Differential-Operator Approximations to the Linear Boltzmann Equation*

ARMAND SIEGEL† ‡
University of Michigan, Ann Arbor, Michigan, and Boston University, Boston, Massachusetts
(Received April 19, 1960)

A measure of deviation from equilibrium of an ensemble of particles is proposed, which is physically appropriate and of especially simple form when expressed in terms of the expansion coefficients of the ensemble distribution function with respect to the system of orthogonal polynomials obtained by using the equilibrium distribution function as weight function. The linear Boltzmann operator can then be expanded in a series of terms which, under certain circumstances, may be regarded as of successively diminishing magnitude in their effect on the rate of approach to equilibrium. This expansion of the operator is different from the expansion due to Kramers (later discussed by Moyal) in derivate moments, commonly used in approximate stochastic treatments of irreversible processes. With the aid of a theorem on definite operators, it is possible to break off the series at any point and thereby obtain a correspondingly accurate approximation to the linear Boltzmann operator, whose temporal solutions tend to the correct equilibrium distribution function. The first approximation is the Fokker-Planck operator, exactly. The next approximation would be the appropriate operator to use when the stochastic variable begins to deviate appreciably from a linear dissipation law, etc. The method is applied to the "Rayleigh process" (ensemble of particles in a rarefied gas medium, the medium itself being in internal equilibrium), and the second approximation to the linear Boltzmann operator for this case is explicitly derived. A possible form for the second approximation in more general processes, suggested by this, is also given.

I. INTRODUCTION AND BASIC THEOREMS

1. Introduction

THE central theoretical tools in the study of time-varying thermal fluctuations have long been the Fokker-Planck equation and its *alter ego* the Langevin equation.[1] The use of these powerful mathematical devices has conferred on the subject a considerable degree of logical cohesion, but they limit its scope to phenomena obeying a linear friction, or dissipation, law (the terminology is defined in footnote 1). From the point of view of experiment this limitation is of no consequence at present, because there are as yet no temporal observations outside the linear friction range. But from the point of view of theory, the extension of our understanding to the nonlinear range appears desirable, because the dominating purpose of theory in this field is to bridge the gap between the fundamental theoretical postulates of kinetic theory and the phenomenological formalism, namely, thermodynamics (*sensu* antonym of thermostatics). The Fokker-Planck-Langevin formalism does make contact with thermodynamics[2,3] on the one side of this gap. But the fundamental theory is certainly nonlinear, hence "nonlinearization" of the Fokker-Planck-Langevin formalism is necessary to further the linkage.

A number of papers and reports on this subject have appeared in recent years.[4–9] Some of these make more or less tentative assumptions regarding the fundamental statistical equations governing the nonlinear systems, and go on to obtain solutions of these equations. Others emphasize only the problem of deriving and justifying the fundamental statistical equations. The present paper is addressed to this latter problem.

The thinking that underlies the present work is as follows: The Fokker-Planck equation may be rigorously derived in the case of the random walk in velocity space.[10] The random walk, as a random impact process, may be regarded as a simplified version of the *Rayleigh process*.[11] The Rayleigh process, which is defined and

* This work supported in part by the U. S. Air Force through the Air Force Office of Scientific Research, Air Research and Development Command.

† John Simon Guggenheim Memorial Foundation Fellow, 1957–58.

‡ Present address: Boston University.

[1] The Fokker-Planck equation for the temporal evolution of the probability density function $P(\xi,t)$ of a scalar variable ξ reads

$$\frac{\partial P(\xi,t)}{\partial t}=-\frac{\partial}{\partial \xi}a\xi P(\xi,t)+\frac{b}{2}\frac{\partial^2}{\partial \xi^2}P(\xi,t).$$

Here $a\xi$ is (if the equation is applicable) minus the ensemble average rate of change of ξ due to "friction" or dissipative effects in general; i.e., $\langle\dot{\xi}\rangle=-a\xi$. For a particle undergoing Brownian motion, $\langle\dot{\xi}\rangle$ is literally due to friction, being attributable to viscosity; more generally, ξ may be any thermodynamic observable in its range of linear dissipation, according to the theories referred to in footnotes 2 and 3. The constant b is (again, if the equation is applicable) a measure of the amplitude of thermal fluctuations, or "noise."

In the mathematically equivalent Langevin formalism, $\xi(t)$ is a random function of time satisfying the Langevin equation

$$\dot{\xi}+a\xi=(b)^{\frac{1}{2}}\epsilon(t),$$

where $\epsilon(t)$ is the "ideal random function" normalized so that

$$\left\langle\left[\int_0^t\epsilon(t)dt\right]^2\right\rangle=t.$$

$\xi(t)$ will then be found to have a probability density satisfying the Fokker-Planck equation as just given.

Introductory treatments of these matters will be found in the well-known review articles by S. Chandrasekhar [Revs. Modern Phys. **15**, 1 (1943)] and by M. C. Wang and G. E. Uhlenbeck [Revs. Modern Phys. **17**, 323 (1945)].

[2] N. Hashitsume, Progr. Theoret. Phys. (Kyoto) **8**, 461 (1952).
[3] L. Onsager and S. Machlup, Phys. Rev. **91**, 1505 (1953).
[4] D. K. C. MacDonald, Phys. Rev. **108**, 541 (1957).
[5] N. G. van Kampen, Phys. Rev. **110**, 319 (1958).
[6] R. O. Davies, Physica **24**, 1055 (1958).
[7] C. T. J. Alkemade, Physica **24**, 1029 (1958).
[8] M. Lax, Revs. Modern Phys. **32**, 25 (1960).
[9] N. G. van Kampen (unpublished report, 1959).
[10] See S. Chandrasekhar or M. C. Wang and G. E. Uhlenbeck, cited in footnote 1.
[11] Lord Rayleigh, *Scientific Papers* (Cambridge University Press, New York, 1902), Vol. 3, p. 273; discussed by C. S. Wang Chang and G. E. Uhlenbeck, *Kinetic Theory of a Gas in Alternating Outside Force Fields*, Engineering Research Institute Report 2457-3-T (University of Michigan, Ann Arbor, Michigan, 1956).

discussed in Sec. II and Appendix A of this paper, is a process which is itself simple enough for easy, explicit mathematical description, yet real enough to embody some basic features of thermal fluctuation phenomena. In the simplified version referred to in the foregoing, the friction dependence is linear, but in the exact formulation it is definitely nonlinear. Thus it might be possible to derive from it *a* counterpart, if not *the* counterpart, of the Fokker-Planck equation for the nonlinear friction region.

The probability density of the random variable in a Rayleigh process obeys a linear Boltzmann equation, the operator of which contains an explicit expansion parameter. When the operator is appropriately expanded in terms of this parameter (this is done in Sec. II, where it is shown that this expansion is different from the customarily employed Kramers or Moyal expansion), the condition that the probability density tends to the known equilibrium form can be applied to the problem of approximating this series. The first approximation is, as it must be, the Fokker-Planck equation. The second approximation is a sixth-order differential operator of precisely defined form containing two independent physical parameters (in addition to that of the first approximation), and an arbitrary parameter which does not affect any experimentally measurable results. The method is, moreover, a general one and yields approximations of arbitrary order. Thus there is a regular sequence of approximations linking the Fokker-Planck and linear Boltzmann equations.

The paper is organized as follows: Since the result may have validity for processes other than the Rayleigh process, the subsections of Sec. I which follow this introductory section discuss the general case of a linear Boltzmann-operator expansion having the necessary properties, and state and prove the theorems for the construction of satisfactory approximations from this expansion. In Sec. II the Rayleigh process is described, certain necessary expressions are derived from it, and the general theorem is applied. In Sec. III certain generalizations suggested by the Rayleigh process analysis are discussed. The equation generalizing the Fokker-Planck equation to cubic friction is then given for a hypothetical process which is mathematically similar to the Rayleigh process but which does not possess an explicit expansion parameter, or for which the parameter is unknown.

2. Précis of Method for Construction of Successive Approximations to the Linear Boltzmann Operator

Consider the linear Boltzmann equation for an ensemble of particles moving in one dimension. Assume no force field, and that the particles have already attained a spatially uniform density. The distribution function will then depend only on velocity and time: We write $P(V,t)$ for the probability density of velocity V, normalized to unity, as a function of time. The linear Boltzmann equation will then be

$$\partial P/\partial t = BP, \tag{1}$$

where B is a linear integral operator. The function BP is given, more explicitly, in terms of a kernel $B(V,V')$ as

$$\int B(V,V')P(V')dV'. \tag{2}$$

The eigenvalues of B must all be negative, except for a nondegenerate zero eigenvalue which has the Maxwell-Boltzmann distribution function

$$F(V) = (2\pi V_R^2)^{-\frac{1}{2}} \exp\left(-\frac{V^2}{2V_R^2}\right) \tag{3}$$

(V_R = root-mean-square value of V in the equilibrium distribution) (4)

as its eigenfunction. This "equilibrium requirement" ensures that an arbitrary initial distribution will always decay into the Maxwell-Boltzmann distribution. It may be equivalently formulated by saying that B must be *negative semidefinite*, in the sense that

$$\int_{-\infty}^{\infty} Y(V)B(V,V')Y(V')F(V')dVdV' \leqq 0 \tag{5}$$

for any polynomial $Y(V)$, with the equality sign holding only for Y = constant.

In the following discussion we take B to be in a Hermitian matrix representation. The vector corresponding to $F(V)$ is then the (unique) eigenvector of B for eigenvalue zero. The negative semidefiniteness requirement will take the form

$$\sum_{r,s} u_r B_{rs} u_s \leqq 0 \tag{6}$$

for all normalizable vectors (u_m), with the equality holding when (u_m) corresponds to $F(V)$. (Details of such a representation will be given in Sec. I.3.)

Suppose B to depend on some parameter λ with respect to which it may be expanded in a convergent series:

$$B = c(\lambda) \sum_{m=0}^{\infty} \lambda^m b_m, \tag{7}$$

where $c(\lambda)$ is a positive c-number function of λ and b_m is a matrix independent of λ. The existence of such an expansion suggests the possibility of approximating B, for small values of λ, by terminating the series at some finite value of m. In so doing, however, it will be important to retain the negative semidefiniteness property in the approximate operator: Lack of this property will imply the existence of at least one eigenvector of B which grows, instead of decaying, expo-

nentially with time; if present in the initial distribution, in however small an admixture, this mode or modes will grow in amplitude indefinitely large with time, hence the Maxwell-Boltzmann distribution will never be reached.

It is not possible to prove from the negative semidefiniteness of B that an operator obtained by terminating the series (7) is negative semidefinite. In fact, we shall find that the model to be discussed below furnishes a counterexample to such a supposition. However, it may be possible to retain the negative semidefiniteness property by a simple construction. This is based on the following factorization theorem[12]: *A positive semidefinite Hermitian matrix can always be written as the product of some (suitably chosen) matrix Q and its adjoint $Q^\dagger$.* Let us put, for the sum in Eq. (7),

$$\mathbf{B}=\sum_{m=0}^{\infty}\lambda^m b_m. \tag{8}$$

Then, applying this theorem to $\mathbf{B}$, which is negative semidefinite like B,

$$\mathbf{B}=-QQ^\dagger. \tag{9}$$

[N.B.: Q is determined only to within a unitary postmultiple; if U is unitary, $QU(QU)^\dagger=QU\cdot U^{-1}Q^\dagger=QQ^\dagger$.] Suppose now that Q also can be expanded in terms of λ:

$$Q=\sum_{l=0}^{\infty}\lambda^l q_l. \tag{10}$$

For the operator obtained by taking the first m_0 terms of Q, write

$$Q^{(m_0)}=\sum_{l=0}^{m_0}\lambda^l q_l. \tag{11}$$

Then

$$\mathbf{B}=-Q^{(m_0)}Q^{(m_0)\dagger}+o(\lambda^{m_0}) \tag{12}$$

where $o(\lambda^{m_0})$ is an operator of order higher than λ^{m_0}; i.e., $\mathbf{B}$ may be approximated to order λ^{m_0} by $-Q^{(m_0)}Q^{(m_0)\dagger}$, a form analogous to the exact factorization $QQ^\dagger$.

We now assert that $-Q^{(m_0)}Q^{(m_0)\dagger}$ is a negative semidefinite matrix. That it is at least negative definite follows from its very form, since for any vector $\mathfrak{u}$, $\mathfrak{u}\cdot Q^{(m_0)}Q^{(m_0)\dagger}\mathfrak{u}=(Q^{(m_0)\dagger}\mathfrak{u})\cdot(Q^{(m_0)\dagger}\mathfrak{u})\geqq 0$. This being so, it will moreover be negative semidefinite in the sense desired if F [Eq. (3)] is an eigenvector, with eigenvalue zero. This can be shown to be true, as follows: $Q^\dagger F=0$ since by hypothesis [remark following Eq. (5)] $(F,QQ^\dagger F)=(Q^\dagger F,Q^\dagger F)=0$. But if $Q^\dagger$ annihilates F, and if the series of Eq. (10) represents Q over some nonzero range of λ, then the individual $q_l^\dagger$ must also annihilate F; since, putting $q_l^\dagger F=F_l$, we must have $\sum\lambda^l F_l=0$ over this range of λ, which means that the individual F_l must vanish; whence $Q^{(m_0)\dagger}$ annihilates F, and F is an eigenvector of $Q^{(m_0)}Q^{(m_0)\dagger}$ for eigenvalue zero, q.e.d.

We use a subscript m_0 to denote a negative semidefinite approximation to $\mathbf{B}$ of order λ^{m_0} as constructed in the foregoing, i.e.,

$$\mathbf{B}_{m_0}=-Q^{(m_0)}Q^{(m_0)\dagger}. \tag{13}$$

It should be noted that $\mathbf{B}_{m_0}$ could, unlike $\mathbf{B}$, be degenerate with respect to the zero eigenvalue, so far as the present proof goes. If this should happen, it would of course be quite unsuitable, since the time-asymptotic distribution, in contradiction to the H theorem, would not be uniquely the Maxwell-Boltzmann distribution, but would depend on the initial distribution. Until more is known, the success of the method sketched above in yielding an approximation suitable in this respect cannot be guaranteed in advance; individual cases to which it is applied will have to be inspected after the event for satisfaction of this criterion.

Construction of $Q^{(m_0)}$

If we substitute the series expressions for Q and for $\mathbf{B}$ into Eq. (9), and equate coefficients of like powers of λ, we obtain an infinite set of equations

$$-\tfrac{1}{2}\sum_{j=0}^{m}(q_j q_{m-j}^\dagger+q_{m-j}q_j^\dagger)=b_m;\quad m=0,1,2,\cdots. \tag{14}$$

The first m_0+1 equations of this set involve only the first $(m_0+1)q$'s, $q_0, q_1\cdots q_{m_0}$. Suppose we have a solution to these first m_0+1 equations, say $q_0', q_1', \cdots q_{m_0}'$; primes are used to allow for the possibility that these may not agree with the first m_0 terms of Q itself (i.e., of the solution to the infinite set of equations), even after allowing for the possibility of an arbitrary common unitary postmultiple. From these we can construct an operator

$$\mathbf{B}_{m_0}'=-Q'^{(m_0)}Q'^{(m_0)\dagger}, \tag{15}$$

where

$$Q'^{(m_0)}=\sum_{l=0}^{m_0}\lambda^l q_l'. \tag{16}$$

$\mathbf{B}_{m_0}'$, like $\mathbf{B}_{m_0}$, agrees with $\mathbf{B}$ to order λ^{m_0}. In this way, if Eqs. (14) can be solved for $m=0, 1, 2\cdots m_0$, we have successive approximations up to m_0th order in λ to $\mathbf{B}$.

3. Estimation of the Degree of Deviation from Equilibrium

The essential nature of B is to drive its operand $P(V,t)$ toward the equilibrium function $F(V)$. Thus when it is expanded, the increasing smallness of its successive terms should be with respect to their effectiveness in this sense. With this in mind, we adopt the

[12] F. D. Murnaghan, *The Theory of Group Representations* (The Johns Hopkins Press, Baltimore, 1938), p. 20.

following as a measure of the degree of deviation of $P(V,t)$ from the equilibrium distribution $F(V)$:

$$\chi_0^2=\int\frac{[P(V,t)-F(V)]^2}{F(V)}dV. \tag{17}$$

It will be noted that this is the same, to within a constant factor, as Pearson's noted χ^2 of statistical theory,[13] with $F(V)$ the hypothetical and $P(V,t)$ the sampling distribution, and with an infinitely fine subdivision of the range of V. However, since there are a number of possible measures of "goodness of fit" of a distribution, mere coincidence with one of these which happens to be famous is not sufficient reason for its adoption; it is necessary to demonstrate the suitability of the choice (17). Our reasons for adopting it are the following:

(1) It emphasizes deviations from equilibrium according to the magnitude of the V values involved: Since $F(V)$ will be essentially localized in the region of equilibrium values of V, it is increasingly small for increasingly large deviations of V from its equilibrium range, and with $F(V)$ in the denominator of the integrand of χ_0^2 these large deviations are increasingly heavily weighted. This is appropriate in a study of the approach to equilibrium, because a given amount of probability added to or taken away from the equilibrium distribution curve in the neighborhood of some V value becomes increasingly important in its effects with increasing deviation from the rms value of V.

(2) It is precisely adapted to formulation in terms of a matrix representation, and therefore to the utilization of the theorem of Sec. I.2: given a set of polynomials $p_r(V)$ orthogonalized with respect to $F(V)$ as weight function (in particular, the Hermite polynomials),

$$\int p_r(V)p_s(V)F(V)dV=(N_r/N_0)\delta_{rs}, \tag{18}$$

where N_r is a normalization constant. Expand $P(V,t)$ in terms of normalized functions $(N_0/N_r)^{\frac{1}{2}}p_r(V)F(V)$:

$$P(V,t)=N_0{}^{\frac{1}{2}}\sum_{r=0}^{\infty}a_r(t)N_r^{-\frac{1}{2}}P_r(V)F(V). \tag{19}$$

Then

$$\chi_0^2=N_0\int[\sum{}'a_r(t)N_r^{-\frac{1}{2}}P_r(V)]^2F(V)dV=\sum{}'a_r^2, \tag{20}$$

where the prime on the summation sign denotes omission of the term $r=0$. Noting that the a_r are the components of the Hilbert-space vector $P(V,t)$, we see that χ_0^2 is now represented by the squared length of the part of $P(V,t)$ orthogonal to the equilibrium function. Thus, when B is put into a matrix representation having the above as basis functions, its tendency to promote equilibrium will be measured in the simplest possible way—by its effect on the components of its vector operand.

We are now in a position to relate the approach to equilibrium to the series expansion of B. The matrix elements of B are

$$B_{rs}=N_0(N_rN_s)^{-\frac{1}{2}}\int\int p_r(V)B(V,V')p_s(V')dVdV'. \tag{21}$$

[Because of the detailed balancing condition, $B(V,V')F(V')=B(V',V)F(V)$, these are Hermitian.] If Eq. (1) is taken in matrix form, and B expanded according to Eq. (7), we find for the rate of decrease of χ_0^2,

$$\frac{d\chi_0^2}{dt}=2c(\lambda)\sum_{m=0}^{\infty}\lambda^m\sum_{r,s}{}'a_r(b_m)_{rs}a_s. \tag{22}$$

Thus if $\lambda\ll1$, the successive matrices $b_m(m=0, 1, 2, \cdots)$ make rapidly decreasing contributions to the trend to equilibrium, at least if the sums which are the coefficients of the λ^m do not increase markedly with m. Thus we have achieved the interpretation of the expansion (7) which was sought at the beginning of this section.

4. Summary of Method

Successive approximations to the Boltzmann operator with respect to its tendency to promote equilibrium may be constructed then along the following lines: (1) B must be put into a matrix representation, using as basis the orthogonal [with respect to $F(V)$ as weighting function] functions $p_r(V)$; (2) the matrix B is then expanded (if possible) in terms of a small parameter λ; (3) the procedure of Sec. I.2 is then used to obtain a negative semidefinite approximation of the desired order in λ; (4) the coefficients of successive powers of λ in the expansion of $d\chi_0^2/dt$ must not increase too rapidly.

We do not discuss requirement (4) in the present paper, but merely apply the method outlined assuming it to be satisfied. We shall show that the zeroth approximation thus obtained is exactly the Fokker-Planck equation, which is evidence that this procedure is justified. Before leaving this subject, however, it might be mentioned that from the general behavior of the successive matrices b_m in the special case to be studied in Sec. II of this paper, it appears [see Eq. (76)] that they satisfy requirement (4) to the extent that $P(V,t)$ approximates to the equilibrium function; whence higher terms of the expansion become important, despite the decreasing values of λ^m, with highly disequilibrated ensembles.

[13] See, e.g., H. Cramér, *Mathematical Methods of Statistics* (Princeton University Press, Princeton, New Jersey, 1945), Chap. 30.

5. Considerations Related to the H Theorem

Brinkman[14] has suggested the application, in the study of the linear Boltzmann process, of the criterion

$$\dot{F} \leqq 0, \tag{23}$$

where

$$F = U - TS, \tag{24}$$

U and S being the internal energy and entropy of the set of particles whose distribution function obeys the postulated linear Boltzmann equation, and T the temperature of the medium whose molecules generate the random motion of the particles. Since Brinkman did not consider the linear Boltzmann equation as such, it is of interest to see how his criterion fits into the present work.

First, to settle on a simple nomenclature, let us agree henceforth to restrict the term "particle" to the individuals whose random motion is being studied, and the term "molecule" to the constituent individuals of the medium. Assume both particles and molecules to be uniformly distributed in space. If $f(v,t)$ is the distribution function of molecules with respect to their velocity variate v, and $P(V,t)$ that of particles with respect to their velocity variate V, then the H function per unit volume of the combined systems is

$$H(t) = \int f(v,t) \log f(v,t) dv + \int P(V,t) \log P(V,t) dV. \tag{25}$$

The application of the *Stosszahlansatz* to collisions of *all* kinds will lead to the usual nonlinear coupled Boltzmann equations for the two distributions, and thereby to the H theorem:

$$\dot{H} = \int \dot{f} \log f dv + \int \dot{P} \log P dV \leqq 0. \tag{26}$$

Now let us specialize to the conditions under which the linear Boltzmann equation holds for the particles—to the "linear Boltzmann regime," as we shall call it. $f(v,t)$ is then negligibly different from the equilibrium Maxwell-Boltzmann distribution. This does not mean that $\dot{f}$ is necessarily to be neglected, but merely that its effect relative to f is. As for $\dot{f}$, it has a contribution due to collisions of molecules with one another, and one due to collisions with the particles. Under the linear Boltzmann regime, the former can be made arbitrarily small independently of the latter—for example, by making the medium sufficiently rarefied while compensating for this by increasing the size of the particles; thus we assume it to be negligible.

For the second term in (26) we shall have

$$\int \dot{P} \log P dV = -\dot{S}/k. \tag{27}$$

As for the first term, the f contribution due to collisions between molecules and particles cannot be neglected, unlike the other contribution, since these are the collisions responsible for the process itself. But since the molecules are in equilibrium,

$$\log f = -\tfrac{1}{2} m v^2/kT + \text{const}, \tag{28}$$

and

$$\int \dot{f} \log f dv = -\frac{\dot{U}(\text{medium})}{kT} = +\frac{\dot{U}}{kT}. \tag{29}$$

We will thus have from (26), with $H = -S/k$,

$$\dot{F} \leqq 0, \tag{30}$$

showing that Brinkman's criterion is necessarily satisfied under the linear Boltzmann regime, i.e., $\dot{F}$ will decrease when $P(V,t)$ satisfies the linear Boltzmann equation.

We shall return to this matter at the end of Sec. II.5.

II. APPLICATION TO THE RAYLEIGH MODEL

1. Linear Boltzmann Operator for the Rayleigh Model

As an illustration of the method described above, we apply it to a simplified case which is still interesting from the kinetic theory point of view; that of the random velocity of a particle suspended in a rarefied gas in internal equilibrium. In order to simplify the mathematical analysis, we introduce certain artifices which are more or less familiar in this classical problem. We study the random motion of *particles* (as defined in Sec. I.5) of mass M. Their random motion results from collisions with the *molecules* of the rarefied gas, of mass m. The molecules have uniform spatial density ρ and a Maxwell-Boltzmann distribution with respect to the molecular velocity $\mathbf{v}$,

$$(2\pi v_R^2)^{-\frac{3}{2}} \exp\left(-\frac{\mathbf{v}^2}{2v_R^2}\right), \tag{31}$$

where

$$v_R^2 = \text{rms value of any component of } \mathbf{v} = kT/m. \tag{32}$$

The mass ratio m/M will at a later stage be assumed less than one, but not necessarily very small. We assume the gas sufficiently rarefied, and the particle concentration low enough, for the initial Maxwell-Boltzmann distribution of the gas molecules not to change appreciably in time, no matter what the initial velocity distribution of the particles. The particle concentration is also to be so low that collisions of particles with one another occur with negligible frequency, i.e., the random velocity changes of the particles are entirely caused by collisions with the molecules.

For further simplicity we take the particles to be infinitely thin disks, each constrained to move only in a direction perpendicular to its plane. This constraint, while artificial, is of a purely passive nature and does not dynamically affect the spontaneous statistical fluc-

[14] H. C. Brinkman, Physica 23, 82 (1957).

tuations responsible for the random process studied; in particular, we shall obtain in lowest approximation exactly the classical Browian motion.

With the foregoing constraint, the distributed variable of the disk distribution is the velocity component along the line of the allowed motion, which we may call V. However, we shall use instead of V the dimensionless

$$y=V/V_R \tag{33}$$

[V_R is defined in Eq. (4); here $V_R=(kT/M)^{\frac{1}{2}}$].

The linear Boltzmann equation of the foregoing process, which is obeyed by $P(y,t)$, the probability density function in y, is

$$\frac{\partial P(y',t)}{\partial t}=\int B(y',y'')P(y'',t)dy'', \tag{34}$$

where the linear Boltzmann operator $B(y',y'')$ has the form

$$B(y',y'')=C(y'|y'')-\delta(y'-y'')\int C(y|y')dy. \tag{35}$$

It is shown in Appendix A that

$$C(y''|y')=V_R(\mu/2\pi)^{\frac{1}{2}}\left(\frac{1+\mu}{2\mu}\right)^2|y''-y'|$$
$$\times\exp\left\{-\frac{1}{8\mu}[(1+\mu)y''-(1-\mu)y']^2\right\}. \tag{36}$$

$C(y''|y')$ is the transition probability density-in-y'' for a particle having initial velocity y'.

2. Expansion in Kramers Series

The operator B can be expanded in Kramers series[15–19]

$$B=\sum_{n=1}^{\infty}\frac{1}{n!}\left(-\frac{\partial}{\partial y'}\right)^n\alpha_n(y'), \tag{37}$$

in terms of the "derivate moments" α_n (Moyal's[18] terminology)

$$\alpha_n(y')=\int_{-\infty}^{\infty}(y''-y')^nC(y''|y')dy''. \tag{38}$$

(B is now an infinite-order differential operator, not an integral operator.)

Some simplification results if $\alpha_n(y')$ is replaced by the function

$$A_n(y')=\frac{1}{V_R\mu^{\frac{1}{2}}n!}\left(\frac{1+\mu}{\mu}\right)^n\alpha_n(y') \tag{39}$$

and if in the integral for $\alpha_n(y')$ the transformation

$$x=\frac{1+u}{2\mu}(y''-y'),\quad y=y' \tag{40}$$

is introduced (x is then the new variable of integration, and y the new argument of α_n). On substituting (36) into the integral for α_n, we then find

$$A_n(y)=\frac{2^n}{(2\pi)^{\frac{1}{2}}n!}\int\exp\left[-\frac{\mu}{2}(x+y)^2\right]x^n|x|dx, \tag{41}$$

and the expansion (37) becomes

$$B=V_R\mu^{\frac{1}{2}}\sum_{n=1}^{\infty}\left(\frac{\mu}{1+\mu}\right)^n\left(-\frac{\partial}{\partial y}\right)^nA_n(y). \tag{42}$$

3. Discussion of the Kramers Expansion

The simple appearance of the Kramers expansion is somewhat misleading from the point of view of the study of the approach to equilibrium, since it is not clear whether or how the successive terms represent decreasing contributions to the equilibrium-seeking tendency. In fact, we shall find that the expansion of B according to the matrix method of Sec. I.3, and which is adapted to the especially simple criterion of deviation from equilibrium there introduced, *is quite different from the Kramers expansion.* Thus the Kramers expansion plays no fundamental role in this work; however, we shall find it useful in deriving actual expressions for matrices.

The foregoing assertion is foreshadowed by the result of the traditional "random walk" analysis of the velocity-space progress of a particle subject to successive independent random impacts, which is a valid approximation in the present kind of system in the limit of vanishingly small μ. But it is certainly not clear how one could, at least in any offhand way, apply this limiting process to Eq. (42) directly. The "random walk" derivation, to be sure, does use the Kramers expansion, but it uses limiting approximations for the derivate moments from the outset, and does not make explicit use of μ.

On the other hand, Wang Chang and Uhlenbeck[11] obtained the Fokker-Planck equation directly from the linear Boltzmann equation by combining with the assumption of vanishing mass ratio the assumption that the velocity variable of the particles never gets much larger than the rms value it would have in equilibrium. [In our case this would amount to assuming $P(y,t)$ negligible for y much larger than μ.] It can be shown that these combined requirements are equivalent to taking the first two terms only in the expansion (42), and simultaneously approximating A_1 and A_2 by their lowest-order terms in y, these being of first and zeroth

[15] H. A. Kramers, Physica **7**, 284 (1940).
[16] S. Chandrasekhar, Revs. Modern Phys. **15**, 1 (1943).
[17] M. C. Wang and G. E. Uhlenbeck, Revs. Modern Phys. **17**, 323 (1945).
[18] J. E. Moyal, J. Roy. Stat. Soc. (London) **B11**, 150 (1949).
[19] J. Keilson and J. E. Storer, Quart. Appl. Math. **10**, 243 (1952).

order, respectively[20]; and this same result will be seen to follow quite simply from our analysis. The work that will be presented here extends the procedure of Wang Chang and Uhlenbeck in that successive approximations, rather than a single limiting approximation, can be obtained.

3. Symmetrization of B; Basis Functions of the Matrix Representation

The kernel of Eq. (34) can be symmetrized by the transformation

$$\bar{B}(y',y'')=\exp\left(\frac{y'^2}{4}\right)B(y',y'')\exp\left(-\frac{y''^2}{4}\right); \quad (43)$$

with the accompanying transformation of the distribution function,

$$\bar{P}(y')=\exp\left(\frac{y'^2}{4}\right)P(y'), \quad (44)$$

the linear Boltzmann equation (34) is unchanged in form:

$$\frac{\partial\bar{P}(y',t)}{\partial t}=\int\bar{B}(y',y'')\bar{P}(y'',t)dy''. \quad (45)$$

Being symmetric, $\bar{B}$ is Hermitian with respect to an unweighted inner product, thus for any two functions $\varphi(y)$, $\psi(y)$, we henceforth define

$$(\varphi,\psi)=\int_{-\infty}^{\infty}\varphi^*(y)\psi(y)dy. \quad (46)$$

This definition of the inner product is of heuristic convenience for the geometrical interpretation, and facilitates the use of standard (at least to mathematical physicists) definitions of the Hermite functions, which we shall use extensively.

The corresponding transformation of B as a *differential* operator [Eq. (42)] is

$$\bar{B}=\exp\left(\frac{y^2}{4}\right)B\exp\left(-\frac{y^2}{4}\right)=V_R\mu^{\frac{1}{2}}\sum_{n=1}^{\infty}\left(\frac{\mu}{1+\mu}\right)^n \times\exp\left(\frac{y^2}{4}\right)\left(-\frac{\partial}{\partial y}\right)^n\exp\left(-\frac{y^2}{4}\right)A_n(y). \quad (47)$$

Since the equilibrium function of the linear Boltzmann operator of the process we are considering is $(2\pi)^{-\frac{1}{2}}\exp(-y^2/2)$, we use as the basis for our matrix representation, for the reasons given in Sec. I.3, the Hermite polynomials that are orthogonal in the sense of Eq. (18) with respect to this as weighting function. But with the definition of inner product we are now using, Eq. (46), the basis consists rather of the Hermite *functions*

$$h_r(y)=\bar{C}^r\exp\left(-\frac{y^2}{4}\right), \quad (48)$$

where $\bar{C}$ is the "creation operator"

$$\bar{C}=\exp\left(\frac{y^2}{4}\right)\left(-\frac{d}{dy}\right)\exp\left(-\frac{y^2}{4}\right). \quad (49)$$

These satisfy

$$(h_r,h_s)=N_r\delta_{rs}, \quad (50)$$

where

$$N_r=(2\pi)^{\frac{1}{2}}r! \quad (51)$$

The Hermite *polynomials* are

$$H_r(y)=\exp\left(\frac{y^2}{4}\right)h_r(y). \quad (52)$$

They are the counterparts of the polynomials $p_r(V)$ of Sec. I.3. If in that section we take N_0 (which is arbitrary at that stage)

$$N_0=(2\pi)^{\frac{1}{2}}, \quad (53)$$

let y correspond directly to V, and put

$$F(y)=(2\pi)^{-\frac{1}{2}}\exp(-y^2/2), \quad (54)$$

then N_0 times formula (18) corresponds exactly to (50).

The matrix elements of B as defined by (21) now become matrix elements of B with respect to the new definition of inner product:

$$B_{rs}(\text{sense of Sec. I.3})=\left(\frac{h_r}{N_r^{\frac{1}{2}}},B\frac{h_s}{N_s^{\frac{1}{2}}}\right) =\bar{B}_{rs}(\text{sense of this section}). \quad (55)$$

Henceforth matrix elements are to be understood as defined according to the *second* of the foregoing equalities.

For χ_0^2, Sec. I.3, we now have

$$\chi_0^2=N_0\int(\bar{P}-\bar{F})^2dy=N_0(\bar{P}-\bar{F},\bar{P}-\bar{F}). \quad (56)$$

4. Matrix Expansion of $\bar{B}$

$\bar{B}$ can now be written in terms of creation operators

$$\bar{B}=V_R\mu^{\frac{1}{2}}\sum_{n=1}^{\infty}\left(\frac{\mu}{1+\mu}\right)^n\bar{C}^nA_n(y). \quad (57)$$

We are now ready to derive the matrix expansion of $\bar{B}$ analogous to Eq. (7). The crucial step for this derivation is to expand the A_n in Hermite polynomials

$$A_n(y)=\sum_{k=0}^{\infty}A_{nk}H_k(y), \quad (58)$$

where the A_{nk} are constant coefficients. It is just this

[20] As pointed out by H. A. Kramers (footnote 15), the derivate moments α_n are even or odd functions as n is an even or odd number.

device that will be seen to make possible a simple expansion of the matrix of $\bar{B}$.

It is shown in Appendix B that the expansion coefficients A_{nk} have the form

$$A_{nk}=(1+\mu)^{-\frac{1}{2}}\left(\frac{\mu}{1+\mu}\right)^{(k-n-2)/2}a_{nk}, \tag{59}$$

where a_{nk} is independent of μ. It should be noted that a_{nk} is nonvanishing only when n and k have the same parity.[20] This gives

$$\bar{B}=\left(\frac{\mu}{1+\mu}\right)^{\frac{1}{2}}V_R\sum_{n=1}^{\infty}\sum_{(k)}\left(\frac{\mu}{1+\mu}\right)^{(n+k-2)/2}a_{nk}\bar{C}^nH_k(y), \tag{60}$$

in which the sum denoted by (k) is over all positive k values having the same parity as n; or, by transforming the indices of summation

$$\bar{B}=\left(\frac{\mu}{1+\mu}\right)^{\frac{1}{2}}V_R\sum_{m=0}^{\infty}\left(\frac{\mu}{1+\mu}\right)^{m}\sum_{p=-m}^{m+1}a_{m+p+1,m-p+1}\times\bar{C}^{m+p+1}H_{m-p+1}(y). \tag{61}$$

The matrix form of $\bar{B}$ is obtained almost immediately from the foregoing. We define the matrix element of $\bar{C}$ as the inner product between *normalized* Hermite functions:

$$C_{rs}=\left(\frac{h_r}{N_r^{\frac{1}{2}}},\bar{C}\frac{h_s}{N_s^{\frac{1}{2}}}\right)=r^{\frac{1}{2}}\delta_{r,s+1}. \tag{62}$$

In order to ascertain the properties of $H_r(y)$ as an operator, it is convenient to introduce the destruction operator

$$\bar{D}=y/2+d/dy, \tag{63}$$

whose effect on the Hermite functions is given by

$$\bar{D}h_r(y)=rh_{r-1}(y), \tag{64}$$

and whose matrix element with respect to normalized Hermite functions is

$$D_{rs}=\left(\frac{h_r}{N_r^{\frac{1}{2}}},\bar{D}\frac{h_s}{N_s^{\frac{1}{2}}}\right)=s^{\frac{1}{2}}\delta_{r,s-1}. \tag{65}$$

Since

$$\bar{C}=y/2-d/dy, \tag{66}$$

we have

$$y=\bar{C}+\bar{D}. \tag{67}$$

The matrix characterization of $H_r(y)$ then follows immediately from its functional form and from Eqs. (62) and (65), if one substitutes for y using Eq. (67)

$$H_r(y)=H_r(\bar{C}+\bar{D}). \tag{68}$$

The important thing about H_r is that its matrix elements with respect to normalized Hermite functions are *independent of* μ.

The upshot is that Eq. (61) may just as correctly stand for the matrix equation giving the matrix expansion of $\bar{B}$, as for a differential operator equation; and in this matrix expansion the coefficients a_{nk} and the matrices $\bar{C}^n$ and H_k are independent of μ. Thus Eq. (61) corresponds to the desired expansion (7), provided

$$\lambda=\mu/(1+\mu). \tag{69}$$

5. Fokker-Planck Equation and the Next Approximation

We write $\bar{B}^{(m_0)}$ for the approximation to $\bar{B}$ obtained by terminating the sum in Eq. (61) at $m=m_0$. Then from the expressions derived in Appendix B we find, for $m_0=0$,

$$\bar{B}^{(0)}=8V_R\left(\frac{\mu}{2\pi(1+\mu)}\right)^{\frac{1}{2}}(-\bar{C}y+\bar{C}^2)$$

$$=-8V_R\left(\frac{\mu}{2\pi(1+\mu)}\right)^{\frac{1}{2}}\bar{C}\bar{D}. \tag{70}$$

This stage corresponds to the case $m_0=0$, $\mathbf{B}_0=b_0=q_0q_0^\dagger$, of Eq. (14), and the approximate operator is exactly factorizable. Being in the form of a negative numerical factor times $-\bar{C}\bar{C}^\dagger=-\bar{C}\bar{D}$, it very transparently exhibits the negative semidefinite property. To get the operator which operates on the true probability density function, we invert the transformation (43) on $\bar{B}^{(0)}$, $\bar{C}$, and $\bar{D}$. We find

$$C=\exp\left(-\frac{y^2}{4}\right)\bar{C}\exp\left(\frac{y^2}{4}\right)=-\frac{d}{dy}, \tag{71}$$

and

$$D=\exp\left(-\frac{y^2}{4}\right)\bar{D}\exp\left(\frac{y^2}{4}\right)=y+\frac{d}{dy}. \tag{72}$$

Thus

$$B^{(0)}=8V_R\left(\frac{\mu}{2\pi(1+\mu)}\right)^{\frac{1}{2}}\cdot\frac{d}{dy}\left(y+\frac{d}{dy}\right), \tag{73}$$

which is the Fokker-Planck operator, as promised. Note the convenience of the form (70), from which it can be seen by inspection that the eigenfunctions of $\bar{B}^{(0)}$ are the $h_r(y)$, and that the eigenvalues (also those of $B^{(0)}$) are $-8nV_R[\mu/2\pi(1+\mu)]^{\frac{1}{2}}$.

$\bar{B}^{(1)}$, the next approximation to $\bar{B}$ which would be obtained by an uncritical inspection of Eq. (61), involves the addition to $\bar{B}^{(0)}$ of the operator

$$V_r\left(\frac{\mu}{1+\mu}\right)^{\frac{1}{2}}[a_{13}\bar{C}H_3+a_{22}\bar{C}^2H_2+a_{31}\bar{C}^3H_1+a_{40}\bar{C}^4H_0]. \tag{74}$$

This must be simplified. From Eq. (68) and the definitions of the polynomials H_r, it may be written as a function of creation and destruction operators. In the resulting expression all $\bar{D}$ operators may be moved to

the right of all $\bar{C}$ operators with the aid of the commutation relation, and the numerical values of the a_{nk} may then be substituted. When this is done, we find

$$\bar{B}^{(1)}=-8V_R\left(\frac{\mu}{2\pi(1+\mu)}\right)^{\frac{1}{2}} \times\left[\bar{C}\bar{D}+\frac{1}{6}\frac{\mu}{1+\mu}(\bar{C}^3\bar{D}-6\bar{C}^2\bar{D}^2+\bar{C}\bar{D}^3)\right]. \quad (75)$$

It is evident by inspection that this expression is Hermitian and has $h_0(y)$ as eigenfunction for eigenvalue zero, as it should. But it is not negative definite, since

$$(h_n,\bar{B}^{(1)}h_n)=-8V_RN_n\left(\frac{\mu}{2\pi(1+\mu)}\right)^{\frac{1}{2}} \times\left[n-\frac{\mu}{1+\mu}n(n-1)\right], \quad (76)$$

which becomes positive for sufficiently large n. Thus the procedure of Sec. I.2 must be used.

The construction of a negative semidefinite operator from $\bar{B}^{(1)}$ as given in Sec. I.2 amounts to "completing the square" of the expression in brackets, Eq. (75), as follows: Conditions on three constants α, β, and γ are found such that if

$$q=\alpha\bar{C}^3+\beta\bar{C}^2\bar{D}+\gamma\bar{C}\bar{D}^2, \quad (77)$$

then

$$\left(\bar{C}+\frac{1}{6}\frac{\mu}{1+\mu}q\right)\left(\bar{C}+\frac{1}{6}\frac{\mu}{1+\mu}q\right)^{\dagger} \quad (78)$$

agrees with the operator in brackets in Eq. (75) to terms of order $\mu/(1+\mu)$. The conditions are found to be

$$\begin{aligned}\alpha+\gamma&=1\\ \beta&=-3.\end{aligned} \quad (79)$$

We thus have arrived at the following operator:

$$\bar{B}_1=-8V_R\left(\frac{\mu}{2\pi(1+\mu)}\right)^{\frac{1}{2}} \times\left\{\bar{C}+\frac{1}{6}\frac{\mu}{1+\mu}[\alpha\bar{C}^3-3\bar{C}^2\bar{D}+(1-\alpha)\bar{C}\bar{D}^2]\right\} \cdot\left\{\bar{D}+\frac{1}{6}\frac{\mu}{1+\mu}[\alpha\bar{D}^3-3\bar{C}\bar{D}^2+(1-\alpha)\bar{C}^2\bar{D}]\right\}. \quad (80)$$

This operator, multiplied by $[(1+\mu)/\mu]^{\frac{1}{2}}$, agrees with $\bar{B}$, multiplied by the same quantity, to terms of order $\mu/(1+\mu)$. It is Hermitian, and negative semidefinite with the transformed [according to Eq. (44)] Maxwell-Boltzmann distribution function as its stable stationary distribution. As a differential operator it is of sixth-order, and therefore is probably easier to handle in the form (80) than in strict differential-operator form. It does not appear likely that any simpler operator can furnish an equivalent approximation.

$\bar{B}_1$ is somewhat arbitrary in that the constant α is arbitrary; only the two Eqs. (79) determine the three constants α, β, and γ. However, the arbitrariness is in a term of higher order than that to which the operator is accurate; when (80) is multiplied out we must, of course, get

$$\bar{B}_1=-8V_R\left(\frac{\mu}{2\pi(1+\mu)}\right)^{\frac{1}{2}} \times\left\{\bar{C}\bar{D}+\frac{1}{6}\frac{\mu}{1+\mu}[\bar{C}^3\bar{D}-6\bar{C}^2\bar{D}^2+\bar{C}\bar{D}^3] +\frac{1}{36}\left(\frac{\mu}{1+\mu}\right)^2[\alpha\bar{C}^3-3\bar{C}^2\bar{D}+(1-\alpha)\bar{C}\bar{D}^2] \times[\alpha\bar{D}^3-3\bar{C}\bar{D}^2+(1-\alpha)\bar{C}^2\bar{D}]\right\}, \quad (81)$$

in which the term inside the braces with coefficient $\mu/(1+\mu)$, which is the first correction term, is independent of α. However, the higher-order term in braces, which has coefficient $[\mu/(1+\mu)]^2$, will not, except by coincidence, agree with the term of like order in the exact operator $\bar{B}$, no matter what the value of α, since the former is in general only part of the term in the exact operator. $\bar{B}_1$ is, however, not meant to be accurate to this order (this higher term in $\bar{B}_1$ would not be accurate even if there were no arbitrariness), and computations should not be carried beyond terms which are determined by the $\mu/(1+\mu)$ term in braces in (81). Then the arbitrariness due to the indeterminateness of α will play no part in the results.

It should not be concluded, from the fact that the $[\mu/(1+\mu)]^2$ term in braces in (81) is to be disregarded where it affects computational results, that it can be dispensed with. By ensuring negative semidefiniteness it prevents runaway solutions; without it, probability modes $h_r(y)$ of very large r value will grow indefinitely in amplitude. It is to be expected that, when used in the proper range of deviations from equilibrium, $\bar{B}_1$ will yield nonarbitrary results. The term which contains the arbitrariness must be included to prevent the intrusion, into solutions of the approximate equation, of spurious effects.

The arbitrariness due to α does not affect the possibility of constructing a sequence of approximations to $\bar{B}$. As constructed according to the prescription in Sec. I.2, $[(1+\mu)/\mu]^{\frac{1}{2}}\bar{B}_N$ will always be correct to order $[\mu/(1+\mu)]^N$. Thus in $\bar{B}_2$ the error due to α will be made good, although a new error will be introduced in a term of higher order.

The results obtained will now be considered in relation to Brinkman's assertion that the condition (23) requires

that all α_n vanish for $n>2$ (our α_n= Brinkman's μ_n). In the first place, the fact that F decreases when $P(V,t)$ obeys the linear Boltzmann equation, as proved in Sec. I.5, makes it impossible from our point of view to agree with this conclusion, since the linear Boltzmann operator in general has nonvanishing derivate moments of all orders. However, in any case, what we have sought is an approximation to an operator which in its exact form does satisfy the requirement. If this operator can be expanded in powers of $\mu/(1+\mu)$, then it is clear that successive approximations to its effect on $P(V,t)$ can be obtained by breaking the series off at successively higher terms, and that these approximations might be useful even if they did not satisfy some of the requirements the exact operator satisfies.

III. MORE GENERAL SYSTEMS

The above work can be generalized to other linear Boltzmann operators in the following two ways: by leaving the a_{nk} general, and by suppressing explicit reference to the expansion parameter $\mu/(1+\mu)$. In the following sections we take up these two modes of generalization successively.

1. Case of General a_{nk}

The a_{nk} are not mutually independent. Let us write the expansion of $\bar{B}$ in the form

$$\bar{B}=V_R\left(\frac{\mu}{1+\mu}\right)^{\frac{1}{2}}\sum_{m=0}^{\infty}\left(\frac{\mu}{1+\mu}\right)^m\bar{b}_m. \tag{82}$$

Each $\bar{b}_m$ in (82) must end in a destruction operator in order that $h_0(y)$ be a stable equilibrium solution. In the case $m=0$ we have, from Eq. (61),

$$\begin{aligned}\bar{b}_0&=[a_{11}\bar{C}y+a_{20}\bar{C}^2]\\&=[a_{11}\bar{C}(\bar{C}+\bar{D})+a_{20}\bar{C}^2].\end{aligned} \tag{83}$$

The stable equilibrium condition here requires that the coefficient of $\bar{C}^2$ vanish, or

$$a_{20}=-a_{11}. \tag{84}$$

This is, of course, just the classic relation between viscosity and diffusion coefficient discovered by Einstein.[21]

A similar relation can be found in the next order: After bringing $\bar{D}$ operators to the right in all terms, we have

$$\begin{aligned}\bar{b}_1=&(a_{13}+a_{22}+a_{31}+a_{40})\bar{C}^4+(3a_{13}+2a_{22}+a_{31})\bar{C}^3\bar{D}\\&+(3a_{13}+a_{22})\bar{C}^2\bar{D}^2+a_{13}\bar{C}\bar{D}^3.\end{aligned} \tag{85}$$

The condition that this annihilate $h_0(y)$ is that the coefficient of $\bar{C}^4$ vanish:

$$a_{13}+a_{22}+a_{31}+a_{40}=0. \tag{86}$$

In this case another restriction on the a_{nk} must be satisfied too, in order that $\bar{b}_1$ be Hermitian. Namely, the coefficients of $\bar{C}^3\bar{D}$ and of $\bar{C}\bar{D}^3$ must be equal:

$$3a_{13}+2a_{22}+a_{31}=a_{13}. \tag{87}$$

With Eqs. (84), (86), and (87) we can eliminate a_{20}, a_{31}, and a_{40}. When this is done, we obtain

$$\begin{aligned}B^{(1)}=V_R\left(\frac{\mu}{1+\mu}\right)^{\frac{1}{2}}\Big\{&a_{11}\bar{C}\bar{D}+\frac{\mu}{1+\mu}[a_{13}\bar{C}^3\bar{D}\\&+(3a_{13}+a_{22})\bar{C}^2\bar{D}^2+a_{13}\bar{C}\bar{D}^3]+o\left(\frac{\mu}{1+\mu}\right)\Big\}.\end{aligned} \tag{88}$$

In the Rayleigh process the a_{nk} are known, and in fact the relations derived above can be verified for the expressions given in Appendix B. However, Eq. (88) may also be applied in the following way. Suppose thermal fluctuations are to be studied beyond the range where linear friction applies, in some system whose linear Boltzmann operator is unknown, but in which the following hypotheses may be justifiable: (a) The (unknown) linear Boltzmann operator is expansible in terms of some parameter analogous to $\mu/(1+\mu)$, and (b) the successive derivate moments are expansible in Hermite polynomials. This amounts to saying that the random process involved is mathematically of the same type as the Rayleigh process. Equation (88) or its equivalent then tells us that in order to study the random process in a consistent way with inclusion of the V^3 term in the friction[22] (i.e., in the first derivate moment), it is sufficeint to know just the coefficient of this term and that of the V^2 term in the noise, i.e., in the second derivate moment; the remaining relevant coefficients a_{31} and a_{40}, which appear in the third and fourth derivate moments, being determined by the former two.

As in Sec. II.5, a negative semidefinite operator agreeing with $\bar{B}^{(1)}$ to order $[\mu/(1+\mu)]^{\frac{1}{2}}$ can be constructed by completing the operator absolute-square. The result is

$$\begin{aligned}\bar{B}_1=&\left(\frac{\mu}{1+\mu}\right)^{\frac{1}{2}}V_R\Big\{a_{11}^{\frac{1}{2}}\bar{C}+\frac{\mu}{1+\mu}a_{11}^{-\frac{1}{2}}\\&\times[\alpha\bar{C}^3+\tfrac{1}{2}(3a_{13}+a_{22})\bar{C}^2\bar{D}+(a_{13}-\alpha)\bar{C}\bar{D}^2]\Big\}\\&\times\Big\{a_{11}^{\frac{1}{2}}\bar{D}+\frac{\mu}{1+\mu}a_{11}^{-\frac{1}{2}}[\alpha\bar{D}^3+\tfrac{1}{2}(3a_{13}+a_{22})\bar{C}\bar{D}^2\\&+(a_{13}-\alpha)\bar{C}^2\bar{D}]\Big\}.\end{aligned} \tag{89}$$

[21] A. Einstein, Ann. Physik 17, 549 (1905).

[22] Carried out to its second term, the expansion of A_1 is

$$\begin{aligned}A_1(y)&=\frac{(1+\mu)^{\frac{1}{2}}}{\mu}\left\{a_{11}H_1+\frac{\mu}{1+\mu}a_{13}H_3+\cdots\right\}\\&=\frac{(1+\mu)^{\frac{1}{2}}}{\mu}\left\{a_{11}y+\frac{\mu}{1+\mu}a_{13}(y^3-3y)+\cdots\right\}.\end{aligned}$$

If $\mu\ll 1$, the H_3 contribution will not come in until $y^3\sim 1/\mu$, and then the $-3y$ term will be negligible compared to y^3. Similarly with all higher H_r contributions. If they contribute significantly at all, the highest power in them will dominate. Thus when $\mu\ll 1$ the Hermite expansion will not be appreciably different from a power series. And this will, of course, hold for the expansion of any derivate moment.

2. Suppression of the Expansion Parameter

Let us put

$$k_1=\left(\frac{\mu}{1+\mu}\right)^{1/4}a_{11}^{\frac{1}{2}}$$

$$k_2=\left(\frac{\mu}{1+\mu}\right)^{5/4}a_{11}^{-\frac{1}{2}}\alpha$$

$$k_3=\left(\frac{\mu}{1+\mu}\right)^{5/4}\tfrac{1}{2}a_{11}^{-\frac{1}{2}}(3a_{13}+a_{22}) \tag{90}$$

$$k_4=\left(\frac{\mu}{1+\mu}\right)^{5/4}a_{11}^{-\frac{1}{2}}(a_{13}-\alpha).$$

Then

$$\begin{aligned}\bar{B}_1&=V_R[k_1\bar{C}+k_2\bar{C}^3+k_3\bar{C}^2\bar{D}+k_4\bar{C}\bar{D}^2]\\&=[k_1\bar{D}+k_2\bar{D}^3+k_3\bar{C}\bar{D}^2+k_4\bar{C}^2\bar{D}].\end{aligned} \tag{91}$$

This form of $\bar{B}_1$ would be usable, if valid, for a system not possessing an expansion parameter, or for which this parameter was unknown. In the absence of as yet unsuspected restrictions, the four k's of Eqs. (90) are mutually independent; they certainly are so for the Rayleigh model, since α, a_{11}, a_{13}, and a_{22} are independent.

The form of the operator of Eq. (91) with arbitrary k's is sufficient for negative semidefiniteness. Let us try to define the conditions under which it is also *necessarily* the next negative semidefinite approximation after the Fokker-Planck operator. In terms of α_{nk}, the kth Hermite coefficient of $\alpha_n(y)$, Eq. (61) reads

$$\bar{B}=\sum_{m=0}^{\infty}\sum_{p=0}^{m+1}\frac{\alpha_{m+p+1,m-p+1}}{(m+p+1)!}\bar{C}^{m+p+1}H_{m-p+1}(y). \tag{92}$$

This is a perfectly general formal expression for any linear Boltzmann operator, since it may be derived without any further assumptions from Eq. (37). If the equilibrium distribution of y is Gaussian, $\exp(-y^2/2)$, sufficient conditions on $\alpha_{nk}/n!$ in order that $\bar{B}$ be Hermitian and promote stable equilibrium are the same as Eq. (84) for the $\alpha_{nk}/n!$ with $n+k=2$, and the same as Eqs. (86) and (87) for those with $n+k=4$. To prove these conditions necessary as well, a variable expansion parameter analogous to $\mu/(1+\mu)$ is needed, in order to make possible the device of setting the coefficient of each power of the parameter equal to zero. However, in the spirit of a phenomenological approach it may be justifiable to hypothesize the existence of such a parameter, when definite knowledge about a given system is not available. Assuming the hypothetical parameter to be small as well—as would be reasonable for any macroscopic variable which fluctuates due to molecular impulses or contributions—Eq. (91) would then be the most general form for the indicated approximation.

ACKNOWLEDGMENTS

This work is the outgrowth of a problem suggested to the author by Professor George E. Uhlenbeck, whose friendly and stimulating advice has moreover been invaluable throughout its course. The author wishes also to thank the John Simon Guggenheim Memorial Foundation for the grant of a fellowship for the period 1957–58, during which roughly the first half of the work was done, while on leave from Boston University.

APPENDIX A

Derivation of the Transition Probability for the Rayleigh Model

We here evaluate the function $C(y''|y')$ of Eq. (36) in the text. To do so we first work in terms of the ordinary velocity V. Let

$C(V''|V')=$ probability density-in-V'' per unit time that a particle with given velocity V' undergoes a collision which changes its velocity to V'' ("transition probability" from V' to V''). (A1)

Let v stand for the component of velocity of a molecule in the direction of the constrained motion of the particles (note that v therefore does *not* stand for the *speed* of the molecules). Given a particle with initial velocity V, and assuming the distribution (31) for the vector velocity $\mathbf{v}$, the probability density-in-v for a collision of the particle with a molecule having velocity component (in the foregoing sense) v, per unit time, is

$$A\rho f_0(v)|V-v|, \tag{A2}$$

if A is the area of the disk of the particle, ρ the spatial density of molecules, and $f_0(v)$ the one-dimensional Maxwell-Boltzmann function

$$f_0(v)=(2\pi v_R^2)^{-\frac{1}{2}}\exp\left(-\frac{v^2}{2v_R^2}\right). \tag{A3}$$

The fraction of all disks which are knocked out of the infinitesimal range dV at V by molecules in dv at v is, per unit time,

$$A\rho f_0(v)|V-v|P(V)dvdV, \tag{A4}$$

where $P(V)$ is the velocity probability density of particles.

The coefficient of $dvdV$ in (A4) is the probability density-in-v-and-V, per unit time, of the process described. However, for use in Eq. (35) we require the transition probability between two values of V, namely, V' and V'' (ultimately, y' and y''). These may be related to v and V by the dynamics of the collision, as follows: Let

$$V'=V. \tag{A5}$$

Assume that molecules are reflected specularly from the disks; then V', V'' are related to the variables V, v by Eq. (A5) and by

$$V''-V'=2\frac{\mu}{1+\mu}(v-V) \tag{A6}$$

where

$$\mu=m/M. \tag{A7}$$

To obtain a transition probability for $V'\to V''$ from the expression (A4) one must (a) write this expression in terms of V' and V''; (b) make it a probability density in V' and V'' by multiplying it by $\partial(v,V)/\partial(V',V'')=(1+\mu)/2\mu$, and (c) divide by $P(V')$ to obtain a probability *conditional* in V'. In this way one obtains

$$C(V''|V')=A\rho\left(\frac{1+\mu}{2\mu}\right)^2 f_0\left(\frac{1+\mu}{2\mu}V''-\frac{1-\mu}{2\mu}V'\right)\cdot|V''-V'|. \tag{A8}$$

An expression formally the same as this would be obtained if the gas were linear instead of three-dimensional, with the particles on a line with the gas molecules, and if every encounter between a particle and a molecule resulted in a collision. The latter is the model originally introduced by Rayleigh,[11] and the present one is mathematically equivalent to it.

Now transform to the variable y [Eq. (33)]. Writing $C(y''|y')$ for the transition probability per unit time for the event $y'\to y''$, which is a probability density in y'', we shall have

$$C(y''|y')=C(V''|V')dV''/dy'' \tag{A9}$$

(in this equation the C's stand for transition probabilities with respect to the arguments in their respective parentheses; since the arguments are different random variables on the two sides of the equation, the C's on the two sides are not meant to be the same functions). Put

$$A\rho=1, \tag{A10}$$

since this combination of constants plays no further part in the analysis. Equation (A9) applied to Eq. (A8) then gives Eq. (36) of the main text.

APPENDIX B

We here evaluate the Hermite expansion coefficients of $A_n(y)$. From Eq. (58) and the normalization constant of the $H_k(y)$,

$$A_{nk}=\frac{1}{(2\pi)^{\frac{1}{2}}k!}\int\exp\left(-\frac{y^2}{2}\right)H_k(y)A_n(y)dy. \tag{B1}$$

We utilize the generating function of the Hermite polynomials,

$$\exp\left(-\frac{z^2}{2}+zy\right)=\sum_{m=0}^{\infty}z^r\frac{H_r(y)}{r!}, \tag{B2}$$

whence A_{nk} is the coefficient of z^k in the power series expansion of

$$I(z)=(2\pi)^{-\frac{1}{2}}\int\exp\left(-\frac{y^2}{2}-\frac{z^2}{2}+zy\right)A_n(y)dy$$
$$=\frac{2^n}{2\pi n!}\int dy\exp\left(-\frac{y^2}{2}-\frac{z^2}{2}+zy\right)\times\int dx\exp\left[-\frac{\mu}{2}(x+y)^2\right]x^n|x|dx. \tag{B3}$$

Inverting the order of integration, it is possible to integrate immediately over y, using the formula

$$\int_{-\infty}^{\infty}\exp(-ay^2-by)dy=\left(\frac{\pi}{a}\right)^{\frac{1}{2}}\exp\left(\frac{b^2}{4a}\right). \tag{B4}$$

This gives

$$I(z)=\frac{2^n}{[2\pi(1+\mu)]^{\frac{1}{2}}n!}\int\exp\left(-\frac{\mu z^2}{2(1+\mu)}-\frac{\mu zx}{1+\mu}\right)\exp\left(-\frac{\mu}{2(1+\mu)}x^2\right)x^n|x|dx. \tag{B5}$$

But here we recognize, in the first exponential in the integrand, the generating function according to formula (B2) of the functions

$$H_r\left(-\left[\frac{\mu}{1+\mu}\right]^{\frac{1}{2}}x\right).$$

From this it follows that

$$A_{nk}=\frac{(-1)^k2^n}{[2\pi(1+\mu)]^{\frac{1}{2}}n!k!}\left(\frac{\mu}{1+\mu}\right)^{k/2}\times\int\exp\left(-\frac{\mu}{2(1+\mu)}x^2\right)H_k\times\left(\left[\frac{\mu}{1+\mu}\right]^{\frac{1}{2}}x\right)x^n|x|dx. \tag{B6}$$

Changing to $[\mu/(1+\mu)]^{\frac{1}{2}}x$ as variable of integration,

$$A_{nk}=\frac{(-1)^k2^n}{[2\pi(1+\mu)]^{\frac{1}{2}}n!k!}\left(\frac{\mu}{1+\mu}\right)^{(k-n-2)/2}\times\int\exp\left(-\frac{x^2}{2}\right)H_k(x)x^n|x|dx. \tag{B7}$$

The integral can be simplified as follows:

$$\int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2}\right) H_k(x) x^n |x| dx$$

$$= \int_0^{\infty} [(-1)^{n+k+2}+1] H_k(x) \exp\left(-\frac{x^2}{2}\right) x^{n+1} dx$$

$$= 2\int_0^{\infty} H_k(x) \exp\left(-\frac{x^2}{2}\right) x^{n+1} dx, \quad \text{if } n+k \text{ even}$$

$$= 0 \qquad \text{if } n+k \text{ odd.} \quad \text{(B8)}$$

The dependence on parity of $n+k$ agrees with the fact that $A_n(y)$ is even or odd according to the parity of n (cf. footnote 20).

The integral in (B8) is

$$J = \int_0^{\infty} H_k(x) \exp\left(-\frac{x^2}{2}\right) x^{n+1} dx$$

$$= \int_0^{\infty} x^{n+1} \left(-\frac{d}{dx}\right)^k \exp\left(-\frac{x^2}{2}\right) dx. \quad \text{(B9)}$$

If $k \leqq n+1$ we integrate by parts k times to get

$$J = (n+1)n \cdots (n-k+2) \int_0^{\infty} x^{n-k+1} \exp\left(-\frac{x^2}{2}\right) dx$$

$$= \frac{(n+1)!}{(n-k+1)!!} \quad (k \leqq n+1), \quad \text{(B10)}$$

where the double factorial $N!!$ is defined by

$$N!! = N(N-2)(N-4)\cdots 3\cdot 1 \quad (N \text{ odd})$$
$$= N(N-2)(N-4)\cdots 4\cdot 2 \quad (N \text{ even}), \quad \text{(B11)}$$

and, by convention, $0!! = (-1)!! = 1$.

If $k > n+1$ we integrate by parts $n+1$ times:

$$J = (n+1)! \left[-\left(-\frac{d}{dx}\right)^{k-n-2} \exp\left(-\frac{x^2}{2}\right)\right]\Bigg|_0^{\infty}$$

$$= (n+1)! H_{k-n-2}(0)$$

$$= (-1)^{(k-n-2)/2} (n+1)! (k-n-3)!!, \quad \text{(B12)}$$

the last form being obtained by adapting, to our definition of Hermite polynomials, formula (13.15), Sec. 10, of *Higher Transcendental Functions* [edited by A. Erdélyi (McGraw-Hill Book Company, Inc., New York, 1953), Vol. II].

If we combine the foregoing results, we get

$$A_{nk} = (1+\mu)^{-\frac{1}{4}} \left(\frac{\mu}{1+\mu}\right)^{(k-n-2)/2} a_{nk}, \quad \text{(B13)}$$

with

$$a_{nk} = (-1)^k \frac{2^{n+1}}{(2\pi)^{\frac{1}{2}}} \frac{n+1}{k!(n-k+1)!!} \quad (k \leqq n+1) \quad \text{(B14)}$$

$$= (-1)^{(n+k+2)/2} \frac{2^{n+1}}{(2\pi)^{\frac{1}{2}}} \frac{(n+1)(k-n-3)!!}{k!}$$
$$(k > n+1). \quad \text{(B15)}$$