PHYSICAL REVIEW A VOLUME 13, NUMBER 1 JANUARY 1976

Fluctuations and the Boltzmann equation. I.

Jonothan Logan and Mark Kac
The Rockefeller University, New York, New York 10021

(Received 12 August 1975)

The evolution of a homogeneous dilute gas is treated as a Markov process in the complete set of K coarse-grained velocity states of all N particles. From the Siegert master equation for the process a Fokker-Planck equation is derived which describes, in the limit $N \rightarrow \infty$, the fluctuations in the occupation numbers $n_i(t)$, whose average behavior is governed by the (appropriately discretized) Boltzmann equation: The continuum limit $K \rightarrow \infty$ corresponds to fluctuations in the usual molecular distribution function $f(\vec{r}\,\vec{v};t)$. On similar reasoning, a Fokker-Planck equation is obtained for the fluctuation process near equilibrium, where the average is governed by the linearized Boltzmann equation. The theory of linear irreversible processes, which offers a statistical description of fluctuations on a thermodynamical basis, is applied to the linearized Boltzmann equation—treated as a linear phenomenological equation—following the development given recently by Fox and Uhlenbeck: The resulting stochastic equation is seen to be equivalent to the Fokker-Planck equation obtained from the master equation, yielding a multidimensional Ornstein-Uhlenbeck process which describes the fluctuations in molecular phase space.

I. INTRODUCTION

Recently Fox and Uhlenbeck[1] have shown how the line of reasoning developed by Langevin for the Brownian motion, and cast in general thermodynamic terms by Onsager and Machlup and others,[2] can be applied to the linearized Boltzmann equation to obtain a statistical description of the fluctuations in the molecular distribution function $f(\vec{r}\,\vec{v};t)$ of a dilute gas near equilibrium. They were able to show that the Chapman-Enskog development, applied to the stochastic Boltzmann equation so obtained, yields the fluctuating hydrodynamical equations of Landau and Lifshitz.[3] Since these can be used in turn to derive the Langevin equation of Brownian motion,[4] a nice display of consistency in the formalism of linear irreversible thermodynamics is thereby put in evidence. Yet the stochastic equation is obtained on a phenomenological basis no more secure than that of the macroscopic equations; thus the need remains to provide a microscopic basis for the linearized Boltzmann equation with fluctuations and to provide an analysis far from equilibrium, where even the phenomenological theory offers no guidance. What is wanted is a development along the lines of the BBGKY (Bogoliubov-Born-Green-Kirkwood-Yvon) hierarchy, say, further refined to adduce from the Liouville equation the properties of the fluctuation process, in addition to the average behavior of the molecular distribution, governed by the Boltzmann equation.[5]

Here we report some progress for the spatially homogeneous gas: In this case one can average over the particle position coordinates and treat the evolution of the remaining N velocity variables as an N-dimensional Markov process,[6] although a rigorous justification of this procedure from the Liouville equation is not yet available; or with somewhat less obvious justification (since contraction tends to destroy the Markov property), as a K-dimensional Markov process in K coarse-grained velocity states, as was first done by Siegert,[7] and which formalism we adopt here. In the limit $N \rightarrow \infty$ a Fokker-Planck equation is obtained from the Chapman-Kolmogorov equation for the process which describes the fluctuations in the occupation numbers $n_1(t), \ldots, n_K(t)$, whose averages obey the discretized Boltzmann equation of the formalism corresponding, in the limit $K \rightarrow \infty$, to fluctuations in the molecular distribution function $f(\vec{r}\,\vec{v};t)$. Close to equilibrium, where the averages obey the linearized Boltzmann equation, a Fokker-Planck equation is obtained which is equivalent to the stochastic equation of the phenomenological theory. These results are indifferent to a substitution for Boltzmann's *Stosszahlansatz* of a more general form, suggested by Uehling and Uhlenbeck to account for the quantum statistics of identical Bose-Einstein/Fermi-Dirac particles.

II. PHENOMENOLOGICAL FLUCTUATION THEORY

The phenomenological fluctuation theory of the Boltzmann equation proceeds from the observation that the molecular distribution function $f(\vec{r}\,\vec{v};t)$ represents an $(N-1)$-fold integral of the complete microscopic distribution so defined that

$$D_N(\vec{r}_1\vec{v}_1, \vec{r}_2\vec{v}_2, \ldots, \vec{r}_N\vec{v}_N; t)\, d\vec{r}_1\, d\vec{v}_1 \cdots d\vec{r}_N\, d\vec{v}_N \tag{2.1}$$

is the density of states in which the N particles are found near points $\vec{r}_1, \ldots, \vec{r}_N$ with velocities

$\vec{v}_1, \ldots, \vec{v}_N$, respectively, and so represents an average, just as macroscopic quantities like the hydrodynamical fluid velocity are obtained as averages over the molecular distribution itself:

$$\vec{u}(\vec{r};t) = \frac{1}{n}\int f(\vec{r}\,\vec{v};t)\vec{v}\,d\vec{v}\,, \tag{2.2}$$

where n represents the local density and $\vec{v}$ the molecular velocity.

The existence of approximate closed dynamical equations in both instances depends upon analogous circumstances which are important also to the fluctuation theory. In the hydrodynamical case the essential circumstance is the existence of a time scale $t_{\mathrm{mf}} \ll \tau \ll t_{\mathrm{equil}}$ on which local equilibrium is approximately attained through collisions while the hydrodynamical variables $n(\vec{r};t)$, $\vec{u}(\vec{r};t)$, and $T(\vec{r};t)$ (local density, hydrodynamical velocity, and temperature) decay relatively slowly to their equilibrium values through the transport of molecules, momentum, and energy through the gas, quantities which are conserved in a collision (1, $\vec{v}$, and v^2 are eigenfunctions of the linearized Boltzmann collision operator with vanishing eigenvalues). The scheme of Chapman and Enskog makes explicit the conception that in a certain definite approximation, this hydrodynamical stage is governed by the evolution of the microscopic variables only through their dependence on the hydrodynamical quantities which function, that is, as secular variables; that apart from this dependence, the gas may be conceived as in local equilibrium everywhere.[8] This idea is implemented through an expansion of the Boltzmann equation in which one seeks solutions of the form

$$f(\vec{r}\,\vec{v};t) \to f(\vec{r}\,\vec{v}\,|\,n\vec{u}T)\,, \tag{2.3}$$

depending on the underlying microscopic variables only parametrically, and expanding the solution in terms of a small quantity equivalent to $\epsilon = t_{\mathrm{mf}}/t_{\mathrm{macroscopic}}$.

The hydrodynamical equations obtained in each successive order are then supposed to represent that collective effect of the myriad molecular collisions which influences the evolution of the special average quantities like (2.2) connected with the conservation laws; the residual influence of the molecular motions amounting to a superposition of completely random noise upon the average behavior, contributing no further systematic effect.

A. Brownian motion

This is the point of view taken by Langevin, who reinterprets the phenomenological equation of motion of a small dense particle in a fluid,[9]

$$m\frac{du}{dt} + fu = 0 \tag{2.4}$$

(where m is the particle mass, u is its velocity, assumed small, and f is the friction constant, a parameter independent of u), as an average equation of motion:

$$\langle \mathcal{F}(t)\rangle_u = -m\beta u + O(u^2) \tag{2.5}$$

where $\mathcal{F}(t)$ represents the net instantaneous force due to collisions, and an average over an ensemble of particles with fixed velocity u is intended ($\beta = f/m$). Substituting $\tilde{F}(t) = \mathcal{F}(t)/m + \beta u$ gives (formally, exactly)

$$\frac{du}{dt} + \beta u = \tilde{F}(t)\,, \tag{2.6}$$

where only the linear term is kept, for small velocities. The term $\tilde{F}(t)$ is supposed, then, to represent that portion of the total molecular force which remains after the systematic part has been subtracted.

Langevin's idea is to represent the "random force" $\tilde{F}(t)$ as a stochastic process, in particular by the completely random stationary Gaussian process determined by

$$\langle \tilde{F}(t)\tilde{F}(t')\rangle = 2D\delta(t'-t)\,, \tag{2.7}$$

where D is a constant[10] independent of u. By definition $\langle \tilde{F}(t)\rangle = 0$; the higher moments are determined by (2.7) and the Gaussian assumption. The stochastic equation (2.6) with (2.7) determines a stationary Gaussian Markov process $u(t)$ whose properties are well known.[11]

As in the case of the hydrodynamical equations [from which (2.4) can be derived], the idealization depends first of all on the circumstance that the forward momentum of the particle is dissipated on a time scale $1/\beta$ very much longer than the time scale of molecular collisions, so that on the intermediate time scale $1/\beta \ll \tau \ll t_{\mathrm{coll}}$ local molecular equilibrium is at each instant approximately attained. Thus the random force is taken to be uncorrelated at different times. The Gaussian character derives from the local Maxwell-Boltzmann distribution and, again, from the fact that the motion is induced by the net influence of an enormous number of molecular collisions, each one of which produces only a trifling effect. It may be seen that the assumptions of the fluctuation theory are closely related to those conditions responsible for the existence of the phenomenological law in the first place.

B. Linear irreversible processes

The circumstances surrounding (2.4) that serve to justify the Langevin theory of Brownian motion do not differ in essentials from those connected with a whole class of familiar linear laws occurring in physics—Fourier's law of heat conduction, Ohm's law, the linear equations of near-equilibrium hydrodynamics, and so on—and the arguments leading to (2.6) and (2.7) can be cast in the general thermodynamic language introduced by Onsager a few years later,[12] as has been done in a number of equivalent treatments.[13]

One considers a system whose thermodynamic state is characterized by a complete set of n extensive variables $\alpha_1(t), \alpha_2(t), \ldots, \alpha_n(t) = \vec{\alpha}(t)$ defined for convenience to vanish at equilibrium, $\vec{\alpha}^{(0)} = 0$. Close to equilibrium the entropy

$$S = S_0 - \frac{k}{2} \sum_{ij}^{n} \alpha_i E_{ij} \alpha_j , \tag{2.8}$$

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 approximation (the first two terms in the Taylor expansion of the entropy) is equivalent to the Gaussian assumption, for inverting the Boltzmann-Planck formula for the entropy gives

$$P(\vec{\alpha}) = (\text{const}) \times \exp\left(-\frac{1}{2} \sum \alpha_i E_{ij} \alpha_j \right) \tag{2.9}$$

for the instantaneous probability of state $\vec{\alpha}$. Generalized thermodynamic forces are defined as

$$X_i = \frac{\partial S}{\partial \alpha_i} = -k \sum_j E_{ij} \alpha_j , \tag{2.10}$$

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

$$X_i = \sum_j R_{ij} J_j = \sum_j R_{ij} \dot{\alpha}_j = -k \sum E_{ij} \alpha_j , \tag{2.11}$$

where R_{ij} is a real nonsingular matrix, the real parts of whose eigenvalues are positive (dissipative rather than divergent systems). Defining

$$G_{ij} = k \sum L_{ik} E_{kj} , \tag{2.12}$$

where L is the inverse of R, one writes

$$\frac{d\alpha_i}{dt} + \sum_j G_{ij} \alpha_j = 0 \tag{2.13}$$

as the generalization of (2.4). Then the generalization of Langevin's equation (2.6) and (2.7) is

$$\frac{d\alpha_i}{dt} + \sum_j G_{ij} \alpha_j = \tilde{F}_i(t) , \tag{2.14}$$

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

$$\langle \tilde{F}_i(t) \tilde{F}_j(t') \rangle = 2 Q_{ij} \delta(t' - t) , \tag{2.15}$$

where Q_{ij} is symmetric, positive definite, and time independent. If the phenomenological variables are even functions of those molecular variables which have odd parity under time reversal and the Hamiltonian is time independent (stationary processes), it is easy to show that $L_{ij} = L_{ji}$ (the Onsager relations), and (2.15) takes the form[14]

$$\langle \tilde{F}_i(t) \tilde{F}_j(t') \rangle = 2 L_{ij} \delta(t' - t) . \tag{2.16}$$

III. FLUCTUATION THEORY FOR LINEARIZED BOLTZMANN EQUATION

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

$$f^{(0)}(v) = n_0 \left(\frac{m}{2\pi k T} \right)^{3/2} e^{-mv^2/2kT} \tag{3.1}$$

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

$$f(\vec{r}\vec{v}; t) = f^{(0)}[1 + h(\vec{r}\vec{v}; t)] \tag{3.2}$$

in the Boltzmann equation

$$\frac{\partial f}{\partial t} + \vec{v} \cdot \frac{\partial f}{\partial \vec{r}} = \int d\vec{v}_1 \int d\Omega \, g I(g, \theta) (f' f_1' - f f_1) \tag{3.3}$$

and neglect terms quadratic in h compared to the linear terms. Thus

$$\frac{\partial h}{\partial t} + \vec{v} \cdot \frac{\partial h}{\partial \vec{r}} = \int d\vec{v}_1 \int d\Omega \, g I(g, \theta) f_1^{(0)} [h' + h_1' - h - h_1] , \tag{3.4}$$

the linearized Boltzmann equation, is obtained. In this now standard notation, f_1 stands for $f(\vec{r}\vec{v}_1; t)$, f' for $f(\vec{r}\vec{v}'; t)$, and so forth, with reference to the four velocities $(\vec{v}, \vec{v}_1) \longrightarrow (\vec{v}', \vec{v}_1')$ characterizing a binary collision; $g = |\vec{v} - \vec{v}_1|$ is the relative velocity, which in the collision turns through the angle θ; the scattering cross section $I(g, \theta)$ is determined by the intermolecular force law (assumed additive and spherically symmetric), and $d\Omega$ is the element of solid angle.

The distribution function appearing in the Boltzmann equation can represent, as remarked earlier, only some sort of average of the exact phase-space density at any instant, just as is the case with the macroscopic variables obtained from the hydrodynamical equations. Indeed, the analogy has been made precise in Bogoliubov's theory of the Boltzmann equation, in which it is argued that the "kinetic stage" ($t \sim t_{\text{mf}}$) issues from the initial stage ($t \sim t_{\mathrm{coll}}$, typical duration of a collision—obviously defined only for short-range forces) in the same general fashion as the hydrodynamical stage issues from the kinetic stage, as described by Chapman and Enskog; that on the intermediate time scale $t_{\mathrm{coll}} \ll \tau \ll t_{\mathrm{mf}}$ the general s-particle distribution depends (approximately) on the molecular variables only parametrically, through the "secular variable" $f(\vec{r}\,\vec{v};t)$:

$$F_s(\vec{r}_1\vec{v}_1, \vec{r}_2\vec{v}_2, \ldots, \vec{r}_s\vec{v}_s; t) \rightarrow F_s(\vec{r}_1\vec{v}_1, \vec{r}_2\vec{v}_2, \ldots, \vec{r}_s\vec{v}_s \,|\, f(\vec{r}\,\vec{v};t)) , \tag{3.5}$$

which is to say that in the kinetic stage the Boltzmann distribution function performs the role occupied by the hydrodynamical variables in the macroscopic regime.[15] With the means provided by Bogoliubov's theory one can then argue in detail, as has recently been done,[16] for the appropriateness of the Boltzmann equation as a subject of the thermodynamical fluctuation theory.

In modern derivations of the Boltzmann equation one obtains from the Liouville equation a hierarchy of functional equations: $f_2(\vec{r}_1\vec{v}_1, \vec{r}_2\vec{v}_2; t)$, the joint distribution of two particles, as a functional of $f(\vec{r}\,\vec{v};t)$; $f_3(\vec{r}_1\vec{v}_1, \vec{r}_2\vec{v}_2, \vec{r}_3\vec{v}_3; t)$ as a functional of f_2, and so forth. A closed equation for f (the Boltzmann equation) is obtained by setting

$$f_2(\vec{r}_1\vec{v}_1, \vec{r}_2\vec{v}_2; t) = f(\vec{r}_1\vec{v}_1; t) f(\vec{r}_2\vec{v}_2; t) , \tag{3.6}$$

or the equivalent, at the last step. This situation is characteristic; since the motions of all the particles are interdependent, the exact specification of the distribution ultimately leads back to the complete (Liouville) distribution, but an approximation is introduced in order to obtain a closed dynamical equation; in consequence some form of average description is obtained. In intuitive derivations of the Boltzmann equation this step has its equivalent in the *Stosszahlansatz,* where one supposes that the frequency of binary collisions between particles with velocities $\vec{v}, \vec{v}_1$ is proportional to the product of their representations in the population, $f(\vec{r}\,\vec{v};t) f(\vec{r}\,\vec{v}_1;t)$, so assuming these are independent. To this statistical assumption can be traced the origin of the time-irreversible character of the Boltzmann equation and its nature as an average equation.[17] Indeed one may think of the fluctuations in $f(\vec{r}\,\vec{v};t)$ as arising in deviations from the mean collision frequency, on account of fluctuations from the average in the actual number of particles in the imaginary collision cylinder.

For the thermodynamical fluctuation theory, then, the linearized Boltzmann equation is to be taken as the linear phenomenological equation for the *average* density in molecular phase space, about which the actual value fluctuates in a random fashion one looks to describe as a generalized Brownian motion.

Fox and Uhlenbeck accomplish this formally by recasting (3.4) in the canonical variables of the thermodynamical fluctuation theory.

One defines the new function

$$a(\vec{r}\,\vec{v};t) = f_{(0)}^{1/2}(v) h(\vec{r}\,\vec{v};t) \tag{3.7}$$

and operators

$$A(\vec{r}\,\vec{v}, \vec{r}\,'\vec{v}\,') = f_{(0)}^{1/2}(v) \vec{v} \cdot \mathrm{grad}\, \delta(\vec{r} - \vec{r}\,')(\vec{v} - \vec{v}\,') , \tag{3.8}$$

$$S(\vec{r}\,\vec{v}, \vec{r}\,'\vec{v}\,') = f_{(0)}^{1/2}(v) f_{(0)}^{1/2}(v') K(\vec{v}, \vec{v}\,') \delta(\vec{r} - \vec{r}\,') , \tag{3.9}$$

where $K(\vec{v}, \vec{v}\,')$ is the symmetric kernel in terms of which the collision operator in (3.4) is written in the Hilbert-Enskog fashion

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

so that the linearized Boltzmann equation is now written

$$\frac{\partial a(\vec{r}\,\vec{v};t)}{\partial t} = -\int\int A(\vec{r}\,\vec{v}, \vec{r}\,'\vec{v}\,') a(\vec{r}\,'\vec{v}\,';t)\, d\vec{r}\,' d\vec{v}\,' - \int\int S(\vec{r}\,\vec{v}, \vec{r}\,'\vec{v}\,') a(\vec{r}\,'\vec{v}\,';t)\, d\vec{r}\,' d\vec{v}\,' . \tag{3.10}$$

If $\vec{r}$ and $\vec{v}$ may be conceived as continuous indices, this appears in the standard form of the average regression equations for a linear process, represented by (2.13). A and S possess odd and even time symmetry, respectively, through their respective dependence on odd and even functions of the molecular velocity $\vec{v}$. Then one needs the entropy matrix appearing in (2.8); Boltzmann's expression

$$S(t) = -k \int\int f(\vec{r}\,\vec{v};t) \ln f(\vec{r}\,\vec{v};t)\, d\vec{r}\, d\vec{v} \tag{3.11}$$

gives for the entropy of a gas close to equilibrium

$$S(t) = S_0 - \frac{k}{2} \int\int a^2(\vec{r}\,\vec{v};t)\, d\vec{r}\, d\vec{v} , \tag{3.12}$$

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

$$E(\vec{r}\,\vec{v}, \vec{r}\,'\vec{v}\,') = \delta(\vec{r} - \vec{r}\,')\delta(\vec{v} - \vec{v}\,') . \tag{3.13}$$

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

$$\frac{\partial h}{\partial t} = -\vec{v}\cdot\frac{\partial h}{\partial \vec{r}} + Lh + \tilde{C}(\vec{r}\,\vec{v};t)\,, \tag{3.14}$$

where L stands for the linearized collision operator in (3.4) and the analog of the random force is the stationary Gaussian process determined by

$$\langle \tilde{C}(\vec{r}\,\vec{v};t)\tilde{C}(\vec{r}\,'\vec{v}\,';t')\rangle = 2K(\vec{v},\vec{v}\,')\delta(\vec{r}-\vec{r}\,')\delta(t-t')\,. \tag{3.15}$$

IV. COARSE-GRAIN MASTER EQUATION

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 $(\vec{v},\vec{v}+\Delta\vec{v})$ 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 $(n_1(t),n_2(t),\ldots,n_K(t)) = \vec{n}(t)$ where

$$\sum_{i=1}^{K} n_i(t) = N\,. \tag{4.1}$$

The idea, apparently first used by Nordsieck, Lamb, and Uhlenbeck,[18] 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 $\vec{n}(t)$.

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, $\vec{n}(t)$ is a K-dimensional vector Markov process with stationary transition probabilities, so that the conditional probability of finding the system in state $\vec{n}$, given initial state $\vec{m}$,

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

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

$$P(\vec{n};\tau) = \sum_{\vec{m}} P(\vec{m};0)P(\vec{m}\,|\,\vec{n};\tau)\,, \tag{4.2}$$

where the sum is taken over all sets of occupation numbers $(m_1,\ldots,m_K)$ consistent with (4.1). By the nature of the process n can change by only a small amount in an infinitesimal time interval, and so one can write

$$P(\vec{m}\,|\,\vec{n};\Delta t) = \delta(\vec{m}\,|\,\vec{n}) + Q(\vec{m}\,|\,\vec{n})\Delta t + O((\Delta t)^2)\,, \tag{4.3}$$

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

$$Q(\vec{m}\,|\,\vec{n}) \geq 0 \quad \text{for } \vec{m}\neq\vec{n} \tag{4.4}$$

(positive probabilities), and

$$Q(\vec{m}\,|\,\vec{m}) = -\sum_{\vec{n}\neq\vec{m}}{}' Q(\vec{m}\,|\,\vec{n}) \tag{4.5}$$

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

$$\sum_{\vec{n}} P(\vec{m}\,|\,\vec{n};\Delta t) = 1\,. \tag{4.6}$$

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

Equations (4.2) and (4.3) give in the limit $\Delta t\to 0$ the master equation

$$\frac{dP(\vec{n};t)}{dt} = \sum_{\vec{m}} P(\vec{m};t)Q(\vec{m}\,|\,\vec{n})\,, \tag{4.7}$$

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.[19] In the coarse-grained description employed here one goes further, lumping states together in discrete cells; sufficient "contraction" of this kind eventually 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.[20])

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

$$P(ij\to kl)\,dt = \alpha_{ij}^{kl} m_i m_j\,dt\,, \tag{4.8}$$

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

$$\alpha_{ij}^{kl} = \alpha_{ji}^{kl} = \alpha_{ij}^{lk} \tag{4.9}$$

and

$$\alpha_{ii}^{kl} = \alpha_{ij}^{kk} = 0 . \tag{4.10}$$

In addition, for central forces one has "microreversibility"

$$\alpha_{ij}^{kl} = \alpha_{kl}^{ij} , \tag{4.11}$$

equality of the cross sections for forward and inverse or "restituting" collisions.[21] If ϵ_i is the energy associated with state i, then one must have

$$\epsilon_i + \epsilon_j = \epsilon_k + \epsilon_l \quad \text{if } \alpha_{ij}^{kl} \neq 0 , \tag{4.12}$$

i.e., for any allowed transition: conservation of energy in a collision. The appropriate transition matrix is

$$Q(\vec{\mathrm{m}}|\vec{\mathrm{n}}) = \frac{1}{4}\Bigg(\sum_{ij\neq kl}{}' \alpha_{ij}^{kl} m_i m_j \delta_{n_i m_i - 1} \delta_{n_j m_j - 1} \delta_{n_k m_k + 1} \delta_{n_l m_l + 1} \times \prod_{r \neq ijk}{}' \delta_{n_r m_r} + \sum_{ij} \alpha_{ij}^{ij} m_i m_j \prod_r \delta_{n_r m_r} \Bigg), \tag{4.13}$$

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

$$\alpha_{ij}^{ij} = - \sum_{ij \neq kl}{}' \alpha_{ij}^{kl} \tag{4.14}$$

so that the unrestricted sum vanishes,

$$\sum_{kl} \alpha_{ij}^{kl} = 0 , \tag{4.15}$$

and gives

$$\sum_{\vec{\mathrm{n}}} P(\vec{\mathrm{m}}|\vec{\mathrm{n}}; \Delta t) = 1 , \tag{4.16}$$

the probability that the system, initially in state $\vec{\mathrm{m}}$, is found in *some* state after time Δt. The equilibrium master distribution, which can be obtained as the limit of the conditional probability $P(\vec{\mathrm{m}}|\vec{\mathrm{n}}; \tau)$ for $\tau \rightarrow \infty$, is

$$W(\vec{\mathrm{n}}) = (\mathrm{const}) \times \frac{N!}{n_1 ! n_2 ! \cdots n_K !} , \tag{4.17}$$

subject to

$$\sum_{i=1}^{K} n_i = N \tag{4.18}$$

and

$$\sum_{i=1}^{K} n_i \epsilon_i = E = N\epsilon , \tag{4.19}$$

which is a consequence of (4.12).

A. Boltzmann equation

From the master equation (4.7) with the above transition matrix one can, by summing over all but 1, 2, ... of the occupation numbers, obtain equations for $\langle n_i(t)\rangle, \langle n_i(t) n_j(t)\rangle, \ldots,$ etc.,[22] where the symbols indicate an average over all possible complete configurations of the system ("master states"). In particular, one gets

$$\begin{aligned} \frac{d\langle n_s\rangle}{dt} &= \frac{1}{2} \sum_{ijk} \alpha_{ij}^{ks} \langle n_i n_j \rangle \\ &= \frac{1}{2} \sum_{\substack{ijk \\ ij \neq ks}}{}' \alpha_{ij}^{ks} \langle n_i n_j \rangle + \frac{1}{2} \sum_k \alpha_{ks}^{ks} \langle n_k n_s \rangle , \end{aligned} \tag{4.20}$$

which can be written, in view of (4.30),

$$\frac{d\langle n_s\rangle}{dt} = \frac{1}{2} \sum_{ij \neq ks}{}' \alpha_{ij}^{ks} \langle n_i n_j \rangle - \frac{1}{2} \sum \alpha_{ks}^{ij} \langle n_k n_s \rangle ; \tag{4.21}$$

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

$$\frac{d\langle n_s\rangle}{dt} = \frac{1}{2} \sum_{ij \neq ks}{}' \alpha_{ij}^{ks} (\langle n_i \rangle \langle n_j \rangle - \langle n_k \rangle \langle n_s \rangle) , \tag{4.22}$$

in which the microreversibility symmetry (4.11) 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* [23]). If, following Boltzmann, one defines

$$H = \sum_s \langle n_s \rangle \ln \langle n_s \rangle , \tag{4.23}$$

it is easy to show that

$$\frac{dH}{dt} = \frac{1}{2} \sum \alpha_{ij}^{ks} (\langle n_i \rangle \langle n_j \rangle - \langle n_k \rangle \langle n_s \rangle) \times \ln\left(\frac{\langle n_k \rangle \langle n_s \rangle}{\langle n_i \rangle \langle n_j \rangle} \right) \leq 0 , \tag{4.24}$$

the H theorem. Note that (4.12) is not required; only the non-negativity of the α_{ij}^{kl} is needed. It follows that if there is an equilibrium state—one characterized by $dH/dt = 0$—then

$$\alpha_{ij}^{kl} = 0 \quad \text{unless } \bar{n}_i \bar{n}_j = \bar{n}_k \bar{n}_l , \tag{4.25}$$

where the bars (here and in the following) designate equilibrium averages. In equilibrium at temperature T, the occupancies follow the canonical distribution

$$\bar{n}_i = Ae^{-\epsilon_i/kT}, \tag{4.26}$$

where A is of course a normalization constant.

For small departures from equilibrium, one can write

$$\langle n_i \rangle = \bar{n}_i(1+h_i), \tag{4.27}$$

substitute in (4.22), neglect quadratic terms compared to h, and obtain, making use of (4.25), the linearized Boltzmann equation

$$\frac{d\langle n_s\rangle}{dt} = \frac{1}{2}\sum_{ijk} \alpha_{ij}^{ks}\bar{n}_i\bar{n}_j[h_i+h_j-h_k-h_s], \tag{4.28}$$

which can be written

$$\frac{d\langle y_s\rangle}{dt} = \frac{1}{\bar{n}_s}\sum_{\sigma} g_{s\sigma}\langle y_s\rangle = 0, \tag{4.29}$$

where $y_s = n_s/\bar{n}_s$ and

$$g_{s\sigma} = -\sum_{jk}\alpha_{\sigma j}^{ks}\bar{n}_\sigma\bar{n}_j + \frac{1}{2}\sum_{ij}\alpha_{ij}^{\sigma s}\bar{n}_i\bar{n}_j + \tfrac{1}{2}\delta_{s\sigma}\sum_{ijk}\alpha_{ij}^{ks}\bar{n}_i\bar{n}_j, \tag{4.30}$$

(note $\sum_\sigma g_{s\sigma} = 0$). It is convenient to define $\nu_s = n_s/N$ ($\bar{\nu}_s = \bar{n}_s/N$) and $a_{ij}^{kl} = N\alpha_{ij}^{kl}$, which are, like y_s, $O(1)$.

The linearized Boltzmann equation with fluctuations from the thermodynamical theory [Eqs. (3.14) and (3.15)] can be written, in this formalism,

$$\frac{d\langle y_s\rangle}{dt} + \frac{1}{\bar{\nu}_s}\sum_{\sigma} g_{s\sigma}\langle y_s\rangle = \tilde{c}_s(t) \tag{4.31}$$

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

$$\langle\tilde{c}_s(t)\tilde{c}_\sigma(t')\rangle = (2/\bar{\nu}_s\bar{\nu}_\sigma)g_{s\sigma}\delta(t'-t), \tag{4.32}$$

all, again, in contemplation of the limit $N\to\infty$. In Sec. IV B we use the means at hand to calculate the properties of the fluctuations close to equilibrium; that is, to determine the needed features of the process

$$Y_s(t) = (1/\bar{\nu}_s\sqrt{N})[n_s(t) - \bar{n}_s] \tag{4.33}$$

from the master equation.

B. Fokker-Planck equation: Fluctuations around equilibrium

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

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

is also; in the limit $N\to\infty$ these ought still to be vector Markov processes, in continuous variables.[25] 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 (4.31) and (4.32).

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

$$\frac{dP(\vec{x};t)}{dt} = -\sum_s \frac{\partial}{\partial x_s}[A_s(\vec{x})P(\vec{x})] + \frac{1}{2}\sum_{s\sigma}\frac{\partial^2}{\partial x_s\partial x_\sigma}[B_{s\sigma}(\vec{x})P(\vec{x})]. \tag{4.34}$$

What is required are the values of the coefficients in (4.34) and a demonstration that the higher coefficients $C_{s\sigma\rho}$, etc., vanish in the indicated limit.

From the familiar theory of the Fokker-Planck equation,[26] recall that the coefficients are defined as limits of conditional averages: First of all,

$$A_s(\vec{x}) = \lim_{\Delta t\to 0}\frac{1}{\Delta t}\langle Y_s - X_s\rangle = \lim_{\Delta t\to 0}\frac{1}{\Delta t}\frac{\langle n_s - m_s\rangle}{\bar{\nu}_s\sqrt{N}}, \tag{4.35}$$

where what is meant is the conditional average

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

which is equal to

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

since the process is stationary, and

$$X_s = (1/\bar{\nu}_s\sqrt{N})[m_s(t) - \bar{m}_s]. \tag{4.36}$$

This can be obtained from (4.3):

$$\begin{aligned} A_s(\vec{x}) &= \lim_{\Delta t\to 0}\frac{1}{\Delta t}\frac{1}{\bar{\nu}_s\sqrt{N}}\sum_{\vec{n}}\{[\delta(\vec{m}\,|\,\vec{n}) + \Delta t Q(\vec{m}\,|\,\vec{n})](n_s - m_s)\} \\ &= \frac{1}{\bar{\nu}_s\sqrt{N}}\sum_{\vec{n}}[Q(\vec{m}\,|\,\vec{n})(n_s - m_s)] \\ &= \frac{1}{\bar{\nu}_s\sqrt{N}}\sum_{\vec{n}}Q(\vec{m}\,|\,\vec{n})n_s, \end{aligned} \tag{4.37}$$

since according to (4.5)

$$\sum_{\vec{n}} Q(\vec{m}\,|\,\vec{n}) = 0. \tag{4.38}$$

Substituting the transition matrix (4.13) gives

$$\sum_{\vec{n}} Q(\vec{m}\,|\,\vec{n})n_s = \frac{1}{4N}\Bigg(\sum_{ijkl\neq s}{}' a_{ij}^{kl} m_i m_j m_s + \sum_{jkl} a_{sj}^{kl}\, m_s m_j(m_s - 1)$$

$$+ \sum_{ikl} a_{is}^{kl} m_i m_s(m_s - 1) + \sum_{ijl} a_{ij}^{sl} m_i m_j(m_s + 1) + \sum_{ijk} a_{ij}^{ks} m_i m_j(m_s + 1)\Bigg)\,. \qquad (4.39)$$

The cubic terms can be grouped together and written

$$\sum_{ijkl} a_{ij}^{kl} m_i m_j m_s\,, \qquad (4.40)$$

where the range of summation is now unrestricted; on account of (4.15) this vanishes. Using the index symmetry, (4.39) can then be written

$$\sum_{\vec{n}} Q(\vec{m}\,|\,\vec{n})n_s = \frac{1}{N}\left(-\frac{1}{2}\sum a_{sj}^{kl} m_s m_j + \frac{1}{2}\sum a_{ij}^{sl} m_i m_j\right). \qquad (4.41)$$

Supposing the system is near equilibrium, we write

$$m_i = \overline{m}_i(1 + x_i/\sqrt{N})\,; \qquad (4.42)$$

to first order in the small quantity we get then, making use of (4.25),

$$\sum_{\vec{n}} Q(\vec{m}\,|\,\vec{n})n_s = \frac{1}{2}\left(\frac{1}{2\sqrt{N}}\sum_{ijk} a_{ij}^{ks}\,\overline{m}_i\overline{m}_j(x_i + x_j - x_s - x_k)\right), \qquad (4.43)$$

and we have finally for (4.37)

$$A_s(\vec{x}) = \frac{1}{2\overline{\nu}_s}\sum a_{ij}^{ks}\nu_i\nu_j(x_i + x_j - x_s - x_k)$$

$$= \frac{-1}{\overline{\nu}_s}\sum_{\sigma} g_{s\sigma}x_\sigma\,, \qquad (4.44)$$

where

$$g_{s\sigma} = -\sum_{jk} a_{\sigma j}^{ks}\overline{\nu}_\sigma\overline{\nu}_j + \frac{1}{2}\sum_{ij} a_{ij}^{\sigma s}\overline{\nu}_i\overline{\nu}_j + \tfrac{1}{2}\delta_{s\sigma}\sum_{ijk} a_{ij}^{ks}\overline{\nu}_i\overline{\nu}_j \qquad (4.45)$$

[the zeroth-order terms cancel on application of (4.25)].

The second coefficient

$$B_{s\sigma}(\vec{x}) = \lim_{\Delta t\to 0}\,(1/\Delta t)\langle (Y_s - X_s)(Y_\sigma - X_\sigma)\rangle$$

$$= \lim_{\Delta t\to 0}\,(1/N\overline{\nu}_s\overline{\nu}_\sigma)\langle (n_s - m_s)(n_\sigma - m_\sigma)\rangle\,, \qquad (4.46)$$

where again *conditional* averages are intended; proceeding as before,

$$B_{s\sigma}(\vec{x}) = \lim_{\Delta t\to 0}\frac{1}{\Delta t}\,\frac{1}{N\overline{\nu}_s\overline{\nu}_\sigma}\sum_{\vec{n}}\{[\delta(\vec{m}\,|\,\vec{n}) + \Delta t\,Q(\vec{m}\,|\,\vec{n})]$$

$$\times(n_s - m_s)(n_\sigma - m_\sigma)\}$$

$$= \frac{1}{N\overline{\nu}_s\overline{\nu}_\sigma}\sum_{\vec{n}} Q(\vec{m}\,|\,\vec{n})[(n_s - m_s)(n_\sigma - m_\sigma)]\,. \qquad (4.47)$$

When this is written out, the quartic terms, grouped together, can be written as the unrestricted sum

$$\sum a_{ij}^{kl} m_i m_j m_s m_\sigma\,,$$

which vanishes; the cubic terms can be written, again substituting (4.42),

$$\frac{1}{4N}\,\frac{1}{\sqrt{N}}\Bigg(-2\sum a_{sj}^{kl}\overline{m}_s\overline{m}_\sigma\overline{m}_j(x_j + x_s) + 2\sum a_{ij}^{sl}\overline{m}_i\overline{m}_j\overline{m}_\sigma(x_i + x_j)$$

$$-2\sum a_{\sigma j}^{kl}\overline{m}_\sigma\overline{m}_j\overline{m}_s(x_\sigma + x_j) + 2\sum a_{ij}^{\sigma l}\overline{m}_i\overline{m}_j\overline{m}_s(x_i + x_j)\Bigg)\,, \qquad (4.48)$$

where use has been made of (4.25) (again the zeroth-order part cancels); and the quadratic terms can be written

$$\frac{1}{4N}\left(-8\sum a_{\sigma j}^{ks}\overline{m}_\sigma\overline{m}_j + 4\sum a_{ij}^{\sigma s}\overline{m}_i\overline{m}_j\right),\quad \sigma\neq s\,. \qquad (4.49)$$

Since

$$\sum_{\vec{n}} Q(\vec{m}\,|\,\vec{n})n_s m_\sigma = m_\sigma\sum_{\vec{n}} Q(\vec{m}\,|\,\vec{n})n_s\,,$$

it follows immediately from the calculation of (4.44) that

$$\sum_{\vec{\mathrm{n}}} Q(\vec{\mathrm{m}}\,|\,\vec{\mathrm{n}})n_s m_\sigma = \frac{1}{N}\left[-\frac{1}{2}\sum a_{sj}^{kl}\bar{m}_s\bar{m}_j\bar{m}_\sigma\left(\frac{x_s+x_j}{\sqrt{N}}\right) + \frac{1}{2}\sum a_{ij}^{sl}\bar{m}_i\bar{m}_j\bar{m}_\sigma\left(\frac{x_i+x_j}{\sqrt{N}}\right)\right] \tag{4.50}$$

and

$$\sum_{\vec{\mathrm{n}}} Q(\vec{\mathrm{m}}\,|\,\vec{\mathrm{n}})n_\sigma m_s = \frac{1}{N}\left[-\frac{1}{2}\sum a_{\sigma j}^{kl}\bar{m}_\sigma\bar{m}_j\bar{m}_s\left(\frac{x_\sigma+x_j}{\sqrt{N}}\right) + \frac{1}{2}\sum a_{ij}^{sl}\bar{m}_i\bar{m}_j\bar{m}_s\left(\frac{x_i+x_j}{\sqrt{N}}\right)\right]; \tag{4.51}$$

finally,

$$\sum_{\vec{\mathrm{n}}} Q(\vec{\mathrm{m}}\,|\,\vec{\mathrm{n}})m_s m_\sigma = m_s m_\sigma \sum_{\vec{\mathrm{n}}} Q(\vec{\mathrm{m}}\,|\,\vec{\mathrm{n}}) = 0\,. \tag{4.52}$$

The first-order terms (4.48) and (4.51) cancel, leaving only the zeroth-order part (4.49); completing it with the (trivial) $s=\sigma$ case, we have finally

$$B_{s\sigma}(\vec{\mathrm{x}}) = \frac{1}{\bar{\nu}_s\bar{\nu}_\sigma}\left(-2\sum a_{\sigma j}^{ks}\bar{\nu}_\sigma\bar{\nu}_j + \sum a_{ij}^{\sigma s}\bar{\nu}_i\bar{\nu}_j + \delta_{s\sigma}\sum a_{ij}^{ks}\bar{\nu}_i\bar{\nu}_j\right) = (2/\bar{\nu}_s\bar{\nu}_\sigma)g_{s\sigma}\,. \tag{4.53}$$

Reasoning in this way, one can go on to show that

$$C_{s\sigma\rho} = 0\,, \tag{4.54}$$

and likewise all the higher coefficients vanish in the required limit; the calculations are outlined in Appendix A.

Thus we obtain the Fokker-Planck equation with

$$A_s(\vec{\mathrm{x}}) = \frac{-1}{\bar{\nu}_s}\sum_\sigma g_{s\sigma}x_\sigma \tag{4.55}$$

and

$$B_{s\sigma}(\vec{\mathrm{x}}) = (2/\bar{\nu}_s\bar{\nu}_\sigma)g_{s\sigma}\,. \tag{4.56}$$

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

$$\frac{dy_s}{dt} + \frac{1}{\bar{\nu}_s}\sum_\sigma g_{s\sigma}y_\sigma = \tilde{c}_s(t)\,, \tag{4.57}$$

where again

$$g_{s\sigma} = -\sum_{jk} a_{\sigma j}^{ks}\bar{\nu}_\sigma\bar{\nu}_j + \frac{1}{2}\sum_{ij} a_{ij}^{\sigma s}\bar{\nu}_i\bar{\nu}_j + \delta_{s\sigma}\sum_{ijk} a_{ij}^{ks}\bar{\nu}_i\bar{\nu}_j\,, \tag{4.58}$$

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

$$\langle\tilde{c}_s(t)\rangle = 0 \tag{4.59}$$

and

$$\langle\tilde{c}_s(t)\tilde{c}_s(t')\rangle = (2/\bar{\nu}_s\bar{\nu}_\sigma)g_{s\sigma}\delta(t'-t)\,, \tag{4.60}$$

the linearized Boltzmann equation with fluctuations (4.31) and (4.32). [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 (4.31) and (4.32).]

C. Fokker-Planck equation: Fluctuations away from equilibrium[27]

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 (4.22), which can as well be written

$$\frac{d\langle\nu_s\rangle}{dt} = \frac{1}{2}\sum a_{ij}^{sl}\langle\nu_i\rangle\langle\nu_j\rangle - \frac{1}{2}\sum a_{sj}^{kl}\langle\nu_s\rangle\langle\nu_j\rangle\,, \tag{4.61}$$

making use of (4.11) (recall $\nu_i = n_i/N$). It is convenient for the purpose to redefine the random variables

$$x_s(t) = [m_s(t) - \langle m_s(t)\rangle]/\sqrt{N}\,,$$
$$y_s(t) = [n_s(t) - \langle n_s(t)\rangle]/\sqrt{N}\,, \tag{4.62}$$

evidently $O(1)$.

First of all we require

$$A_s(\vec{\mathrm{x}}) = \lim_{\Delta t\to 0}(1/\Delta t)\langle y_s(t+\Delta t) - y_s(t)\rangle\,, \tag{4.63}$$

where again the conditional average

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

is intended. This we can write

$$\lim_{dt\to 0}\frac{1}{dt}\left(\frac{1}{\sqrt{N}}\langle n_s(t+dt) - n_s(t)\,|\,\vec{\mathrm{n}}(t) = \vec{\mathrm{m}}\rangle - \frac{1}{\sqrt{N}}\frac{d\langle n_s\rangle}{dt}\,dt\right)$$
$$= \frac{1}{\sqrt{N}}\left(-\frac{1}{2}\frac{1}{N}\sum a_{sj}^{kl}(\langle m_j\rangle + \sqrt{N}x_j)(\langle m_s\rangle + \sqrt{N}x_s) + \frac{1}{2}\frac{1}{N}\sum a_{ij}^{sl}(\langle m_i\rangle + \sqrt{N}x_i)(\langle m_j\rangle + \sqrt{N}x_j) - \frac{d\langle n_s\rangle}{dt}\right)\,. \tag{4.64}$$

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

$$\sqrt{N}\left(-\frac{1}{2}\sum a_{sj}^{kl}\langle\nu_s\rangle\langle\nu_j\rangle+\frac{1}{2}\sum a_{ij}^{sl}\langle\nu_i\rangle\langle\nu_j\rangle-\frac{d\langle n_s\rangle}{dt}\right)-\frac{1}{2}\sum a_{sj}^{kl}(\langle\nu_j\rangle x_s+\langle\nu_s\rangle x_j)+\frac{1}{2}\sum a_{ij}^{sl}(\langle\nu_i\rangle x_j+\langle\nu_j\rangle x_i)\,; \tag{4.65}$$

the quantity in large parentheses, of order $\sqrt{N}$, vanishes on substitution of the Boltzmann equation (4.61) for the time derivative, and what remains $[O(1)]$ can be rewritten, so yielding

$$A_s(\vec{x})=-\sum_{\sigma}\gamma_{s\sigma}x_{\sigma}\,, \tag{4.66}$$

where

$$\gamma_{s\sigma}=-\sum a_{\sigma j}^{sl}\langle\nu_j\rangle+\frac{1}{2}\sum a_{ij}^{s\sigma}\langle\nu_s\rangle+\tfrac{1}{2}\delta_{s\sigma}\sum a_{sj}^{kl}\langle\nu_j\rangle\,. \tag{4.67}$$

Then we need

$$B_{s\sigma}(\vec{x})=\lim_{\Delta t\to 0}(1/\Delta t)\langle[y_s(t+\Delta t)-y_s(t)]\times[y_{\sigma}(t+\Delta t)-y_{\sigma}(t)]\,|\,\vec{y}(t)=\vec{x}\rangle\,; \tag{4.68}$$

proceeding in the same way, we obtain

$$B_{s\sigma}=\frac{1}{2}\sum a_{ij}^{s\sigma}\langle\nu_i\rangle\langle\nu_j\rangle+\frac{1}{2}\sum a_{s\sigma}^{kl}\langle\nu_s\rangle\langle\nu_{\sigma}\rangle-\sum a_{\sigma j}^{sl}\langle\nu_{\sigma}\rangle\langle\nu_j\rangle-\sum a_{sj}^{\sigma l}\langle\nu_s\rangle\langle\nu_j\rangle\,. \tag{4.69}$$

Since the calculation outlined in Appendix A applies equally to the present case, we have again $C_{s\sigma\rho}=0$, etc., and thus the Fokker-Planck equation (4.34), but now with time-dependent coefficients, since the $\langle\nu_i\rangle$—solutions of the Boltzmann equation—do of course vary with time.

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 Sec. IV B, 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 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.

ACKNOWLEDGMENTS

We are grateful to G. E. Uhlenbeck for useful discussions and helpful suggestions. This work was supported in part by National Science Foundation Grant No. NPS74-13252 and in part (M.K.) by AFOSR 72-218.

APPENDIX A: HIGHER TERMS FOR THE FOKKER- PLANCK EQUATION

Proceeding as before, one obtains

$$C_{s\sigma\rho}(\vec{x})=\lim_{\Delta t\to 0}\langle(Y_s-X_s)(Y_{\sigma}-X_{\sigma})(Y_{\rho}-X_{\rho})\rangle$$

$$=\frac{1}{N^{3/2}\overline{\nu}_s\overline{\nu}_{\sigma}\overline{\nu}_{\rho}}\sum_{\vec{n}}Q(\vec{m}\,|\,\vec{n})(n_sn_{\sigma}n_{\rho}-n_sn_{\sigma}m_{\rho}-n_sn_{\rho}m_{\sigma}-n_{\sigma}n_{\rho}m_s+n_sm_{\sigma}m_{\rho}+n_{\sigma}m_sm_{\rho}+n_{\rho}m_sm_{\sigma}-m_sm_{\sigma}m_{\rho})\,. \tag{A1}$$

The only term of new type is

$$\sum_{\vec{n}}Q(\vec{m}\,|\,\vec{n})n_sn_{\sigma}n_{\rho}\,.$$

When this is written out, one observes that the contribution of the quintic terms vanishes, again on account of (4.15); there remain quartic and cubic terms. Since the latter are smaller by a factor N, only the former need be considered for the limit, so we have

$$\sum_{\vec{n}} Q(\vec{m}\,|\,\vec{n}) n_s n_\sigma n_\rho = \frac{1}{4}\frac{1}{N}\Big(-2\sum a_{sj}^{kl} m_s m_j m_\sigma m_\rho + 2\sum a_{ij}^{sl} m_i m_j m_\sigma m_\rho - 2\sum a_{\sigma j}^{kl} m_\sigma m_j m_s m_\rho$$
$$+2\sum a_{ij}^{\sigma l} m_i m_j m_s m_\rho - 2\sum a_{\rho j}^{kl} m_\rho m_j m_s m_\sigma + 2\sum a_{ij}^{\rho l} m_i m_j m_s m_\sigma\Big). \tag{A2}$$

Since everything is symmetric with respect to $s\sigma\rho$, one may write just the contributions to (A1) involving a_{sj}^{kl}, etc.

From (A2) one obtains

$$\langle\langle n_s n_\sigma n_\rho\rangle\rangle_s = -2\sum a_{sj}^{kl} m_s m_j m_\sigma m_\rho + 2\sum a_{ij}^{sl} m_i m_j m_\sigma m_\rho ; \tag{A3}$$

From (4.47) *et seq*. one obtains

$$\langle\langle n_s n_\sigma m_\rho\rangle\rangle_s = -2\sum a_{sj}^{kl} m_s m_j m_\sigma m_\rho + 2\sum a_{ij}^{sl} m_i m_j m_\sigma m_\rho , \tag{A4}$$

$$\langle\langle n_s n_\rho m_\sigma\rangle\rangle_s = -2\sum a_{sj}^{kl} m_s m_j m_\sigma m_\rho + 2\sum a_{ij}^{sl} m_i m_j m_\sigma m_\rho , \tag{A5}$$

and

$$\langle\langle n_\sigma n_\rho m_s\rangle\rangle_s = 0 . \tag{A6}$$

From (4.41) one obtains

$$\langle\langle n_s m_\sigma m_\rho\rangle\rangle_s = -2\sum a_{sj}^{kl} m_s m_j m_\sigma m_\rho + 2\sum a_{ij}^{sl} m_i m_j m_\sigma m_\rho , \tag{A7}$$

$$\langle\langle n_\sigma m_s m_\rho\rangle\rangle_s = 0 , \tag{A8}$$

$$\langle\langle n_\rho m_s m_\sigma\rangle\rangle_s = 0 ; \tag{A9}$$

and from (4.5) it follows that

$$\langle m_s m_\sigma m_\rho\rangle = 0 . \tag{A10}$$

Combining (A3)–(A5) and (A7) gives zero. Since the other cases can be generated by interchanging $s\sigma\rho$, we conclude that

$$\sum_{\vec{n}} Q(\vec{m}|\vec{n})[(n_s - m_s)(n_\sigma - m_\sigma)(n_\rho - m_\rho)] = 0 ,$$

and [to $O(1)$],

$$C_{s\sigma\rho}(\vec{X}) = \lim_{\Delta t\to 0} (1/\Delta t)\langle (Y_s - X_s)(Y_\sigma - X_\sigma)(Y_\rho - X_\rho)\rangle = 0 . \tag{A11}$$

Proceeding in the way just illustrated, one can see that the reasoning above applies likewise to the higher coefficients, that to $O(1)$,

$$D_{s\sigma\rho\gamma}(\vec{X}) = 0 , \tag{A12}$$

etc. If we call $A, B, C\ldots$ the first, second, third coefficients, etc., at the Rth coefficient there will be a contribution [corresponding to (A3)]

$$\langle\langle n_1 n_2 \ldots n_s \ldots n_R\rangle\rangle_s = C_s^{(R)} ;$$

there will be $R-1$ terms of the type

$$\langle\langle m_1 m_2 \ldots n_i \ldots n_s \ldots m_R\rangle\rangle_s = -C_s^{(R)} ,$$

$\frac{1}{2}(R-1)(R-2)$ terms of the type,

$$\langle\langle m_1 m_2 \ldots n_i \ldots n_j \ldots n_s \ldots m_R\rangle\rangle_s = C_s^{(R)} ,$$

and so on.

The total contribution to the Rth coefficient of this type will be

$$C_s^{(R)} + \sum_{k=1}^{R-1} (-1)^k \binom{R-1}{k} C_s^{(R)} = C_s^{(R)}\left[1 + \sum_{k=1}^{R-1} (-1)^k \binom{R-1}{k}\right] = 0 ,$$

since the quantity in brackets vanishes according to an elementary identity. All other contributions to the Rth coefficient can be obtained by substituting σ for s, etc., so the whole thing is zero.

APPENDIX B: QUANTUM STATISTICS

Uehling and Uhlenbeck[28] propose that for a dilute gas of identical Bose-Einstein/Fermi-Dirac particles the Boltzmann equation be replaced by one in which the *Stosszahlansatz* is modified to take account of the quantum statistics of identical particles. In the language of the present formalism, the corresponding form would be, instead of (4.8),

$$P(ij \to kl)dt = \alpha_{ij}^{kl} m_i m_j (1 + \theta m_k)(1 + \theta m_l) , \tag{B1}$$

where $\theta = 0, 1$, and -1 for MB, BE, and FD particles, respectively.

This leads to a generalized Boltzmann equation; instead of (4.22) one obtains

$$\frac{d\langle n_s\rangle}{dt} = \frac{1}{2}\sum \alpha_{ij}^{ks}\langle n_i\rangle\langle n_j\rangle(1 + \theta\langle n_k\rangle)(1 + \theta\langle n_s\rangle) - \alpha_{ks}^{ij}\langle n_k\rangle\langle n_s\rangle(1 + \theta\langle n_i\rangle)(1 + \theta\langle n_j\rangle) ; \tag{B2}$$

the equilibrium values satisfy [instead of (4.25)],

$$\bar{n}_i\bar{n}_j(1 + \theta\bar{n}_k)(1 + \theta\bar{n}_l) = \bar{n}_k\bar{n}_l(1 + \theta\bar{n}_i)(1 + \theta\bar{n}_j) \quad \text{if } \alpha_{ij}^{kl} \neq 0 , \tag{B3}$$

and one has

$$\bar{n}_i = \frac{1}{A^{-1}e^{\epsilon_i/kT} - \theta} \qquad \text{(B4)}$$

in general instead of the canonical distribution (4.26).

If one substitutes the above *Stosszahlansatz* in the master equation the whole theory goes through in the same way. One obtains by contraction the Uehling–Uhlenbeck Boltzmann equation and so forth. In addition, the results of the fluctuation theory (the Fokker-Planck equations) are unaffected (the calculations are mildly tedious but straightforward). This remains true for unphysical values of the parameter θ, and evidently for a wider class of models (corresponding to different collision forms), the extent of which remains to be established.

---

[1] R. Fox and G. E. Uhlenbeck, Phys. Fluids 13, 1893 (1970); 13, 2881 (1970). Equivalent results are reported by M. Bixon and R. Zwanzig, Phys. Rev. 187, 267 (1969). See also D. Montgomery, Phys. Fluids 12, 804 (1969) and F. L. Hinton, Phys. Fluids 13, 857 (1970), which corrects some errors in the former. The idea of adding fluctuations to the Boltzmann equation, conceived as an equation for the average distribution, seems to have appeared first in B. B. Kadomtsev, Zh. Eksp. Teor. Fiz. 32, 943 (1957) [Sov. Phys.—JETP 5, 771 (1957)], and is utilized by A. A. Abrikosov and I. M. Khalatnikov, Zh. Eksp. Teor. Fiz. 34, 198 (1958) [Sov. Phys.—JETP 7, 135 (1958)]. Here we follow the treatment of Fox and Uhlenbeck.

[2] L. Onsager and S. Machlup, Phys. Rev. 91, 1505 (1953); S. Machlup and L. Onsager, Phys. Rev. 91, 1512 (1953). An equivalent theory (spectral densities instead of time evolution) appeared in H. B. Callen and R. F. Greene, Phys. Rev. 86, 702 (1952); H. B. Callen, M. L. Barasch, and J. L. Jackson, Phys. Rev. 88, 1387 (1952); and also in M. S. Green, J. Chem. Phys. 20, 1281 (1952). The essence of the theory was presented also in N. Hashitsume, Prog. Theor. Phys. 8, 461 (1952); H. Takahasi, J. Phys. Soc. Jpn. 7, 439 (1952); T. Yamamoto, Prog. Theor. Phys. 10, 11 (1953) and Busseiron Kenkyu 57, 66 (1952) (in Japanese). The theory is outlined in textbooks such as S. R. deGroot and P. Mazur, *Non-Equilibrium Thermodynamics* (North-Holland, Amsterdam, 1962), where references to the large literature on the subject can be found.

[3] L. D. Landau and E. M. Lifshitz, Zh. Eksp. Teor. Fiz. 32, 618 (1957) [Sov. Phys.—JETP 5, 512 (1957)]; L. D. Landau and E. M. Lifshitz, *Fluid Mechanics* (Addison-Wesley, Reading, Mass., 1959) (translation of first Russian edition). A fault in the derivation is corrected in Fox and Uhlenbeck, Ref. 1. See also Bixon and Zwanzig (Ref. 1), Hinton (Ref. 1), and G. E. Kelly and M. B. Lewis, Phys. Fluids 14, 1925 (1971).

[4] See, for example, E. H. Hauge and A. Martin-Löf, J. Stat. Phys. 7, 259 (1973), and references therein.

[5] The two-particle, two-time correlation fucntion, at least, can be obtained by means like those introduced in N. Rostoker and M. N. Rosenbluth, Phys. Fluids 3, 1 (1960) and N. Rostoker, Nucl. Fusion 1, 101 (1961); see for example S. Tsugé, Phys. Lett. 26A, 235 (1968); Hinton, Ref. 1 and references therein, and Sh. M. Kogan, Teor. Mat. Fiz. 10, 143 (1972) [Theor. Math. Phys. 10, 94 (1972)].

[6] The resulting Chapman-Kolmogorov equation, which might be called the fine-grain master equation, was utilized as a basis for the kinetic theory of the Boltzmann equation in M. Kac, in *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability* (University of California Press, 1956), Vol. 3, p. 171, and M. Kac, *Probability and Related Topics in Physical Sciences* (Interscience, New York, 1959), Chap. 3. The fine-grain master equation is used to calculate the covariance of the fluctuation process for a somewhat artificial two-state model by M. Kac in *The Boltzmann Equation*, edited by E. G. D. Cohen and W. Thirring (Springer-Verlag, Vienna, 1973); Acta Phys. Austriaca Suppl. X, 379 (1973); and for a one-dimensional gas of "Maxwell molecules" [$\phi(\vec{r}) \sim r^{-4}$], in another work to appear in the present series, reported by J. L. Logan, Bull. Am. Phys. Soc. 18, 1591 (1973). Some further results in this line are contained in H. P. McKean, (unpublished).

[7] A. J. F. Siegert, Phys. Rev. 76, 1708 (1949); also J. E. Moyal, J. R. Stat. Soc. B 11, 150 (1949). This formalism has been used to study fluctuations in electron-hole levels in solids by several authors; see K. M. van Vliet, Phys. Rev. 133, A1182 (1964), and references therein.

[8] These ideas are reviewed in G. E. Uhlenbeck and G. W. Ford, *Lectures in Statistical Mechanics* (American Mathematical Society, Providence, R. I., 1963), Chaps. IV, VI, and VII, and by E. G. D. Cohen, in *Transport Phenomena in Fluids*, edited by H. J. Hanley (Dekker, New York, 1969), Chap. 6; in *Recent Advances in Engineering Science*, edited by A. C. Eringen (Gordon and Breach, New York, 1968), Vol. III, p. 125; and in *Fundamental Problems in Statistical Mechanics*, edited by E. G. D. Cohen (North-Holland, Amsterdam, 1962), p. 110.

[9] Hauge and Martin-Löf (Ref. 4) point out (as had been done earlier by Lorentz, but then forgotten) that the limit of validity involves the densities of the Brownian particle and fluid medium, $\rho/\rho_B \ll 1$, not the molecular and particle masses, as is often stated, for example, in the papers cited below; also, the particle must be so small that hydrodynamical currents are not set up, which would then lead to a non-Markovian diffusion process. Microscopic derivations of the Langevin theory for harmonically bound systems have been given by G. E. Uhlenbeck and S. Goudsmit, Phys. Rev. 34, 145 (1929), and in various, more recent works; see, for example, G. W. Ford, M. Kac, and P. Mazur, J. Math. Phys. 6, 504 (1965); P. Mazur and I. Oppenheim, Physica (Utr.) 50, 241 (1970), and references therein.

[10] Recall that the diffusion in space is given, for long times, by the Einstein diffusion law, $\langle x^2 \rangle = (4D/\beta^2)t$.

[11] See, for example, M. C. Wang and G. E. Uhlenbeck, Rev. Mod. Phys. 17, 323 (1945).

[12] L. Onsager, Phys. Rev. 37, 405 (1931); 38, 2265 (1931). The needed generalization had to wait the availability of the natural mathematical language for it, the theory of stationary Markov processes; the needed generality of formulation seems to have been provided first by J. G. Kirkwood, J. Chem. Phys. 14, 180 (1946), and the physical basis, in the application of the second law to fluctuations, by H. Nyquist, Phys. Rev. 32, 110 (1928), suggested in the first instance by Smoluchowski and Szilard— which is not to say the actual development of the theory paid any heed to this logical order.

[13] See Ref. 2. The diverse intentions and methods of these papers provide a nice gloss on the note in Ref. 12; the words Brownian motion, for example, appear nowhere in Onsager and Machlup (1953), while the (equivalent) work of Yamamoto (1953) appeared under the title "Statistical Mechanics of General Brownian Motion Underlying Irreversible Processes."

[14] Fox and Uhlenbeck (Ref. 1) provide the generalization of the treatments in Ref. 2 required for cases of mixed time symmetry, needed for hydrodynamics and the Boltzmann equation. The development here for even parity (what Onsager and Machlup call "α variables") is sufficient for the case at hand, the Boltzmann equation for a spatially homogeneous gas—so without the (odd parity) streaming operator.

[15] N. N. Bogoliubov, in *Studies in Statistical Mechanics* (North-Holland, Amsterdam, 1962), Vol. I. Pt. A, or any of the papers in Ref. 8.

[16] P. S. Lee and T. Y. Wu, Int. J. Theor. Phys. 7, 267 (1973).

[17] M. J. Klein, in *The Boltzmann Equation*, edited by E. G. D. Cohen and W. Thirring (Springer-Verlag, Vienna, 1973); Acta Phys. Austriaca Suppl. X, 379 (1973), p. 53, traces the course of arguments that led finally to this understanding from the first publication of the Boltzmann equation in Wien. Berichte 66, 275 (1872) [English translation by S. G. Brush, *Kinetic Theory II* (Pergamon, New York, 1966)], in which Boltzmann, thinking (3.3) exact and the distribution function a representation of the actual number of particles near a point, makes it serve in his program for a "purely analytical, mechanical proof" of the second law.

[18] A. Nordsieck, W. E. Lamb, and G. E. Uhlenbeck, Physica (Utr.) 7, 344 (1940).

[19] The detailed or fine-grain master-equation approach of Ref. 6.

[20] R. Brout, Physica (Utr.) 22, 509 (1956) deals with a scheme for deriving the fine-grain master equation from the Liouville equation through a systematic expansion in "collision configurations" of successive combinatorial orders; the argument is outlined in Kac (1959), Ref. 6.

[21] 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 R. C. Tolman, *The Principles of Statistical Mechanics* (Oxford U.P., New York, 1938), Secs. 38–42; and J. M. Blatt and V. F. Weisskopf, *Theoretical Nuclear Physics* (Wiley, New York, 1952), Sec. X2E.

[22] The most convenient way is to define the generating function

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

then the master equation (4.7) is equivalent to

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

and, for example,

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

See P. M. Mathews, I. I. Shapiro, and D. L. Falkoff, Phys. Rev. 120, 1 (1960), and Siegert, Ref. 7, Appendix III.

[23] L. Boltzmann, *Lectures in Gas Theory* (1896, 1898) [English translation by S. G. Brush (University of California, 1964), p. 427].

[24] Further details of this work, which was first reported by Logan, Ref. 6, are contained in J. L. Logan, Ph.D. thesis (The Rockefeller University, New York, 1974), Chap. 4 (unpublished).

[25] 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 operators will not be considered.

[26] See for example R. L. Stratonovich, *Topics in the Theory of Random Noise*, translated by R. A. Silverman (Gordon and Breach, New York, 1963), Vol. I, Chap. 4.

[27] While this paper was in preparation, we received an interesting series of papers by J. Keizer, J. Chem. Phys. (to be published), which set forth a phenomenological theory of fluctuations away from equilibrium for chemical reactions and other processes described by rate equations in terms of elementary events, and which prompted use to attempt to extend the results of the preceding section to the case far from equilibrium, with the results appearing below.

[28] E. A. Uehling and G. E. Uhlenbeck, Phys. Rev. 43, 552 (1933). See also E. J. Hellund and E. A. Uehling, Phys. Rev. 56, 818 (1939). The nature of the limit in which the generalized Boltzmann equation actually applies is somewhat unclear; see J. de Boer and R. B. Bird, in J. O. Hirschfelder, C. F. Curtiss, and R. Bird, *Molecular Theory of Gases and Liquids* (Wiley, New York, 1954), p. 668, and S. Fujita, Physica (Utr.) 27, 940 (1961). The corresponding quantum-mechanical master equation is derived in first-order perturbation theory in D. I. Blokhintsev, *Quantum Mechanics*, translated by J. B. Sykes and M. J. Kearsley (Reidel, Dordrecht, Holland, 1964), Sec. 120. A very helpful discussion appears in Tolman, Ref. 21, Chap. 10.