975a, b) has used it to study the Boltzmann equation along the lines of this section. Much of this section is taken from Logan and Kac (1976), with a few changes.

where $\delta(\boldsymbol{m}|\boldsymbol{n}) = 1$ or 0 according to whether $\boldsymbol{m} = \boldsymbol{n}$ or otherwise, and the transition matrix characterizing the process must satisfy

$$Q(\boldsymbol{m}|\boldsymbol{n}) \geq 0 \qquad \text{for } \boldsymbol{m} \neq \boldsymbol{n}$$

(positive probabilities), and

$$Q(\boldsymbol{m}|\boldsymbol{m}) = -\sum_{\boldsymbol{n} \neq \boldsymbol{m}}{}' Q(\boldsymbol{m}|\boldsymbol{n}) \tag{114}$$

(summation over all $\boldsymbol{m}$, $\boldsymbol{n}$ consistent with the restriction) because

$$\sum_{\boldsymbol{n}} P(\boldsymbol{m}|\boldsymbol{n}; \Delta t) = 1.$$

The conservation laws are implemented by requiring that $Q(\boldsymbol{m}|\boldsymbol{n}) \neq 0$ only for allowed transitions.

From (112) it follows that

$$\frac{\mathrm{d}P(\boldsymbol{n}; t)}{\mathrm{d}t} = \lim_{\Delta t \to 0} \sum_{\boldsymbol{m}} \frac{P(\boldsymbol{m}; t)P(\boldsymbol{m}|\boldsymbol{n}; t + \Delta t) - P(\boldsymbol{m}; t)P(\boldsymbol{m}|\boldsymbol{n}; t)}{\Delta t}$$

$$= \sum_{\boldsymbol{m}} P(\boldsymbol{m}; t) \lim_{\Delta t \to 0} \left[\frac{P(\boldsymbol{m}|\boldsymbol{n}; t + \Delta t) - P(\boldsymbol{m}|\boldsymbol{n}; t)}{\Delta t} \right]$$

using (113), this gives

$$\frac{\mathrm{d}P(\boldsymbol{n}; t)}{\mathrm{d}t} = \sum_{\boldsymbol{m}} P(\boldsymbol{m}; t) Q(\boldsymbol{m}|\boldsymbol{n}) \tag{115}$$

– the *master equation* (the Chapman–Kolmogorov equation for the process).

This approach is intrinsically limited to the spatially homogeneous case, for the treatment of the elementary events (the molecular collisions) as random transitions depends on the suppression – or averaging – of the position coordinates. Thus one proceeds from an N-particle dynamical problem to an N-particle stochastic process, where one follows each of the individual particle momenta but ignores the coordinates*. In the coarse-grained description employed here one goes further, lumping states together in discrete cells; sufficient "contraction" of this kind eventually

* The resulting Chapman–Kolmogorov equation or "fine-grain" master equation was used as a basis for the kinetic theory of the Boltzmann equation for a one-dimensional model gas in Kac (1956, 1959). The fine-grain master equation is used to calculate the covariance of the fluctuation process for an artificial two-state process in Kac (1973), and for the one dimensional model in Logan (1973; and forthcoming). Some further results in this line appear in McKean (1975). A nice elementary discussion of the master equation approach can be found in Dresden (1962).

destroys the Markov property, but one proceeds on the assumption that it holds in the indicated limit. (No adequate derivation of the detailed master equation from the Liouville equation is yet available, nor any useful theorems for the coarse-grain limit, nor any general information on the conditions under which the Markov property is inherited on contraction or projection.)*

Following Siegert, we take the molecules as point particles, and for the elementary collision law Boltzmann's *Stosszahlansatz*

$$P(ij \to kl)\mathrm{d}t = \alpha_{ij}^{kl} m_i m_j \mathrm{d}t \tag{116}$$

where m_i, m_j are the occupation numbers of cells i,j before the collision and α_{ij}^{kl} is a constant, determined by the intermolecular potential. The probability of a collision between particles in states i,j is proportional to the product of the numbers of particles in the respective states; the probability per collision of a scattering into k, l is proportional to the "cross section" α_{ij}^{kl} [evidently $O(1/N)$], which is non-negative for $ij \neq kl$ and obviously satisfies

$$\alpha_{ij}^{kl} = \alpha_{ji}^{kl} = \alpha_{ij}^{lk}$$

and

$$\alpha_{ii}^{kl} = \alpha_{ij}^{kk} = 0.$$

In addition, for central forces one has "microreversibility"

$$\alpha_{ij}^{kl} = \alpha_{kl}^{ij} \tag{117}$$

equality of the cross sections for forward and inverse or "restituting" collisions**. If ε_i is the energy associated with state i, then one must have

$$\varepsilon_i + \varepsilon_j = \varepsilon_k + \varepsilon_l, \qquad \text{if } \alpha_{ij}^{kl} \neq 0 \tag{118}$$

i.e., for any allowed transition: conservation of energy in a collision.

The transition matrix to go in the master equation (115), embodying these assumptions, is

$$Q(\boldsymbol{m}|\boldsymbol{n}) = \tfrac{1}{4}\Bigg[\sum_{ij \neq kl}{}' \alpha_{ij}^{kl} m_i m_j \delta_{n_i m_i - 1} \delta_{n_j m_j - 1} \delta_{n_k m_k + 1} \delta_{n_l m_l + 1} \times \prod_{r \neq ijk}{}' \delta_{n_r m_r} + \sum_{ij} \alpha_{ij}^{ij} m_i m_j \prod_r \delta_{n_r m_r} \Bigg] \tag{119}$$

* Brout (1956) deals with a scheme to derive the fine-grain master equation from the Liouville equation by a systematic expansion in "collision configurations" of successive combinatorial orders. The argument is outlined in Kac (1959).

**For spherically symmetric molecules interacting through central forces, this obtains, and it is all we require of the intermolecular forces. In general, however, the inverse collisions may not even exist; see Tolman (1938), sects. 38–42, and Blatt and Weisskopf (1952).

since the collision $ij \to kl$ leaves all occupation numbers unchanged except $n_k = m_k + 1$, $n_i = m_i + 1$, $n_l = m_l - 1$, and $n_j = m_j - 1$ (the factor $\frac{1}{4}$ accounts for double counting of i, j and k, l). We define

$$\alpha_{ij}^{ij} = -\sum_{ij \neq kl}{}' \alpha_{ij}^{kl}$$

so that the unrestricted sum vanishes,

$$\sum_{kl} \alpha_{ij}^{kl} = 0 \tag{120}$$

and gives

$$\sum_{\boldsymbol{n}} P(\boldsymbol{m}|\boldsymbol{n}; \Delta t) = 1$$

the probability that the system, initially in state $\boldsymbol{m}$, is found in *some* state after time Δt.

By definition the stationary distribution satisfies the equation

$$\sum_{\boldsymbol{m}} W(\boldsymbol{m}) P(\boldsymbol{m}|\boldsymbol{n}; t) = W(\boldsymbol{n})$$

or equivalently

$$\sum_{\boldsymbol{m}} W(\boldsymbol{m}) Q(\boldsymbol{m}|\boldsymbol{n}) = 0 \tag{121}$$

($W(\boldsymbol{m})$ is thus a left eigenvector of $Q(\boldsymbol{m}|\boldsymbol{n})$ belonging to the eigenvalue 0). In summing over states $\boldsymbol{m}$, the restrictions

$$\sum_{i=1}^{K} m_i = N \tag{122}$$

$$\sum_{i=1}^{K} m_i \varepsilon_i = E = N\varepsilon \tag{123}$$

must be obeyed [(123) follows from (118)]. We assume the stationary distribution is unique. For a Markov process with a finite number of states (Markov chain) this will be the case when each allowable state of the system is reachable from every state.

The stationary distribution can be obtained as the limit of the conditional probability $P(\boldsymbol{m}|\boldsymbol{n}; \tau)$ for $\tau \to \infty$, and is easily found to be

$$W(\boldsymbol{n}) = (\text{const}) \times \frac{N!}{n_1! n_2! \ldots n_K!} \tag{124}$$

subject to (122) and (123). Because of the constraints, the equilibrium average occupations are given by the canonical distribution

$$\bar{n}_i = A\mathrm{e}^{-\beta\varepsilon_i} \tag{125}$$

in the limit $N \to \infty$. The constants A and β are determined by the constraint equations

$$A\sum_i e^{-\beta\varepsilon_i} = N, \qquad A\sum_i \varepsilon_i e^{-\beta\varepsilon_i} = N\varepsilon.$$

Boltzmann obtained his formula by maximizing log $W(\boldsymbol{m})$ under the constraints, and much later Darwin and Fowler used the method of steepest descent to rederive it; here we give yet another derivation. Let λ and μ be chosen so that

$$\sum_{\boldsymbol{m}} \left(\sum m_i\right) W(\boldsymbol{m}) e^{\lambda \sum m_i} e^{\mu \sum \varepsilon_i m_i} = N$$

and

$$\sum_{\boldsymbol{m}} (m_i\varepsilon_i) W(\boldsymbol{m}) e^{\lambda \sum m_i} e^{\mu \sum \varepsilon_i m_i} = E.$$

The summation on $\boldsymbol{m}$ is now unrestricted so that

$$\sum_{\boldsymbol{m}} W(\boldsymbol{m}) e^{\lambda\sum m_i} e^{\mu\sum \varepsilon_i m_i} = \frac{C}{N!}\left(\sum_i e^{\lambda + \mu\varepsilon_i}\right)^N$$

and the equations determining λ and μ become

$$\sum_i e^{\lambda + \mu\varepsilon_i} = 1, \qquad \sum_i \varepsilon_i e^{\lambda + \mu\varepsilon_i} = \varepsilon.$$

Having chosen λ and μ so that the constraints are satisfied on the *average* with respect to the artificially introduced probability distribution

$$C' W(\boldsymbol{m}) e^{\lambda \sum m_i} e^{\mu \sum \varepsilon_i m_i}$$

(C' being the normalizing constant) we can now compute the joint probability distribution of

$$\sum m_i/N \quad \text{and} \quad \sum \varepsilon_i m_i \Big/ N$$

again with respect to the artificial distribution. To do this, we calculate the joint characteristic function

$$\exp(i\xi \sum m_j/N + i\eta \sum m_j\varepsilon_j/N)$$

$$= \frac{\left(\sum_j \exp[\lambda + i\xi/N + \varepsilon_j(\mu + i\eta/N)]\right)^N}{\left(\sum_j \exp(\lambda + \mu\varepsilon_j)\right)^N}$$

which in the limit $N \to \infty$ becomes

$$e^{i\xi + i\varepsilon\eta}.$$

That is, the joint distribution of $(1/N)\sum m_j$ and $(1/N)\sum m_j\varepsilon_j$ becomes, in the limit $N \to \infty$, infinitely sharp about the mean values 1 and ε. For the purposes of calculating mean values it is thus permissible to replace the distribution *with* constraints by the artificial one without them.

11.1 *The Boltzmann equation*

From the master equation (115) with the transition matrix (119) one can, by summing over all but 1, 2, ..., of the occupation numbers, obtain equations for $\langle n_i(t)\rangle$, $\langle n_i(t)n_j(t)\rangle$, ..., etc.*, where the symbols indicate an average over all possible complete configurations of the system ("master states"). In particular, one gets

$$\frac{d\langle n_s\rangle}{dt} = \tfrac{1}{2}\sum_{ijk} \alpha_{ij}^{ks}\langle n_i n_j\rangle$$
$$= \tfrac{1}{2}\sum_{\substack{ijk \\ ij \neq ks}}{}' \alpha_{ij}^{ks}\langle n_i n_j\rangle + \tfrac{1}{2}\sum_k \alpha_{ks}^{ks}\langle n_k n_s\rangle \qquad (126)$$

which can be written in view of (120)

$$\frac{d\langle n_s\rangle}{dt} = \tfrac{1}{2}\sum_{ij \neq ks}{}' \alpha_{ij}^{ks}\langle n_i n_j\rangle - \tfrac{1}{2}\sum \alpha_{ks}^{ij}\langle n_k n_s\rangle$$

similarly one gets an equation for $(d/dt)\langle n_i n_j\rangle$ in terms of $\langle n_i n_j n_k\rangle$ and so on, corresponding to the BBGKY hierarchy. If one neglects the correlations $\langle (n_i - \langle n_i\rangle)(n_j - \langle n_j\rangle)\rangle$ in comparison with the product of the averages

* The most convenient way is to define the generating function

$$\phi(z/t) = \sum_{n} \sum_{i=1}^{K} z_i^{n_i} P(\boldsymbol{n}; t)$$

then the master equation (115) is equivalent to

$$\frac{\partial\phi(z;t)}{\partial t} = \sum_{m,n} P(\boldsymbol{m};t) Q(\boldsymbol{m}|\boldsymbol{n}) \sum_{i=1}^{K} z_i^{n_i}$$

and, for example

$$\langle n_1\rangle = z_1\left(\frac{\partial\phi}{\partial z_1}\right)_{z_2 = z_3 = \ldots = z_K = 1}$$

See Mathews et al. (1960) and Siegert (1949), appendix III.

$\langle n_i\rangle\langle n_j\rangle$, which one justifies on the grounds that the former is relatively $O(1/N)$, one obtains the coarse-grain Boltzmann equation

$$d\langle n_s\rangle/dt = \tfrac{1}{2}\sum_{ij\neq ks}' \alpha_{ij}^{ks}(\langle n_i\rangle\langle n_j\rangle - \langle n_k\rangle\langle n_s\rangle) \tag{127}$$

in which the microreversibility symmetry (117) has been exploited; here, and subsequently, summation is directed over the running indices ijk. This discrete version of the equation appears in Boltzmann's *Lectures on Gas Theory*. If, following Boltzmann, one defines

$$H = \sum_s \langle n_s\rangle \ln\langle n_s\rangle$$

it is easy to show that

$$dH/dt = \tfrac{1}{2}\sum \alpha_{ij}^{ks}(\langle n_i\rangle\langle n_j\rangle - \langle n_k\rangle\langle n_s\rangle) \times \ln\frac{\langle n_k\rangle\langle n_s\rangle}{\langle n_i\rangle\langle n_j\rangle} \leq 0 \tag{128}$$

the H-theorem. Note that (118) is not required; only the nonnegativity of the α_{ij}^{kl} is needed. It follows that if there is an equilibrium state – one characterized by $dH/dt = 0$ – then

$$\alpha_{ij}^{kl} = 0 \qquad \text{unless} \quad \bar{n}_i\bar{n}_j = \bar{n}_k\bar{n}_l \tag{129}$$

where the bars (here and in the following) designate equilibrium averages.

For small departures from equilibrium, one can write

$$\langle n_i\rangle = \bar{n}_i(1 + h_i)$$

substitute in (127), neglect quadratic terms compared to h, and obtain, making use of (129), the linearized Boltzmann equation

$$d\langle n_s\rangle/dt = \tfrac{1}{2}\sum_{ijk} \alpha_{ij}^{ks}\bar{n}_i\bar{n}_j[h_i + h_j - h_k - h_s] \tag{130}$$

which can be written*

$$\frac{d\langle y_s\rangle}{dt} + \frac{1}{n_s}\sum_\sigma g_{s\sigma}\langle y_\sigma\rangle = 0 \tag{131}$$

where $y_s = n_s/\bar{n}_s$ and (note $\sum_\sigma g_{s\sigma} = 0$)

$$g_s = -\sum_{kl}\alpha_{\sigma j}^{ks}\bar{n}_\sigma\bar{n}_j + \tfrac{1}{2}\sum_{ij}\alpha_{ij}^{\sigma s}\bar{n}_i\bar{n}_j + \tfrac{1}{2}\delta_{s\sigma}\sum_{ijk}\alpha_{ij}^{ks}\bar{n}_i\bar{n}_j\,. \tag{132}$$

It is convenient to define $v_s = n_s/N$ ($\bar{v}_s = \bar{n}_s/N$) and $\alpha_{ij}^{kl} = N\alpha_{ij}^{kl}$ which are, like y_s, $O(1)$.

* The corresponding eq. (4.29) in Logan and Kac (1976) is misprinted.

The linearized Boltzmann equation with fluctuations from the thermodynamical theory [eqs. (110)] can then be written, in this formalism

$$\frac{d\langle y_s\rangle}{dt} + \frac{1}{\bar{v}_s}\sum_\sigma g_{s\sigma}\langle y_\sigma\rangle = \tilde{c}_s(t) \tag{133a}$$

with $\tilde{c}_s(t)$ the stationary, purely random Gaussian process determined by $\langle\tilde{c}_s(t)\rangle = 0$ and

$$\langle\tilde{c}_s(t)\tilde{c}_\sigma(t')\rangle = (2/\bar{v}_s\bar{v}_\sigma)g_{s\sigma}\delta(t' - t) \tag{133b}$$

all, again, in contemplation of the limit $N \to \infty$.

11.2 *The Fokker–Planck equation: fluctuations around equilibrium*

Now we are ready to use the master equation to *derive* the needed features of the process

$$Y_s(t) = \frac{n_s(t) - \bar{n}_s}{\bar{v}_s\sqrt{N}}\,. \tag{134}$$

By assumption $\boldsymbol{n}(t) = (n_1(t), \ldots, n_K(t))$ is a K-dimensional vector Markov process, hence

$$\frac{n_1(t) - \bar{n}_1}{\sqrt{N}}, \ldots, \frac{n_K(t) - \bar{n}_K}{\sqrt{N}}$$

is also; in the limit $N \to \infty$ these ought still to be vector Markov processes, in continuous variables*. Since, as is demonstrated in the following, the process is continuous in the requisite technical sense, it is possible to derive from the master equation the corresponding Fokker–Planck equation, which turns out to be equivalent to (133).

Since the explicit limit $N \to \infty$ is essential, it is convenient to work in the manifestly $O(1)$ quantities introduced above, in terms of which the Fokker–Planck equation is

$$\frac{dP(\boldsymbol{x};t)}{dt} = -\sum_s\frac{\partial}{\partial x_s}[A_s(\boldsymbol{x})P(\boldsymbol{x})] + \frac{1}{2}\sum_{s\sigma}\frac{\partial^2}{\partial x_s\partial x_\sigma}[B_{s\sigma}(\boldsymbol{x})P(\boldsymbol{x})]. \tag{135}$$

What is required are the values of the coefficients in (135) and a demonstration that the higher coefficients $C_{s\sigma p}$, etc., vanish in the indicated limit.

* Since the limit $K \to \infty$ is also required, we are considering the continuous limit of an infinite-dimensional random process; the (possibly delicate) questions connected with the rigorous justification of the various limiting operations will not be considered.

Recall from sect. 6 that the coefficients are defined as limits of conditional averages: First of all

$$A_s(x) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \langle Y_s - X_s \rangle = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \frac{\langle n_s - m_s \rangle}{\bar{\nu}_s \sqrt{N}} \tag{136}$$

where what is meant is the conditional average

$$\langle Y_s(t + \Delta t) - X_s(t) | X(t) = x \rangle$$

which is equal to

$$\langle Y_s(\Delta t) - X_s(0) | X(0) = x \rangle$$

since the process is stationary, and

$$X_s = (1/\bar{\nu}_s \sqrt{N})[m_s(t) - \bar{m}_s].$$

This can be obtained from (113)

$$\begin{aligned} A_s(x) = & \lim_{\Delta t \to 0} \frac{1}{\Delta t} \frac{1}{\bar{\nu}_s \sqrt{N}} \\ & \times \sum_{n} \{[\delta(m|n) + \Delta t Q(m|n)](n_s - m_s)\} \\ = & \frac{1}{\bar{\nu}_s \sqrt{N}} \sum_{n} [Q(m|n)(n_s - m_s)] \\ = & \frac{1}{\bar{\nu}_s \sqrt{N}} \sum_{n} Q(m|n) n_s \end{aligned} \tag{137}$$

since according to (114)

$$\sum_{n} Q(m|n) = 0.$$

Substituting the transition matrix (119) gives

$$\begin{aligned} \sum_{n} Q(m|n) n_s = (1/4N) \Bigg\{ & \sum_{ijkl=s}{}' \alpha_{ij}^{kl} m_i m_j m_s \\ & + \sum_{jkl} \alpha_{sj}^{kl} m_s m_j (m_s - 1) + \sum_{ijkl} \alpha_{is}^{kl} m_i m_s (m_s - 1) \\ & + \sum_{ijl} \alpha_{ij}^{sl} m_i m_j (m_s + 1) = \sum_{ijk} \alpha_{ij}^{ks} m_i m_j (m_s + 1) \Bigg\}. \end{aligned} \tag{138}$$

The cubic terms can be grouped together and written

$$\sum_{ijkl} \alpha_{ij}^{kl} m_i m_j m_s$$

where the range of summation is now unrestricted; on account of (120) this vanishes. Using the index symmetry, (138) can then be written

$$\sum_{n} Q(m|n) n_s = (1/N) \left(-\tfrac{1}{2} \sum \alpha_{sj}^{kl} m_s m_j + \tfrac{1}{2} \sum \alpha_{ij}^{sl} m_i m_j \right). \tag{139}$$

Supposing the system is near equilibrium, we write

$$m_i = \bar{m}_i (1 + x_i / \sqrt{N}). \tag{140}$$

To first order in the small quantity we get then, making use of (129)

$$\sum_{n} Q(m|n) n_s = \frac{1}{2} \left[\frac{1}{2\sqrt{N}} \sum_{ijk} \alpha_{ij}^{ks} \bar{m}_i \bar{m}_j (x_i + x_j - x_s - x_k) \right]$$

and we have finally for (137)

$$A_s(x) = \frac{1}{2\bar{\nu}_s} \sum \alpha_{ij}^{ks} \nu_i \nu_j (x_i + x_j - x_s - x_k) = \frac{-1}{\bar{\nu}_s} \sum_{\sigma} g_{s\sigma} x_\sigma \tag{141}$$

where

$$g_{s\sigma} = -\sum_{jk} \alpha_{\sigma j}^{ks} \bar{\nu}_\sigma \bar{\nu}_j + \tfrac{1}{2} \sum_{ij} \alpha_{ij}^{\sigma s} \bar{\nu}_i \bar{\nu}_j + \tfrac{1}{2} \delta_{s\sigma} \sum_{ijk} \alpha_{ij}^{ks} \bar{\nu}_i \bar{\nu}_j \tag{142}$$

[the zeroth-order terms cancel on application of (129)].

The second coefficient

$$\begin{aligned} B_{s\sigma}(x) &= \lim_{\Delta t \to 0} (1/\Delta t) \langle (Y_s - X_s)(Y_\sigma - X_\sigma) \rangle \\ &= \lim_{\Delta t \to 0} (1/N \bar{\nu}_s \bar{\nu}_\sigma) \langle (n_s - m_s)(n_\sigma - m_\sigma) \rangle \end{aligned} \tag{143}$$

can be calculated in very much the same way (see Logan and Kac 1976, sect. 4C), with the result

$$\begin{aligned} B_{s\sigma}(x) &= \frac{1}{\bar{\nu}_s \bar{\nu}_\sigma} \left(-2 \sum \alpha_{\sigma j}^{ks} \bar{\nu}_\sigma \bar{\nu}_j + \sum \alpha_{ij}^{\sigma s} \bar{\nu}_i \bar{\nu}_j + \delta_{s\sigma} \sum \alpha_{ij}^{ks} \bar{\nu}_i \bar{\nu}_j \right) \\ &= (2/\bar{\nu}_s \bar{\nu}_\sigma) g_{s\sigma} \end{aligned} \tag{144}$$

Then one can go on to show that

$$C_{s\sigma\rho} = 0 \tag{145}$$

and likewise all the higher coefficients vanish.

Thus we obtain the Fokker–Planck equation with

$$A_s(x) = (-1/\bar{\nu}_s) \sum_{\sigma} g_{s\sigma} x_\sigma \tag{146}$$

and

$$B_{s\sigma}(x) = (2/\bar{v}_s\bar{v}_\sigma)g_{s\sigma}. \tag{147}$$

This can be shown, following familiar arguments, to be equivalent to the stochastic equation

$$\frac{dy_s}{dt} + \frac{1}{\bar{v}_s}\sum_\sigma g_{s\sigma}y_\sigma = \tilde{c}_s(t) \tag{148}$$

where again

$$g_{s\sigma} = -\sum_{jk}\alpha^{ks}_{\sigma j}\bar{v}_\sigma\bar{v}_j + \tfrac{1}{2}\sum\alpha^{\sigma s}_{ij}\bar{v}_i\bar{v}_j + \delta_{s\sigma}\sum_{ijk}\alpha^{ks}_{ij}\bar{v}_i\bar{v}_j \tag{149}$$

and $\tilde{c}_s(t)$ is the stationary, purely random Gaussian process defined by

$$\langle\tilde{c}_s(t)\rangle = 0$$

and

$$\langle\tilde{c}_s(t)\tilde{c}_s(t')\rangle = (2/\bar{v}_s\bar{v}_\sigma)g_{s\sigma}\delta(t'-t) \tag{150}$$

the linearized Boltzmann equation with fluctuations (133). (The equivalence is easily established on the observation of the equality of the complete set of moments determined in the two cases. Recall that the vanishing of the higher coefficients corresponds to the Gaussian property; the existence and character of the first two are equivalent to (133)).

11.3 *The Fokker–Planck equation: fluctuations away from equilibrium*

By the means just elaborated one can derive a Fokker–Planck equation for the fluctuation process away from equilibrium, where the mean occupation numbers are the time-dependent averages $\langle n_i(t)\rangle$ satisfying the nonlinear Boltzmann equation (127), which can as well be written

$$d\langle v_s\rangle/dt = \tfrac{1}{2}\sum\alpha^{si}_{ij}\langle v_i\rangle\langle v_j\rangle - \tfrac{1}{2}\sum\alpha^{kl}_{sj}\langle v_s\rangle\langle v_j\rangle \tag{151}$$

making use of (117) (recall $v_i = n_i/N$). It is convenient for the purpose to redefine the random variables

$$x_s(t) = [m_s(t) - \langle m_s(t)\rangle]\sqrt{N} \tag{152a}$$

$$y_s(t) = [n_s(t) - \langle n_s(t)\rangle]/\sqrt{N} \tag{152b}$$

evidently O(1).

First of all we require

$$A_s(x) = \lim_{\Delta t\to 0}(1/\Delta t)\langle y_s(t+\Delta t) - y_s(t)\rangle$$

where again the conditional average

$$\langle y_s(t+\Delta t) - y_s(t)|\boldsymbol{y}(t) = \boldsymbol{x}\rangle$$

is intended. This we can write

$$\lim_{\Delta t\to 0}\frac{1}{\Delta t}\left(\frac{1}{\sqrt{N}}\langle n_s(t+dt) - n_s(t)|\boldsymbol{n}(t) = \boldsymbol{m}\rangle - \frac{1}{N}\frac{\Delta n_s}{\Delta t}\Delta t\right)$$
$$= \frac{1}{\sqrt{N}}\left(-\frac{1}{2}\frac{1}{N}\sum\alpha^{kl}_{sj}(\langle m_j\rangle + \sqrt{N}\,x_j)(\langle m_s\rangle + \sqrt{N}\,x_s)\right.$$
$$\left. + \frac{1}{2}\frac{1}{N}\sum\alpha^{sl}_{ij}(\langle m_i\rangle + \sqrt{N}\,x_i)(\langle m_j\rangle + \sqrt{N}x_j) - \frac{d\langle n_s\rangle}{dt}\right).$$

Ignoring the terms of order $1/\sqrt{N}$ in contemplation of the limit $N\to\infty$, what remains can be written

$$\sqrt{N}\left(-\frac{1}{2}\sum\alpha^{kl}_{sj}\langle v_s\rangle\langle v_j\rangle + \frac{1}{2}\sum\alpha^{sl}_{ij}\langle v_i\rangle\langle v_j\rangle - \frac{d\langle n_s\rangle}{dt}\right)$$
$$-\frac{1}{2}\sum\alpha^{kl}_{sj}(\langle v_j\rangle x_s + \langle v_s\rangle x_j) + \frac{1}{2}\sum\alpha^{sl}_{ij}(\langle v_i\rangle x_j + \langle v_j\rangle x_i)$$

the quantity in large parentheses, of order $\sqrt{N}$, vanishes on substitution of the Boltzmann equation (151) for the time derivative, and what remains [O(1)] can be rewritten, so yielding

$$A_s(x) = -\sum_\sigma\gamma_{s\sigma}x_\sigma \tag{153}$$

where

$$\gamma_{s\sigma} = -\sum\alpha^{sl}_{\sigma j}\langle v_j\rangle + \tfrac{1}{2}\sum\alpha^{s\sigma}_{ij}\langle v_s\rangle + \tfrac{1}{2}\delta_{s\sigma}\sum\alpha^{kl}_{sj}\langle v_j\rangle.$$

Then we need

$$B_{s\sigma}(x) = \lim_{\Delta t\to 0}(1/\Delta t)\langle[y_s(t+\Delta t) - y_s(t)]$$
$$\times[y_\sigma(t+\Delta t) - y_\sigma(t)]|\boldsymbol{y}(t) = \boldsymbol{x}\rangle.$$

Proceeding in the same way, we obtain

$$B_{s\sigma} = \tfrac{1}{2}\sum\alpha^{s\sigma}_{ij}\langle v_j\rangle\langle v_i\rangle + \tfrac{1}{2}\sum\alpha^{kl}_{s\sigma}\langle v_s\rangle\langle v_\sigma\rangle$$
$$-\sum\alpha^{sl}_{\sigma j}\langle v_\sigma\rangle\langle v_j\rangle - \sum\alpha^{\sigma l}_{sj}\langle v_s\rangle\langle v_j\rangle. \tag{154}$$

The calculations for the higher terms of the preceding case carry over almost unchanged, so we have again $C_{s\sigma\rho} = 0$, etc., and thus the Fokker–Planck equation (135), but now with time-dependent coefficients, since the $\langle v_i\rangle$ – solutions of the Boltzmann equation – do of course vary with time. This result was also obtained by Van Kampen (1974).

The Fokker–Planck equation just obtained describes the fluctuations, arbitrarily far from equilibrium, about the means, whose course is governed by the Boltzmann equation; indeed the latter serves in the derivation. This may be contrasted with the character of the calculation in the previous section, in which no use is made of the linearized Boltzmann equation, which governs the means near equilibrium, but essential use is made of certain properties of the equilibrium distribution, around which the actual values fluctuate. Here, no use is made of the character of the equilibrium state; indeed nowhere need it be assumed that such a state exists. As in the near-equilibrium case the variance of the fluctuation process is determined by the mean (which is itself determined by the nonlinear Boltzmann equation); thus an important feature of the usual fluctuation-dissipation theorem is preserved.

It may be noticed that in both instances the derivation differs somewhat from the more usual case, in that the higher moments vanish by virtue of the limit $N \to \infty$, which induces vanishing coefficients of terms of first order in the infinitesimal dt, which would otherwise contribute. The limit $K \to \infty$ is not explicitly required in any of the foregoing, but serves to rationalize the Markov assumption for the K-dimensional process.

References

Abrikosov, A.A. and I.M. Khalatnikov, 1958, Zh. Eksp. Teor. Fiz. **34**, 198 [Sov. Phys. JETP 7, 135].

Bixon, M. and R. Zwanzig, 1969, Phys. Rev. **187**, 267.

Blatt, J.M. and V.F. Weisskopf, 1952, Theoretical nuclear physics (Wiley, New York) sect. X2E.

Brout, R., 1956, Physica **22**, 509.

Callen, H.B. and R.F. Greene, 1952, Phys. Rev. **86**, 702.

Callen, H.B., M.L. Barasch and J.L. Jackson, 1952, Phys. Rev. **88**, 1387.

Chandrasekhar, S., 1943, Rev. Mod. Phys. **15**, 1.

Cohen, E.G.D., 1969, The kinetic theory of dilute gases, in Transport phenomena in fluids (H.J. Hanley, ed.) (Dekker, New York) ch. 6.

Doob, J.L., 1942, Ann. Math. (Princeton) **43**, 351.

Doob, J.L., 1944, Ann. Math. Stat. **15**, 229.

Doob, J.L., 1953, Stochastic processes (Wiley, New York).

Dresden, M., 1962, A study of models in non-equilibrium statistical mechanics, in Studies in statistical mechanics, vol. 1 (J.de Boer and G.E. Uhlenbeck, eds.) (North-Holland, Amsterdam) p. 303.

Ehrenfest, P., 1907, Phys. Z. **8**, 311 (reprinted in Klein 1959).

Ehrenfest, P., 1912, The conceptual foundations of the statistical approach in mechanics, translated by M.J. Moravcsik (Cornell University Press, Ithaca, New York) 1959. Originally published in Encycklopädie der mathematischen wissenschaften IV, 2 (6), (B.G. Teubner, Leipzig).

Ehrenfest, P. and T. Ehrenfest, 1906, Math.-Naturw. Blätter **3** (reprinted in Klein 1959).

Einstein, A., 1905, Ann. Physik **17**, 549.

Einstein, A., 1906, Ann. Physik **19**, 371.

Feller, W., 1950, An introduction to probability theory and its applications, vol. I (Wiley, New York) first and subsequent editions.

Feller, W., 1966, An introduction to probability theory and its applications, vol. II (Wiley, New York).

Fox, R. and M. Kac, 1977, Biosystems **8**, 187.

Fox, R. and G.E. Uhlenbeck, 1970, Phys. Fluids **13**, 1893.

Gerlach, W., 1927, Naturwiss. **15**, 15.

Green, M.S., 1952, J. Chem. Phys. **20**, 1281.

de Groot, S.R. and P. Mazur, 1962, Non-equilibrium thermodynamics (North-Holland, Amsterdam).

Hashitsume, N., 1952, Prog. Theor. Phys. **8**, 461.

Hess, F.G., 1954, Amer. Math. Monthly **61**, 323.

Hinton, F.L., 1970, Phys. Fluids **13**, 857.

Kac, M., 1947, Amer. Math. Monthly **54**, 369.

Kac, M., 1956, Foundations of kinetic theory, in Proceedings of the third Berkeley symposium on mathematical statistics and probability, vol. 3 (Neyman, J. ed.) (University of California Press, Berkeley) p. 171.

Kac, M., 1959, Probability and related topics in physical sciences (Interscience, New York).

Kac, M., 1973, Some probabilistic aspects of the Boltzmann equation, in The Boltzmann equation: theory and applications (E.G.D. Cohen and W. Thirring, eds.) (Springer-Verlag, Vienna) p. 379 (Acta Phys. Austriaca, Supp. 10).

Kadomtsev, B.B., 1957, Zh. Eksp. Teor. Fiz. **32**, 943 [Sov. Phys. JETP **5**, 771].

Van Kampen, N.G., 1961, Can. J. Phys. **39**, 551.

Van Kampen, N.G., 1972, Fluctuations, in Irreversibility in the many-body problem (J. Biel and J. Rae, eds.) (Plenum, New York).

Van Kampen, N.G., 1973, Physica **67**, 1.

Van Kampen, N.G., 1974, Phys. Lett. **50A**, 237.

Van Kampen, N.G., 1976a, Fluctuations in closed and open non-linear systems, in: Statistical physics: proceedings of the international conference, August, 1975, Budapest, Hungary (L. Pál and P. Szépfalusy, eds.) (North-Holland, Amsterdam) p. 29.

Van Kampen, N.G., 1976b, Phys. Repts. **24C**, 171.

Van Kampen, N.G., 1976c, Adv. Chem. Phys. **34**, 245.

Keizer, J., 1975, J. Chem. Phys. **63**, 398.

Keizer, J., 1976, J. Chem. Phys. **64**, 1679.

Kelly, G.E. and M.B. Lewis, 1971, Phys. Fluids **14**, 1925.

Kirkwood, J.G., 1946, J. Chem. Phys. **14**, 180.

Klein, M.J., ed., 1959, Paul Ehrenfest, collected scientific papers (North-Holland, Amsterdam).

Klein, M.J., 1970, Paul Ehrenfest (North-Holland, Amsterdam).

Klein, M.J., 1973, The development of Boltzmann's statistical ideas, in The Boltzmann equation: theory and applications (E.G.D. Cohen and W. Thirring, eds.) (Springer-Verlag, Vienna) pp. 53–106 (Acta Phys. Aust. Supp. 10).

Kolmogorov, A., 1933, Grundbegriffe der wahrscheinlichkeitsrechnung (Ergebnisse der Mathematik **2**(3), Berlin).

Kramers, H.A., 1940, Physica **7**, 284.

Kubo, R., 1965, Statistical mechanics (North-Holland, Amsterdam).

Kubo, R., K. Matzuo and K. Kitahara, 1973, J. Stat. Phys. **9**, 51.
Kurtz, T., 1972, J. Chem. Phys. **57**, 2976.
Landau, L.D. and E.M. Lifshitz, 1957, Zh. Eksp. Teor. Fiz. **32**, 618 [Sov. Phys. JETP **5**, 512].
Landau, L.D. and E.M. Lifshitz, 1960, Electrodynamics of continuous media (Pergamon, New York) ch. 13.
Langevin, P., 1908, Comptes rendus **146**, 530.
Logan, J.L., 1973, Bull. Am. Phys. Soc. **18**, 1591.
Logan, J.L. and M. Kac, 1976, Phys. Rev. A **13**, 458.
Loschmidt, J., 1876, Wiener Berichte **73**, 139.
Machlup, S. and L. Onsager, 1953, Phys. Rev. **91**, 1512.
Mathews, P.M., I.I. Shapiro and D.L. Falkoff, 1960, Phys. Rev. **120**, 1.
Mayer, J.E., 1947, J. Chem. Phys. **15**, 187.
Mayer, J.E., 1948, J. Chem. Phys. **16**, 665.
McKean, H.P., 1967, J. Combin. Theory **2**, 358.
McKean, H.P., 1975, Commun. Pure Appl. Math. **28**, 435.
Montgomery, D., 1969, Phys. Fluids **12**, 804.
Mori, H., 1973, Prog. Theor. Phys. **49**, 1516.
Moyal, J.E., 1949, J.R. Stat. Soc. B **11**, 150.
Nelson, E., 1967, Dynamical theories of Brownian motion (Princeton University Press, Princeton).
Nordsieck, A., W.E. Lamb and G.E. Uhlenbeck, 1940, Physica (Utr.) **7**, 344.
Nyquist, H., 1928, Phys. Rev. **32**, 110.
Onsager, L., 1931a, Phys. Rev. **37**, 405.
Onsager, L., 1931b, Phys. Rev. **38**, 2265.
Onsager, L. and S. Machlup, 1953, Phys. Rev. **91**, 1505.
Rytov, S.M., 1953, Theory of electrical fluctuation and heat radiation (Academy of Sciences Press, Moscow).
Siegert, A.J.F., 1949, Phys. Rev. **76**, 1708.
Skorobogatov, G.A., 1975a, Vestnik Leningrad Univ. **4**, 72.
Skorobogatov, G.A., 1975b, Vestnik Leningrad Univ. **10**, 72.
Smoluchowski, M., 1906, Ann. Physik **21**, 756.
Smoluchowski, M., 1916, Phys. Zeit. **17**, 557 and 585.
Stratonovich, R.L., 1963, Topics in the theory of random noise, vol. 1, translated by R.A. Silverman (Gordon and Breach, New York).
Takahasi, H., 1952, J. Phys. Soc. Japan, **7**, 439.
Tokuyama, M. and H. Mori, 1976, Prog. Theor. Phys. **56**, 1073.
Tolman, R.C., 1938, The principles of statistical mechanics (Oxford University Press, New York).
Uhlenbeck, G.E. and G.W. Ford, 1963, Lectures in statistical mechanics (American Mathematical Society, Providence, Rhode Island).
Uhlenbeck, G.E. and R. Fox, 1970, Phys. Fluids **13**, 2881.
Uhlenbeck, G.E. and S. Goudsmit, 1929, Phys. Rev. **34**, 145.
Uhlenbeck, G.E. and L.S. Ornstein, 1930, Phys. Rev. **36**, 823.
Uhlenbeck, G.E., P.C. Hemmer and M. Kac, 1963, J. Math. Phys. **4**, 229.
Van Vliet, K.M., 1964, Phys. Rev. **133**, A1182.
Wang, M.C. and G.E. Uhlenbeck, 1945, Rev. Mod. Phys. **17**, 323.
Wang Chang, C.S. and G.E. Uhlenbeck, 1970, in Studies in statistical mechanics, vol. 5 (J. deBoer and G.E. Uhlenbeck, eds.) (North-Holland, Amsterdam), part A.
Yamamoto, T., 1952, Busseiron Kenkyu **57**, 66 [in Japanese].
Yamamoto, T., 1953, Prog. Theor. Phys. **10**, 11.

CHAPTER 2

*On an Enriched Collection of Stochastic Processes**

ELLIOTT W. MONTROLL** AND BRUCE J. WEST

Physical Dynamics, Inc.
La Jolla, California 92038, *USA*

* This work was partially supported by the Fluid Dynamics Branch of the Office of Naval Research.

** Permanent Address, Institute for Fundamental Studies, Physics Dept., University of Rochester, Rochester, NY.

Fluctuation Phenomena
Eds. E.W. Montroll and J.L. Lebowitz