to the rate $\exp(-t/\tau_d)$, is found to be

$$\tau_d \simeq T_1[26(n^2 l_n/D)+8(nl_n/D)^{1/2}]^{-1} \equiv T_1/\delta.$$

Here $T_1=2D/C$ is the period of the fundamental mode, $l_n=\eta/\rho c$ is of the order of the mean-free path, n is the mode number ($n=1$ being the fundamental), and D is the length of the side of the square tube considered. The first term in the brackets comes from the losses outside the boundary layer and the second term from the losses inside the layer.

The quantity $l_n=\eta/\rho c$ depends on both the density and the temperature of the gas. At room temperature and a pressure of approximately 1 mm Hg, we have for He, $l_n \simeq 10^{-2}$ cm. Then, if $D \simeq 1$ cm, we obtain for the first mode $n=1$, $\tau_d \simeq T_1$. Under these conditions the decay of the acoustic waves will be quite rapid, and it would be very difficult to observe acoustic resonances in a plasma at room temperature and pressure ≈ 1 mm Hg or less. This may explain why the acoustic resonances described by Strickler and Stewart were observed only at comparatively high pressures ($\simeq 13$ mm Hg), in which case we get $\tau_d \simeq 5\ T_1$.

In the afterglow experiments of Berlande, Goldan, and Goldstein, the number density of the neutrals was $N_n \simeq 7\times10^{17}$ cm^{-3}. The plasma was submerged in a bath with a temperature of 4.2°K, and at the corresponding low gas temperature in the discharge the coefficient of viscosity is estimated to be lower than the value at room temperature ($\eta \simeq 1.9\times10^{-4}$ cgs units) by more than a factor of ten (using, say, Sutherland's formula[12]). Therefore $l_n=\eta/\rho c$ will be less than 10^{-4} cm. The corresponding value of the decay time is then found to be $\tau_d \geq 10\ T_1$, which is consistent with the experimental observations.

[12] J. O. Hirschfelder, C. F. Curtiss, and R. B. Bird, *Molecular Theory of Gases and Liquids* (John Wiley & Sons, Inc., New York, 1954), p. 565.

PHYSICAL REVIEW VOLUME 145, NUMBER 1 6 MAY 1966

Nonperturbative Proof of Some Results in Classical Nonequilibrium Statistical Mechanics

RAPHAEL ARONSON
Department of Nuclear Engineering, New York University, New York, New York
(Received 3 December 1965)

Some results of interest in classical nonequilibrium statistical mechanics, previously proved to all orders in the interaction, are proved without recourse to perturbation expansions, making it possible to avoid convergence questions. One theorem so proved is that when the interaction is switched on slowly, a Maxwellian distribution goes into the canonical distribution at the same temperature. The two steps in the Prigogine theory of the approach to equilibrium that originally depended on a perturbation proof are also demonstrated nonperturbatively. Lastly, the statement that in equilibrium cluster expansions in the density there are no articulation points in the graphs contributing to the reduced distributions is proved from a time-dependent point of view.

I. INTRODUCTION

In classical-statistical mechanics, as in quantum-statistical mechanics, perturbation theory has been a convenient tool for investigations into nonequilibrium and transport problems. It has been possible in some cases to sum perturbation series in the interaction strength to all orders to obtain sensible results. However, because it is extremely difficult even for finite interactions to prove convergence, it is useful to avoid perturbation expansions when possible. A procedure which treats the time development of the classical distribution function nonperturbatively is developed here. It is used to justify certain results originally obtained by perturbation methods. The approach is formally quite similar to time-dependent scattering theory.

Two problems, the calculation of thermal averages and the approach to equilibrium, are reconsidered. It has been shown that it is convenient to compute thermal averages of many-time quantities by introducing an interaction picture.[1] The distribution function can be regarded as resulting under the mechanical motion of the system out of an initial unperturbed distribution function when the interaction is turned on slowly.[2] If the distribution function is canonical, the initial distribution function before the interaction is switched on is Maxwellian at the same temperature, and conversely. A proof not involving perturbation theory is given here as an alternative to the perturbation proof given previously.[1]

A second application is to the Prigogine treatment of the approach to equilibrium. The theory, as developed

[1] R. Aronson, J. Math. Phys. **7**, 221 (1966).
[2] The term "slowly" means the same as the more commonly used "adiabatically," which we avoid to eliminate possible confusion.

by Prigogine, Henin, Résibois, and Andrews,[3,4] starts from certain assumptions about the analytic behavior of certain quantities and shows that the equilibrium distribution is approached asymptotically. There are certain stages in the proof in which perturbation expansions are used. It is shown here that these expansions can be avoided, so that the result in no way depends on the assumed convergence of a perturbation series.

The demonstration involves the projection operators introduced by Zwanzig[5] as an alternative to a diagrammatic expansion. The relaxation of the correlations to their equilibrium values once the equilibrium velocity distribution is established turns out to be intimately related to the evolution of the distribution function under the slow switching on of the interaction.

One further application is made, a proof from a time-dependent point of view of the theorem in equilibrium theory that Mayer diagrams contributing to the reduced distributions have no articulation points.[6] It was pointed out by Andrews[7] that the theorem does in fact come out of nonequilibrium theory asymptotically, to all orders in the interaction, and the result has been proved.[1] This theorem is also proved here nonperturbatively.

The basic Liouville-operator formalism is given in Sec. II. The integral equation connecting the density distributions before and after switching is derived in Sec. III. In Sec. IV the effect of slow switching on the equilibrium distribution is determined. Section V discusses the Prigogine theory and states the theorems which are proved nonperturbatively in Sec. VI. In Sec. VII the theorem about the absence of articulation points is stated and proved.

II. BASIC FORMALISM

The Liouville equation for a system of N particles can be written[3]

$$i\frac{\partial\rho(t)}{\partial t}=L\rho(t)\,, \tag{2.1}$$

where $\rho(t)=\rho[\{\mathbf{r}(t)\},\,\{\mathbf{p}(t)\}]$ is the distribution function in phase space and L is the Liouville operator

$$L=-i\sum_{i=1}^{N}\boldsymbol{\partial}_iH\cdot\boldsymbol{\nabla}_i+i\sum_{i=1}^{N}\boldsymbol{\nabla}_iH\cdot\boldsymbol{\partial}_i\equiv L(H)\,, \tag{2.2}$$

so that

$$LF=i\{H,F\}\,. \tag{2.3}$$

Here H is the Hamiltonian of the system and $\boldsymbol{\partial}_i$ is the gradient with respect to the momentum of the ith particle. The Poisson bracket of H and F is denoted by $\{H,F\}$. L is Hermitian.

Consider H to be the sum of an unperturbed Hamiltonian H_0 and a perturbation H_1. Then the Liouville operator L can be decomposed in the same way:

$$L=L_0+L_1=L(H_0)+L(H_1)\,. \tag{2.4}$$

We can introduce an interaction picture by defining

$$\varphi(t)=e^{iL_0t}\rho(t)\,, \tag{2.5}$$

$$L_1(t)=e^{iL_0t}L_1e^{-iL_0t}. \tag{2.6}$$

Substituting (2.5) and (2.6) into (2.1), we obtain the Liouville equation in the interaction picture

$$i\frac{\partial\varphi(t)}{\partial t}=L_1(t)\varphi(t)\,. \tag{2.7}$$

Equation (2.7) can be solved to give

$$\varphi(t)=U(t,t')\varphi(t')\,, \tag{2.8}$$

where U is the unitary time-ordered operator

$$U(t,t')=\left(\exp-i\int_{t'}^{t}L_1(t'')dt''\right)_+. \tag{2.9}$$

The average of a phase function $f=f(\{\mathbf{r}\},\,\{\mathbf{p}\})$ at time t is given by

$$\langle f(t)\rangle=\int d\Gamma f\rho(t)=\int d\Gamma f_L(t)\rho\,, \tag{2.10}$$

where $d\Gamma$ is an element of phase space. Here $f=f(0)$ and $f_L(t)$ is the operator corresponding to f at time t:

$$f_L(t)=e^{iLt}fe^{-iLt}=f(\{\mathbf{r}(t)\},\,\{\mathbf{p}(t)\})\,. \tag{2.11}$$

The first form in (2.10) corresponds to a Schrödinger picture and the second form to a Heisenberg picture.

In the interaction picture, f develops according to the unperturbed motion, so that

$$f(t)=e^{iL_0t}fe^{-iL_0t}. \tag{2.12}$$

Then the average is also given by

$$\langle f(t)\rangle=\int d\Gamma f(t)\varphi(t)\,. \tag{2.13}$$

In obtaining (2.13) from (2.10) we use the fact that L is a differential operator, so that

$$\int d\Gamma L\psi=\int d\Gamma L_0\psi=0 \tag{2.14}$$

---

[3] I. Prigogine, *Non-Equilibrium Statistical Mechanics* (Interscience Publishers, Inc., New York, 1962), especially Chaps. 8, 11, and 12.

[4] F. Henin, P. Résibois, and F. C. Andrews, J. Math. Phys. **2**, 68 (1961).

[5] R. Zwanzig, in *Lectures in Theoretical Physics, Boulder*, edited by W. E. Brittin, B. W. Downs, and J. Downs (Interscience Publishers, Inc., New York, 1960), Vol. 3.

[6] G. Stell, in *The Equilibrium Theory of Classical Fluids*, edited by H. L. Frisch and J. L. Lebowitz (W. A. Benjamin, Inc., New York, 1964). Stell's definition of an articulation point is adopted here.

[7] F. C. Andrews, Physica **27**, 1054 (1961).

for any function ψ of interest. It also follows from (2.14) that

$$\int d\Gamma U(t,t')\psi = \int d\Gamma\psi. \tag{2.15}$$

From (2.7) and (2.8), or alternatively, directly from (2.9), we obtain the differential equation for $U(t,t')$:

$$i(\partial/\partial t)U(t,t') = L_1(t)U(t,t'). \tag{2.16}$$

The boundary condition, from (2.9), is $U(t,t)=1$. Equation (2.16) can be written in integral form:

$$U(t,t') = 1 - i\int_{t'}^{t} dt'' L_1(t'')U(t'',t'). \tag{2.17}$$

We note also that

$$U(t',t) = U^{-1}(t,t'). \tag{2.18}$$

III. THE INTEGRAL EQUATION

It is convenient here, as in calculations for the quantum-mechanical many-body problem which this approach parallels closely, to think of the interaction being switched on slowly from zero to its full value. It will be shown in a later section that this artificial procedure is equivalent to a more physical description of the relaxation of the correlations to equilibrium in a real system. Here it is introduced only as a computational convenience because it turns out easier to take averages over the initial preswitching distribution. The problem is to determine which initial state goes into which final state.

The formal procedure has similarities both to bound-state and to continuum situations in quantum mechanics.[8] Like the bound-state eigenvectors, the distribution functions are normalizable and we can derive a detailed parallel to the Gell-Mann and Low theorem.[9] As in a scattering problem,[8] there is a continuum of eigenvalues (of L), and in particular, turning the interaction on changes the distribution but not the eigenvalue of interest ($L_0=0$ goes to $L=0$).

To see how this goes, let the interaction H_1 be turned on slowly. That is,

$$H = H_0 + H_1 \exp(-\epsilon|t|), \tag{3.1}$$

with $\epsilon>0$, where we will eventually take the limit $\epsilon \to 0$. We will denote this limit as $0+$. Then $L_1(t)$ in Sec. II is to be replaced by $L_1(t)\exp-\epsilon|t|$ in (2.7) and (2.9). To show explicitly that we have not yet passed to the limit, we will write the new unitary operator as $U_\epsilon(t,t')$ when the distinction is necessary. The operator $U(t,t')$ is now defined by

$$U(t,t') = \lim_{\epsilon\to 0+} U_\epsilon(t,t'). \tag{3.2}$$

Inserting (2.17) into (2.8), we have for $t'=-\infty$

$$\varphi_\epsilon(t) = \varphi(-\infty) - i\int_{-\infty}^{t} L_1(t')e^{-\epsilon|t'|} \times U_\epsilon(t',-\infty)\varphi(-\infty)dt', \tag{3.3}$$

where $\varphi_\epsilon(t)$ is the distribution that develops from $\varphi(-\infty)$ with ϵ finite. Both U and U_ϵ have the usual group property

$$U_\epsilon(t,t') = U_\epsilon(t,t'')U_\epsilon(t'',t'). \tag{3.4}$$

Using (3.4) and (2.6) in (3.3) and passing to the limit, we have

$$\varphi(t) = \varphi(-\infty) - i\lim_{\epsilon\to 0+}\int_{-\infty}^{t} dt' e^{-\epsilon|t'|}e^{iL_0t'}L_1e^{-iL_0t'} \times U(t',0)U(0,-\infty)\varphi(-\infty). \tag{3.5}$$

But as can be shown quite easily,

$$e^{-iL_0t}U(t,0) = e^{-iLt}. \tag{3.6}$$

Also, from (2.8),

$$U(0,-\infty)\varphi(-\infty) = \varphi(0) \equiv \varphi. \tag{3.7}$$

Then (3.5) becomes

$$\varphi = \varphi(-\infty) - i\lim_{\epsilon\to 0+}\int_{-\infty}^{0} dt e^{i(L_0-i\epsilon)t}L_1e^{-iLt}\varphi. \tag{3.8}$$

This is our basic integral equation. We are only concerned with solutions of (3.8) which are also solutions of (3.5), i.e., which vanish when $\varphi(-\infty)=0$. For any $\varphi(-\infty)$ such a solution, if it exists, is unique.

IV. RESULT OF SLOW SWITCHING

There are two questions to be answered. First, given a stationary distribution φ, what is the unperturbed distribution from which it develops as the interaction is turned on slowly? And second, the converse. Given a stationary unperturbed distribution, what is the final distribution when the interaction is turned on up to its full strength?

A. Stationary Final Distribution

Let φ be a function of the total Hamiltonian, i.e., $\varphi=\varphi(H)$. Then

$$L\varphi = 0, \tag{4.1}$$

and from (2.7), φ is stationary. Thus,

$$L_1e^{-iLt}\varphi = (L-L_0)\varphi = -L_0\varphi, \tag{4.2}$$

[8] See, for instance, S. Schweber, *An Introduction to Quantum Field Theory* (Row, Peterson, and Company, Evanston, 1961), Chap. 11; or T. Y. Wu and T. Ohmura, *Quantum Theory of Scattering* (Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1962), Sec. 4.

[9] M. Gell-Mann and F. Low, Phys. Rev. 84, 350 (1951).

so that the integral equation (3.8) becomes

$$\begin{aligned}\varphi&=\varphi(-\infty)+i\lim_{\epsilon\to0+}\int_{-\infty}^{0}dte^{i(L_0-i\epsilon)t}L_0\varphi\\&=\varphi(-\infty)+\lim_{\epsilon\to0+}(L_0-i\epsilon)^{-1}L_0\varphi\\&=\varphi(-\infty)+(1-P_0)\varphi,\end{aligned}\tag{4.3}$$

where P_0 is the projection operator onto the subspace spanned by the zero-eigenvalue eigenfunctions of L_0. It follows that

$$\varphi(-\infty)=P_0\varphi\tag{4.4}$$

is the unperturbed distribution that goes into φ under the perturbation.

We are not able to evaluate $P_0\varphi$ in general. However, when φ is the canonical distribution function and H_0 is the kinetic energy and H_1 depends on position only, it is possible to evaluate $P_0\varphi$. The projection operator P_0 has been evaluated for this case elsewhere.[1] It is done somewhat more explicitly in Appendix A. The result is that P_0 consists of a spatial averaging operator plus an operator which projects on to a set of measure zero in phase space. The second part describes singular invariant distributions. We are concerned with nonsingular distributions φ, however, so the contribution of the singular components must be taken vanishingly small and does not contribute to any integrals or averages over φ. To put it in another way, the probability of a particle in the distribution being in any of these singular regions of phase space is vanishingly small.

The second part of P_0 will therefore be ignored in everything that follows and P_0 will be taken to be a spatial averaging operator. The results to follow will thus be strictly true only for reduced distributions. We will speak loosely, however, and regard the results as being true for the distribution functions themselves.

Take φ to be the equilibrium distribution:

$$\varphi=\frac{\exp[-\beta(H_0+H_1)]}{\int d\Gamma\exp[-\beta(H_0+H_1)]},\tag{4.5}$$

where H_0 is the total kinetic energy

$$H_0=\sum_{i=1}^{N}T_i,\tag{4.6}$$

T_i being the kinetic energy of particle i, and

$$H_1=H_1(\{\mathbf{r}\}).\tag{4.7}$$

Then, if Ω is the volume of the system,

$$P_0\varphi=\Omega^{-N}\int\{d\mathbf{r}\}\exp-\beta H_1\frac{\exp(-\beta H_0)}{\int d\Gamma\exp[-\beta(H_0+H_1)]},\tag{4.8}$$

so that

$$\varphi(-\infty)=\varphi_0\equiv\frac{\exp(-\beta H_0)}{\int d\Gamma\exp(-\beta H_0)}.\tag{4.9}$$

Thus, the distribution before the interaction is turned on is canonical at the same temperature as the final distribution.

This result is the one we need to compute thermal averages in terms of the free-particle motion only. It permits one to write[1]

$$\begin{aligned}&\langle(f_1(t_1)\cdots f_n(t_n))_+\rangle\\&\quad=\langle(f_1(t_1)\cdots f_n(t_n)U(\infty,-\infty))_+\rangle_0\\&\quad\equiv\int d\Gamma(f_1(t_1)\cdots f_n(t_n)U)\infty,-\infty))_+\varphi_0,\end{aligned}\tag{4.10}$$

where, in the last two members, the $f_i(t_i)$ are given by (2.12).

B. Stationary Initial Distribution

Let $\varphi(-\infty)$ be a function of H_0 only. Then

$$L_0\varphi(-\infty)=0.\tag{4.11}$$

It is shown in Appendix B that[10]

$$[L_0,U(t,-\infty)]+L_1(t)U(t,-\infty)=0.\tag{4.12}$$

Then

$$L\varphi=U(0,-\infty)L_0\varphi(-\infty)=0,\tag{4.13}$$

so that the necessary condition (4.4) holds and φ is stationary. Since $U(0,-\infty)\varphi(-\infty)$ defines a unique distribution, it follows under the conditions (4.6) and (4.7) that if φ_0 satisfies (4.9), then φ satisfies (4.5) and is canonical at the same temperature. This result has previously been shown by perturbation theory.[1]

It should be noted that (4.13) is merely a statement of the adiabatic theorem of Gell-Mann and Low[9] for the case in which (2.15) is satisfied. However, the key result (4.12) has here been proved without recourse to perturbation theory.

V. APPROACH TO EQUILIBRIUM

Prigogine, Henin, Résibois, and Andrews demonstrated[3,4] that under certain assumptions a classical system actually is driven to equilibrium. Under these assumptions the velocity distribution becomes asymptotically a function of the kinetic energy H_0. If molecular chaos is assumed, the function is restricted to be Maxwellian. Once the equilibrium velocity distribution is established, the correlations become asymptotically equal to those obtained from equilibrium statistical

---

[10] An alternative nonperturbative proof for $t=0$ for the quantum-mechanical case is given by Schweber, Ref. 8, p. 326.

mechanics. The demonstration of this set of results involved perturbation theory at two stages:

(1) To prove a theorem that the sum of all diagonal fragments involving n particles gives zero when applied to a function of the unperturbed Hamiltonian.

(2) To demonstrate that once the velocity distribution function reaches its equilibrium value the asymptotic correlations are those given by equilibrium theory.

We will describe the approach just outlined in somewhat more detail, showing where these two theorems fit in, and will prove them here without recourse to perturbation expansions.

The original calculations did not use the unphysical but convenient idea of switching on the interaction, but rather a resolvent method[3] which involves a causality condition which makes the integrals converge properly. We will reformulate the Liouville equation in terms of the resolvent in order to make contact with the work of Prigogine *et al.*

The interaction Hamiltonian is now assumed to be H for all times. The kinetic energy is H_0. The problem is treated as an initial value problem with the distribution $\rho(t)$ for positive times being generated from an initial distribution $\rho(0)$ and taken to vanish for $t<0$. The convenient approach is then through the Laplace transform. The typical transform relation is

$$\rho(z)=\int_0^\infty e^{izt}\rho(t)dt. \tag{5.1}$$

Here $z=is$, where s is the usual Laplace transform variable. The transform converges, if $\rho(t)$ is bounded, when $\mathrm{Im}z>0$, in which case $\rho(z)$ is analytic. The inverse of (5.1) is

$$\rho(t)=\frac{1}{2\pi}\int_{-\infty+i\gamma}^{\infty+i\gamma} dze^{-izt}\rho(z)\,, \tag{5.2}$$

where γ is real and all singularities are below the contour of integration. The transform of the solution of (2.1) is found in terms of the resolvent operator $(L-z)^{-1}$:

$$\rho(z)=-i(L-z)^{-1}\rho_0\,, \tag{5.3}$$

where $\rho_0=\rho(t=0)$.

In order to write equations which do not involve perturbation expansions directly, it is convenient to introduce the decomposition of the Liouville equation given by Zwanzig.[5]

Let P be an operator which projects onto some subspace of the phase space. The Liouville equation (2.1) can be replaced by the two equations

$$i\frac{\partial\rho_1(t)}{\partial t}=PL\rho_1(t)+PL\rho_2(t)\,, \tag{5.4}$$

$$i\frac{\partial\rho_2(t)}{\partial t}=(1-P)L\rho_2(t)+(1-P)L\rho_1(t)\,, \tag{5.5}$$

where $\rho_1=P\rho$ and $\rho_2=(1-P)\rho$. Equation (5.5) is solved for ρ_2 in terms of ρ_1 and the solution is inserted into (5.4). The resulting equations are

$$\rho_2(t)=-i\int_0^t dt'e^{-i(1-P)Lt'}\times(1-P)L\rho_1(t-t')+e^{-i(1-P)Lt}\rho_2(0)\,, \tag{5.6}$$

$$i\frac{\partial}{\partial t}\rho_1(t)=PL\rho_1(t)-i\int_0^t dt'PLe^{-i(1-P)Lt'}\times(1-P)L\rho_1(t-t')+PLe^{-i(1-P)Lt}\rho_2(0)\,. \tag{5.7}$$

In the applications to be made, P is a spatial averaging operator. As discussed previously, P and P_0 may be used interchangeably as long as we are dealing with nonsingular distributions.

The Zwanzig relations (5.6) and (5.7) become respectively in transform space

$$\rho_2(z)=-[(1-P)L-z]^{-1}(1-P)L\rho_1(z)-i[(1-P)L-z]^{-1}\rho_{20}\,, \tag{5.8}$$

$$z\rho_1(z)=PL\rho_1(z)-PL[(1-P)L-z]^{-1}(1-P)L\rho_1(z)+i\rho_{10}-iPL[(1-P)L-z]^{-1}\rho_{20}\,. \tag{5.9}$$

Here ρ_{10} and ρ_{20} are the initial values of $\rho_1(t)$ and $\rho_2(t)$, respectively.

The term $PL\rho_1$ in (5.9) contains the operator PLP, which vanishes because L is a sum of spatial gradients.

Because L is Hermitian, the operators $[L-z]^{-1}$ and $[(1-P)L-z]^{-1}$ are analytic in z off the real axis. Then $\rho(z)$ is analytic for $\mathrm{Im}z\neq0$. It defines an analytic function $\rho^+(z)$ in the upper half-plane and another one $\rho^-(z)$ in the lower half-plane and is discontinuous across the real axis. Consider $\rho^+(z)$ to be continued analytically into the lower half-plane. Since the inversion integral (5.2) is carried out over a contour for which $\mathrm{Im}z>0$, $\rho(z)$ in (5.2) may be replaced by $\rho^+(z)$. Then the contour may be closed in the lower half-plane. The integral is evaluated in terms of the singularities of $\rho^+(z)$ for $\mathrm{Im}z<0$. This may be done for ρ_1 and ρ_2 separately.[11]

Every term in (5.8) and (5.9) is analytic for $\mathrm{Im}z>0$ and can be continued analytically into lower half-plane. Henin *et al.*[12] make certain assumptions which imply that if we define

$$\rho_2'(z)=-i[(1-P)L-z]^{-1}\rho_{20}\,, \tag{5.10}$$

then all the singularities of $\rho_2'^+(z)$ are a finite distance below the real axis. The same is assumed for the operator

$$\psi(z)=PL[(1-P)L-z]^{-1}(1-P)LP. \tag{5.11}$$

[11] For any function $f(z)$ analytic off the real axis, we define $f^+(z)$ as the analytic continuation from the upper half-plane.

[12] The assumptions of Henin *et al.* are (see Prigogine, Ref. 3, Chap. 11, for definitions) that $F^+(z)$, $\psi^+(z)$, and $C_{\gamma\gamma'}{}^+(z)$ have their only singularities for $\mathrm{Im}z<-Y<0$. Assuming (5.10) *ab initio* seems just as plausible and leads to a simpler calculation.

For very large times, i.e., as $t\to\infty$, the only singularities that can contribute to (5.2) are those on the real axis. Since $\psi^+(z)$ and $\rho_2'^+(z)$ are analytic on the real axis, it is evident from (5.8) that the only relevant singularity of $\rho_1^+(z)$ and therefore of $\rho_2^+(z)$ is at the origin.

It has been shown[3] that these considerations imply that for large times the non-Markoffian equation (5.4) reduces to the Markoffian equation.

$$(\partial/\partial t)\rho_1(t)=i\mathbf{\Omega}\psi^+(0)\rho_1(t)\,, \tag{5.12}$$

where $\mathbf{\Omega}$ is an operator involving ψ^+ and its derivatives at the origin. In addition, if $f(H_0)$ is an arbitrary well-behaved function of H_0 and

$$\{\psi^+(0)\}_n f(H_0)=\lim_{\epsilon\to 0+}\{PL[(1-P)L-i\epsilon]^{-1}(1-P)LP\}_n f(H_0)=0, \tag{5.13}$$

then ρ_1 is asymptotically a function of H_0 only. The symbol $\{\ \}_n$ denotes the n-particle part of the expression inside. Under the additional assumption of molecular chaos, $\rho_1(t)$ becomes Maxwellian for long times.

While the rest of the argument about the asymptotic behavior of $\rho_1(t)$ did not involve expansions in the interaction strength, Eq. (5.13) was proved by a perturbation procedure.[3] We shall give a new nonperturbative proof in the next section.

The second part of the procedure is the demonstration that once ρ_1 is Maxwellian, the perturbation expansion of the reduced distribution functions found from $\rho_2(t)$ become asymptotically the same as those found from $\varphi-\varphi_0$ in Sec. IV. We will show in the next section that the functions themselves are asymptotically identical, not just the perturbation expansions. The theorem will be stated in the form

$$\lim_{t\to\infty}\rho_2(t)=\varphi-\varphi_0. \tag{5.14}$$

The equality is in the sense used in Sec. IV. It holds for the reduced distributions, not for the over-all distribution functions themselves.

VI. PROOF

To prove (5.13), it is sufficient to show that

$$\psi^+(0)f(H_0)=0. \tag{6.1}$$

That is because the kinetic energy H_0 is an additive function of the kinetic energies of the individual particles, so $f(H_0)=\{g(H_0)\}_n$, with the momenta of the other particles regarded as fixed parameters. The expression $\{\psi^+(0)g(H_0)\}_n$ can thus be thought of as referring to an n-particle system with all the n particles entering into the expression in brackets.

It is convenient to define an operator $V(z)$ and a function $\varphi(z)$ by

$$V(z)=P(L-z)^{-1}LP\,, \tag{6.2}$$

$$\varphi(z)=-z(L-z)^{-1}\varphi_1\,, \tag{6.3}$$

where φ_1 is a function of H_0.

In the following, repeated use is made of the identity $A^{-1}-B^{-1}=A^{-1}(B-A)B^{-1}=B^{-1}(B-A)A^{-1}$.

Decomposing L into a free-particle part L_0 and an interaction part L_1, and recognizing that $L_0\varphi_1=0$, one finds that

$$\varphi(z)=\varphi_1-(L_0-z)^{-1}L_1\varphi(z)\,. \tag{6.4}$$

From (6.3) and (6.4),

$$(L-z)^{-1}L\varphi_1=(L_0-z)^{-1}L_1\varphi(z)\,. \tag{6.5}$$

It is also easy to show that

$$\psi(z)=zV(z)+\psi(z)V(z)\,. \tag{6.6}$$

Using (6.2), (6.5), and (6.6), one has

$$\psi(z)\varphi_1=[z+\psi(z)]V(z)\varphi_1=[z+\psi(z)]P(L_0-z)^{-1}L_1\varphi(z)\,. \tag{6.7}$$

Now take $z=i\epsilon$ and let $\epsilon\to 0+$. From (6.3), $\varphi i(\epsilon)$ approaches a limit $\varphi^+(0)$ as $\epsilon\to 0+$, and

$$L\varphi^+(0)=0. \tag{6.8}$$

In the limit, Eq. (6.4) becomes identical to (3.8) with $\varphi_1=\varphi(-\infty)$ and we can make the identification

$$\varphi^+(0)=\varphi. \tag{6.9}$$

Then

$$V^+(0)\varphi_1=P\lim_{\epsilon\to 0+}(L_0-i\epsilon)^{-1}L_1\varphi=-P(1-P)\varphi=0. \tag{6.10}$$

Substituting (6.10) into (6.7) in the limit gives the desired result (6.1), and therefore (5.13).

This proof is a restatement of a well-known result concerning the diagrammatic representation of the perturbation expansion. It says that unlinked parts do not contribute, i.e., that diagrams with unlinked parts vanish. The operator corresponding to an unlinked part is just $\psi^+(0)$.

The second result we wish to prove nonperturbatively is that once the equilibrium velocity distribution is established the correlations relax to their equilibrium values. For that purpose we now restrict φ_1 to be the equilibrium velocity distribution φ_0. If we replace $\rho_1(t)$ by φ_0, or in terms of the transform, put

$$\rho_1(z)=iz^{-1}\varphi_0 \tag{6.11}$$

in (5.8), then it will be shown that

$$\lim_{t\to\infty}\rho_2(t)=\varphi-\varphi_0. \tag{6.12}$$

The distribution φ calculated from (6.4) and (6.9) is now the equilibrium distribution. By virtue of (6.11), Eq. (5.8) becomes

$$\rho_2(z)=-[(1-P)L-z]^{-1}(1-P)Liz^{-1}\varphi_0+\rho_2'(z)\,, \quad (6.13)$$

where $\rho_2'(z)$ is given by (5.10).

The same argument that gives (6.6) gives

$$[(1-P)L-z]^{-1}(1-P)L\varphi_0 = (L-z)^{-1}L\varphi_0+[(1-P)L-z]^{-1}zV(z)\varphi_0. \quad (6.14)$$

Substituting (6.14) into (6.13) and using (6.4) and (6.5), we have

$$z\rho_2(z)=i[\varphi(z)-\varphi_0]-i[(1-P)L-z]^{-1} \times zV(z)\varphi_0+z\rho_2'(z)\,. \quad (6.15)$$

Now in the limit $z=i\epsilon$, $\epsilon\rightarrow 0+$, the last term in (6.15) vanishes, according to the discussion following (5.11). The preceding term also vanishes, by (6.10). The inverse transform (5.2) therefore gives (6.12).

This result indicates the practical equivalence of the slow switching procedure with the adoption of a causality condition.

VII. ARTICULATION POINTS

Consider a term in an equilibrium cluster expansion in the density in which two clusters have one and only one particle in common. Suppose that the coordinates of the particles in at least one of these clusters, except possibly for the linking particle, do not appear in the final result of interest. Then such a term does not contribute to the result. In the corresponding Mayer diagram, the point representing the linking particle is known as an articulation point. Andrews[7] pointed out that the same result comes out of nonequilibrium theory. It is formulated in his case as the vanishing of certain terms in the expansion in the interaction. The result has been proved in terms of a switching on procedure using perturbation theory.[1]

The result in all cases depends on the assumption, which we shall make and which is not terribly restrictive, that the interaction between a pair of particles is invariant under interchange of their coordinates.

Let cluster A be the cluster whose coordinates do not appear in the final result, minus the linking particle. Let P_C be the operator which averages over the positions of the particles in cluster A. Let $L_{\neq}$ be the part of L which does not refer to any of the particles in cluster A.

The correlations between the particles are given by

$$\varphi-\varphi_0=-\lim_{\epsilon\rightarrow 0+}(L-i\epsilon)^{-1}L\varphi_0\,, \quad (7.1)$$

according to (6.4), (6.5), and (6.9), where φ_0 is the Maxwellian and φ the distribution function in equilibrium. The theorem can then be stated as

$$\lim_{\epsilon\rightarrow 0+}P_C(L-i\epsilon)^{-1}L\varphi_0=\lim_{\epsilon\rightarrow 0+}(L_{\neq}-i\epsilon)^{-1}L_{\neq}\varphi_0. \quad (7.2)$$

That is, the asymptotic correlations between the remaining particles in no way involve the particles in cluster A.

To prove the theorem, note that

$$\begin{aligned} P_C(L-z)^{-1}L\varphi_0 &= (L_{\neq}-z)^{-1}P_C[1-\mathfrak{L}(L-z)^{-1}]L\varphi_0 \\ &= (L_{\neq}-z)^{-1}L_{\neq}\varphi_0+(L_{\neq}-z)^{-1} \\ &\quad \times P_C\mathfrak{L}[1-(L-z)^{-1}L]\varphi_0, \end{aligned} \quad (7.3)$$

where

$$\mathfrak{L}=L-L_{\neq}. \quad (7.4)$$

In the limit, the last term of (7.3) becomes, by virtue of (7.1),

$$\lim_{\epsilon\rightarrow 0+}(L_{\neq}-i\epsilon)^{-1}P_C\mathfrak{L}[1-(L-i\epsilon)^{-1}L]\varphi_0 = \lim_{\epsilon\rightarrow 0+}(L_{\neq}-i\epsilon)^{-1}P_C\mathfrak{L}\varphi. \quad (7.5)$$

Now consider an inversion of the coordinates of the particles of cluster A in the position of the other particle with which they interact. By hypothesis the interactions are invariant under this transformation. Since $\mathfrak{L}$ is a homogeneous linear function of the gradients with respect to the positions of the particles of cluster A, it changes sign under the inversion. On the other hand $L_{\neq}$ is invariant, as is the projection operator P_C. The distribution φ is canonical and therefore depends on the coordinates only through the interactions, so it is also invariant under the inversion. Thus, by symmetry,

$$P_C\mathfrak{L}\varphi=0. \quad (7.6)$$

Taking the limit of (7.3) and using (7.5) and (7.6) gives the desired result (7.2).

The proof evidently holds also when the two groups of particles do not interact with each other at all, i.e., when the clusters are unlinked. When cluster A consists of all the particles in the system, $L_{\neq}=0$ and $P_C=P$, so (7.2) becomes equivalent to (6.10).

VIII. DISCUSSION

It was shown how perturbation theory may be avoided in the proof of certain results pertaining to development of the classical distribution function in time. That means that these results hold for singular as well as nonsingular interactions and one can completely avoid awkward convergence questions.

In Sec. IV it was shown that under the slow switching the unique distribution which develops in time into the equilibrium distribution φ is the Maxwellian φ_0 at the same temperature. At first sight a unique initial distribution seems incompatible with the relaxation of an arbitrary initial distribution into the equilibrium distribution. The resolution of the paradox lies in the fact that (3.7) is exact and describes the mechanical motion only, with no statistical averaging involved. The averaging must then appear in the initial condition. An initial distribution which develops into the equilibrium distribution according to (3.7) perforce obeys the molecular chaos condition and cannot be arbitrary. It is, in fact, equal to the equilibrium velocity distribution. The relation between the two is the burden of much of Sec. VI. The slow switching procedure reproduces mathematically the actual relaxation of an initial Maxwellian to the equilibrium distribution function.

It is tempting to give the results of Sec. IV a thermodynamic interpretation. One would say that the initial canonical (Maxwellian) distribution in the absence of the interaction remains canonical as the interaction is switched on slowly and that an isothermal equilibrium process is occurring. If this interpretation is correct, the same result must occur for the density matrix in a quantum system.

When neither H_0 nor H has bound states, it follows from the quantum analog of (4.12) that the unperturbed density matrix $\rho_0=\exp(-\beta H_0)/\mathrm{Tr}\exp(-\beta H_0)$ goes into the density matrix $\rho=\exp(-\beta H)/\mathrm{Tr}\times\exp(-\beta H)$ when the perturbation is turned on adiabatically. In particular, the Maxwellian density matrix goes into the canonical density matrix for the interacting system. When there are bound states, the theorem does not hold exactly, but only in a correspondence limit.

Since the adiabatic limiting process causes certain complications which are not present classically, and since $\exp(-\beta H)$ appears in quantum-mechanical averages along with the time-development operator $\exp iHt$, it is convenient to treat β as an imaginary time in quantum-mechanical calculations. It is remarkable that in the classical case, where the time development operator is $\exp iLt$ so that β cannot be treated as an imaginary time, it is possible to regard the distribution function as developing in real time from an unperturbed distribution. As a result the switching procedure is useful classically.

There has been no attempt here to weaken any of the assumptions entering into the theory of Prigogine, Henin, Résibois, and Andrews other than the assumption that the perturbation series converges. However, the present nonperturbative proof of those of their results that originally involved expansions in the interaction puts on solid ground what were perhaps the weakest links in their chain of argument.

ACKNOWLEDGMENTS

The author wishes to acknowledge the hospitality of the Aspen Institute for Humanistic Studies, where most of this work was done. He also wishes to thank Dr. Jerome Percus for his comments on the manuscript, and Dr. George Stranahan for a helpful discussion.

APPENDIX A: PROJECTION OPERATOR P_0

For a system of N identical free particles of mass m, the Hamiltonian is

$$H_0=\sum_{i=1}^{N}\mathbf{p}_i^2/2m. \tag{A1}$$

The corresponding free-particle Liouville operator is

$$L_0=-im^{-1}\sum_{i=1}^{N}\mathbf{p}_i\cdot\nabla_i. \tag{A2}$$

It is easy to see that the functions

$$\psi(\{\mathbf{r}\},\{\mathbf{p}\};\{\mathbf{k}\},\{\boldsymbol{\pi}\})=\Omega^{-N/2}\prod_{i=1}^{N}e^{i\mathbf{k}_i\cdot\mathbf{r}_i}\delta(\mathbf{p}_i-\boldsymbol{\pi}_i)\,, \tag{A3}$$

where Ω is the volume, form a complete orthonormal set of eigenfunctions of L_0 with eigenvalues

$$\lambda(\{\mathbf{k}\},\{\boldsymbol{\pi}\})=m^{-1}\sum_{i=1}^{N}\boldsymbol{\pi}_i\cdot\mathbf{k}_i. \tag{A4}$$

Then $f(\{\mathbf{r}\},\{\mathbf{p}\})$ has the expansion

$$f(\{\mathbf{r}\},\{\mathbf{p}\})=\sum_{\{\mathbf{k}\}}\int\{d\boldsymbol{\pi}\}\tilde{f}(\{\mathbf{k}\},\{\boldsymbol{\pi}\})\times\psi(\{\mathbf{r}\},\{\mathbf{p}\};\{\mathbf{k}\},\{\boldsymbol{\pi}\})\,. \tag{A5}$$

The eigenvalue $\lambda(\{\mathbf{k}\},\{\boldsymbol{\pi}\})$ vanishes (a) for all $\{\boldsymbol{\pi}\}$, when $\{\mathbf{k}\}=0$; and (b), when $\{\mathbf{k}\}\neq\mathbf{0}$, on the $(3N-1)$-dimensional hyperplane in $\boldsymbol{\pi}$ space given by $\sum_i\boldsymbol{\pi}_i\cdot\mathbf{k}_i=0$.

The projection on the hyperplane can be described by a projection operator

$$\begin{aligned}\Delta(\mathbf{k},\mathbf{p})&=1\quad\text{when}\quad\textstyle\sum_i\mathbf{k}_i\cdot\mathbf{p}_i=0\\&=0\quad\text{when}\quad\textstyle\sum_i\mathbf{k}_i\cdot\mathbf{p}_i\neq0.\end{aligned} \tag{A6}$$

Then

$$P_0f(\{\mathbf{r}\},\{\mathbf{p}\})=\sum_{\{\mathbf{k}\}\neq0}\Delta(\mathbf{k},\mathbf{p})\Omega^{-N/2}\times\prod_{i=1}^{N}e^{i\mathbf{k}_i\cdot\mathbf{r}_i}\tilde{f}(\{\mathbf{k}\},\{\mathbf{p}\})+\int\{d\boldsymbol{\pi}\}\tilde{f}(0,\{\boldsymbol{\pi}\})\psi(\{\mathbf{r}\},\{\mathbf{p}\};0,\{\boldsymbol{\pi}\})\,. \tag{A7}$$

It is assumed that we are dealing with distributions which are nonsingular functions of the velocities and that the wave number transforms of all the quantities of interest exist. Then $\Omega^{-N/2}(\prod_{i=1}^{N}\exp i\mathbf{k}_i\cdot\mathbf{r}_i)\tilde{f}(\{\mathbf{k}\},\{\mathbf{p}\})$ is finite. On the other hand, $\sum_{\{\mathbf{k}\}\neq0}\Delta(\mathbf{k},\mathbf{p})$ projects on to a set of measure zero in phase space. Thus, the first term on the right in (A7) cannot contribute to any integrals. It is ignored here, with the proviso that the results obtained really hold only for the reduced distributions.

Physically, this term describes an invariant distribution in which the motion of the system is in a sense perpendicular to the spatial inhomogeneities. But for a continuous distribution the probability of the particles being so distributed is vanishingly small.

The last term on the right is just the spatial average of $f(\{\mathbf{r}\},\{\mathbf{p}\})$, so

$$P_0f(\{\mathbf{r}\},\{\mathbf{p}\})=\Omega^{-N}\int\{d\mathbf{r}\}f(\{\mathbf{r}\},\{\mathbf{p}\})\,. \tag{A8}$$

In the limit as the volume becomes infinite, $\sum_{\{\mathbf{k}\}\neq 0}\Delta(\mathbf{k},\mathbf{p})$ is replaced by $\int\{d\mathbf{k}\}\Delta(\mathbf{k},\mathbf{p})$. The transform $\tilde{f}(\{\mathbf{k}\},\{\mathbf{p}\})$ is assumed to exist in the limit $\{\mathbf{k}\}\rightarrow 0$. Then the integrand is finite everywhere and vanishes except on a set of measure zero, so the integral vanishes.

APPENDIX B: PROOF THAT $L_1(t)U(t,-\infty)+[L_0,U(t,-\infty)]=0$

We give here a nonperturbative proof of the analog of the adiabatic theorem of Gell-Mann and Low.

The differential equation for $U_\epsilon(t,-\infty)$ is

$$i\frac{\partial}{\partial t}U_\epsilon(t,-\infty)=e^{\epsilon t}L_1(t)U_\epsilon(t,-\infty)\quad\text{for}\quad t<\tau<\infty. \tag{B1}$$

Taking the commutator with L_0, we have

$$i\frac{\partial}{\partial t}[L_0,U_\epsilon(t,-\infty)]=e^{\epsilon t}[L_0,L_1(t)]U_\epsilon(t,-\infty)+e^{\epsilon t}L_1(t)[L_0,U_\epsilon(t,-\infty)]. \tag{B2}$$

But

$$\begin{aligned}e^{\epsilon t}[L_0,L_1(t)]U_\epsilon(t,-\infty)&=-ie^{\epsilon t}\left[\frac{\partial}{\partial t}L_1(t)\right]U_\epsilon(t,-\infty)\\&=-i\frac{\partial}{\partial t}[e^{\epsilon t}L_1(t)U_\epsilon(t,-\infty)]+ie^{\epsilon t}L_1(t)\frac{\partial}{\partial t}U_\epsilon(t,-\infty)+i\epsilon e^{\epsilon t}L_1(t)U_\epsilon(t,-\infty).\end{aligned} \tag{B3}$$

Define

$$f_\epsilon(t)=[L_0,U_\epsilon(t,-\infty)]+e^{\epsilon t}L_1(t)U_\epsilon(t,-\infty). \tag{B4}$$

Then $f_\epsilon(-\infty)=0$. By (B1), $f_\epsilon(t)$ can also be written

$$f_\epsilon(t)=[L_0,U_\epsilon(t,-\infty)]+i\frac{\partial}{\partial t}U_\epsilon(t,-\infty). \tag{B5}$$

Substituting (B3) into (B2) and using (B4) and (B5), we get

$$i\frac{\partial}{\partial t}f_\epsilon(t)=e^{\epsilon t}L_1(t)f_\epsilon(t)+i\epsilon e^{\epsilon t}L_1(t)U_\epsilon(t,-\infty), \tag{B6}$$

with the solution

$$\begin{aligned}f_\epsilon(t)&=\epsilon\int_{-\infty}^{t}dt'e^{\epsilon t'}U_\epsilon(t,t')L_1(t')U_\epsilon(t',-\infty)\\&=\epsilon\int_{-\infty}^{t}dt'e^{\epsilon t'}(L_1(t')U_\epsilon(t,-\infty))_+.\end{aligned} \tag{B7}$$

Equation (B7) can be written somewhat differently. We introduce an interaction-strength parameter λ; so instead of L_1 we write λL_1. Since then

$$U_\epsilon(t,-\infty)=\left(\exp-i\lambda\int_{-\infty}^{t}L_1(t')e^{\epsilon t'}dt'\right)_+, \tag{B8}$$

$$\frac{\partial}{\partial\lambda}U_\epsilon(t,-\infty)=-i\int_{-\infty}^{t}dt'e^{\epsilon t'}(L_1(t')U_\epsilon(t,-\infty))_+, \tag{B9}$$

so that (B7) becomes

$$[L_0,U_\epsilon(t,-\infty)]+e^{\epsilon t}L_1(t)U_\epsilon(t,-\infty)=i\epsilon\lambda\frac{\partial}{\partial\lambda}U_\epsilon(t,-\infty). \tag{B10}$$

This is precisely the form of Gell-Mann and Low's results. In our case, however, $(\partial/\partial\lambda)U(t,-\infty)$ is finite and well determined for all ϵ. Passing to the limit $\epsilon\rightarrow 0+$, we obtain the desired result

$$[L_0,U(t,-\infty)]+L_1(t)U(t,-\infty)=0. \tag{B11}$$

This proof avoids the assumption that

$$\frac{\partial}{\partial t}U(t,-\infty)=i[L_0,U(t,-\infty)], \tag{B12}$$

whose direct proof in turn involves perturbation theory.