Journal of Statistical Physics, Vol. 5, No. 1/2, 1972

Nonequilibrium Statistical Mechanics of Finite Classical Systems—I

C. J. Myerscough[1]

Received June 22, 1971

It is pointed out that the fine-grained probability density of statistical mechanics is of interest only through coarse-grained densities—integrals over nonzero volumes of phase space. This suggests the definition of a smoothed probability density: the unsmoothed density convoluted with a kernel having a small "spread" around zero velocity. If this kernel is of Gaussian form, the smoothed density satisfies a closed and exact equation for its evolution differing from the Liouville equation by the addition of one term. This equation is applied to the simple example of a noninteracting system. We need make no assumption about the size of the system in our discussion, though if the system is large enough, the assumption that it is infinite gives the same results. Reduced distribution functions are then discussed, and a treatment of the Landau damping of electron plasma oscillations is given that is free from the usual difficulties occasioned by the breakdown of the linearization.

KEY WORDS: Coarse-grained probability density; smoothed probability density; avoidance of infinite system limit; weak-interaction master equation; diagram techniques; nonequilibrium statistical mechanics; kinetic theory.

1. INTRODUCTION

There have been two distinct phases in kinetic theory and the statistical development of mechanics since the pioneering work of Boltzmann[1] nearly a century ago. Until about thirty years ago, this was the only work that could be used in practical applications,[2] but much related theoretical

[1] Central Electricity Research Laboratories, Leatherhead, Surrey, England. Formerly at Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, England.

work was being done[3] with the aim of underpinning the foundations of the subject. Since then, there has been a vast expansion in both of these fields, but the gap between them has widened. In kinetic theory, many *master equations* for the evolution of classical systems have been derived, and many new phenomena, particularly in the field of plasma physics, explained.[4–8] The fundamental problems have, however, been considered more and more by pure mathematicians, who have developed the subject of ergodic theory.[9]

One cause of this gap is that while most results of ergodic theory have concerned systems of finite size (though possibly very large), all recent derivations of master equations have asserted the necessity of making the number of particles in the system actually infinite. Often this has been deemed necessary in order to make the Poincaré recurrence time for the system infinite. However, details of the mathematical difficulties inherent in proceeding from a phase space of finite size and dimensionality to one of infinite size and dimensionality are usually left rather vague.

In this paper and its successor, we aim to show how various well-known problems of kinetic theory and statistical mechanics may be tackled without invoking this limit. We first set up the basic equations and boundary conditions with which we shall be concerned in a study of finite systems; we then briefly discuss the meaning of the phrase "a system tends to equilibrium," stressing that the basic concept involved is that of the "coarse-grained probability density" first introduced by Gibbs. We set out what we should aim to prove, and show that this is not contradicted by the existence of a finite Poincaré recurrence time for the system.

Next appears the idea of a smoothed probability density. This is much easier to handle than the coarse-grained density, for it satisfies a closed and exact equation for its evolution similar to the Liouville equation for the unsmoothed density. This equation can be used instead of the Liouville equation to discuss almost any problem of classical statistical mechanics. We begin with the simple one of a noninteracting system. The results are very similar to those derived using the infinite limit, provided the time the system takes to reach equilibrium is much less than the time a particular particle of the system takes to cross its containing vessel.

We then discuss smoothed reduced probability distribution functions, setting up an analog of the BBGKY hierarchy. Finally, in this paper, we show how Landau's famous solution[4] of the Vlasov equation for electron plasma oscillations may be set out in terms of smoothed densities. The long-term behavior of the electric field is the usual exponential decay; however, we find that the linearization of Vlasov's equation used to derive this need not break down eventually if the initial perturbation is small enough.

In the next paper, we shall apply our ideas to the work of the Brussels school, studying particularly the weak-interaction problem, and obtaining

an analog of the Brout–Prigogine master equation, again avoiding the use of the infinite limit. We discuss the conditions of validity of this equation, and suggest directions for future work. Earlier versions of some of the results of these papers were given at the 1969 Cornell Symposium on Kinetic Equations, and appear in the proceedings.[21]

2. FINE-GRAINED AND COARSE-GRAINED PROBABILITY DENSITIES

We consider a classical dynamical system of N identical particles of unit mass. Each particle has s translational degrees of freedom. Let

$$\mathbf{x}_r = (x_{r1}, ..., x_{rs}), \qquad \mathbf{v}_r = (v_{r1}, ..., v_{rs}) \tag{1}$$

be respectively the position and the velocity of the rth particle at time t. The set of pairs $(\mathbf{x}_r, \mathbf{v}_r)$ for any r is the $2s$-dimensional μ-space of the system.

The configuration of the system at time t is specified by the two Ns-vectors

$$\mathbf{X} = \{\mathbf{x}_1, ..., \mathbf{x}_N\}, \qquad \mathbf{V} = \{\mathbf{v}_1, ..., \mathbf{v}_N\} \tag{2}$$

i.e., by a point in the N-dimensional space Γ of pairs $(\mathbf{X}, \mathbf{V})$ (the Γ-space). Throughout this work, we shall use a capital letter for a Ns-vector, with the corresponding small letter for its Ns-vector components, as in (2). We write

$$\mathbf{A} \cdot \mathbf{B} = \mathbf{a}_1 \cdot \mathbf{b}_1 + \cdots + \mathbf{a}_N \cdot \mathbf{b}_N \tag{3}$$

$$\partial\phi/\partial\mathbf{X} = \{\partial\phi/\partial\mathbf{x}_1, ..., \partial\phi/\partial\mathbf{x}_N\} \tag{4}$$

In statistical mechanics, we are never given the exact positions and velocities of particles of the system. Instead, we assume that there exists a real and positive-valued function $\rho(\mathbf{X}, \mathbf{V}, t)$ such that the probability of finding the system (or rather, the point representing the configuration of the system) in a given subset Ω of Γ at time t is

$$P(\Omega, t) = \int_{\Omega} \rho(\mathbf{X}, \mathbf{V}, t)\, d^{Ns}X\, d^{Ns}V \tag{5}$$

[Hence $P(\Gamma, t) = 1$.]

The only physical significance of ρ is through expressions of the form (5). No experiment can measure ρ itself; all experiments we can do give estimates of some $P(\Omega, t)$. Suppuse, for example, that we have a number G of independent copies of the system (a Gibbs ensemble). Then, the number $n(\Omega, t)$ of these that are in Ω at time t is a random variable whose mean is $GP(\Omega, t)$. This if n is measured, n/G gives an estimate of P. The proportional error in

this estimate will decrease with n; indeed, it is well known from probability theory to behave like $n^{-1/2}$ when n is large. For a given accuracy of measurement of $P(\Omega, t)$, we can stipulate a lower bound to GP. Assuming that ρ is continuous, for small enough sets Ω, $P(\Omega, t)$ is proportional to $\mu(\Omega)$, the $2Ns$-dimensional measure of Ω. For a given accuracy of measurement of P, there is a lower bound to $G\mu(\Omega)$. Since there is always an upper limit to the size of ensemble that we can use in out experiments, to fix any standard of accuracy in measurement at once establishes a nonzero lower bound to the measure of sets Ω such that $P(\Omega, t)$ is physically meaningful.

We therefore consider the behavior of the "coarse-grained probability density"

$$\rho(\Omega, t) = [1/\mu(\Omega)]\, P(\Omega, t) \tag{6}$$

for $\mu(\Omega) \neq 0$, rather than that of the "(fine-grained) probability density" ρ itself.

If the system moves according to Hamilton's equations

$$d\mathbf{X}/dt = \partial\mathscr{H}/\partial\mathbf{V}, \qquad \partial\mathbf{V}/\partial t = -\partial\mathscr{H}/\partial\mathbf{X} \tag{7}$$

where $\mathscr{H}(\mathbf{X}, \mathbf{V}) = H(\mathbf{X}) + V^2/2$ is the total energy, and ρ varies with time according to Liouville's equation

$$0 = \mathscr{L}\rho \equiv (\partial\rho/\partial t) + \mathbf{V} \cdot (\partial\rho/\partial\mathbf{X}) + \mathbf{F}(\mathbf{X}) \cdot (\partial\rho/\partial\mathbf{V}) \tag{8}$$

where $\mathbf{F}(\mathbf{X}) = -\partial H/\partial\mathbf{X}$ gives the force on each particle.

3. BOUNDARY CONDITIONS

The system is confined within a cubic vessel with perfectly reflecting sides of length l; Γ consists of all points $(\mathbf{X}, \mathbf{V})$ satisfying

$$0 < x_{ij} < l, \qquad i = 1,..., N; \quad j = 1, 2, 3,..., s \tag{9}$$

and μ consists of the set of such points for any given i.

When the motion of the system causes one of its x coordinates to reach an extreme value, the corresponding velocity is at once reversed in sign, and the motion continues. This corresponds to the following condition on ρ at the edges of Γ:

$$\rho(\mathbf{X}, \mathbf{V}, t) = \rho(\mathbf{X}, \mathbf{V}', t) \tag{10}$$

where, for all r, j such that $x_{rj} = 0$ or l,

$$v'_{rj} = -v_{rj} \tag{11}$$

To incorporate such explicit boundary conditions we introduce periodic boundary conditions. Instead of representing the system by one point $(\mathbf{X}, \mathbf{V})$, represent it by the infinite set of points $(\mathbf{X}'', \mathbf{V}'')$, where

$$\begin{aligned} x''_{ij} &= 2n_{ij}l \pm x_{ij} \\ v''_{ij} &= \pm v_{ij} \end{aligned} \tag{12}$$

for any set of integers n_{ij} $(i = 1,..., N;\ j = 1,..., s)$. (Some of these points may coincide.) Extend the definition of $\mathscr{H}(\mathbf{X}, \mathbf{V})$ to all space by writing

$$\mathscr{H}(\mathbf{X}'', \mathbf{V}'')[=\mathscr{H}(\mathbf{X}'', \mathbf{V})] = \mathscr{H}(\mathbf{X}, \mathbf{V}) \tag{13}$$

for all such $(\mathbf{X}'', \mathbf{V}'')$. $\mathscr{H}$ will be continuous at the boundary of Γ, and at the boundaries of all the other "cells" obtained from Γ by changing any of the x_{ij} by any multiple of l. $\mathscr{H}$ is even and has period $2l$ in any of the x_{ij}.

Let all the points $(\mathbf{X}'', \mathbf{V}'')$ defined by (12) move according to (7) with the $\mathscr{H}$ of (13), and subject to no other constraints. The motion of a point moving within Γ and kept inside Γ by perfectly reflecting walls is exactly the same as the motion of the point $(\mathbf{X}'', \mathbf{V}'')$ which lies inside Γ. As soon as this point reaches the edge of Γ, it leaves Γ, but simultaneously and at the same position in space, another point enters Γ with all velocity components the same except for those normal to the boundary, which are reversed in sign. The first point will suffer no discontinuities in its position or velocity as it leaves Γ.

It is clear that the probability of finding one of the points representing the system inside a volume Ω of $(\mathbf{X}, \mathbf{V})$ space wholly contained within one cell is given by (5) if we stipulate that for all (X, V) in the interior of Γ and all $(\mathbf{X}'', \mathbf{V}'')$ defined by (12),

$$\rho(\mathbf{X}'', \mathbf{V}'', t) = \rho(\mathbf{X}, \mathbf{V}, t) \tag{14}$$

This ρ is hence a probability density; it is even and periodic in $\mathbf{X}$, and must satisfy (8) at all interior points of cells. The final remark of the last paragraph shows that ρ must be continuous across cell boundaries, so we may define it on these boundaries also. Furthermore, if $(\mathbf{X}, \mathbf{V})$ is a boundary point, $(\mathbf{X}, \mathbf{V}')$ given by (11) is an $(\mathbf{X}'', \mathbf{V}'')$ given by (12). So (14) implies (10). Within Γ, $\rho(\mathbf{X}, \mathbf{V}, t)$ satisfies the Liouville equation (7) and the boundary condition (10).

Thus we need consider only certain periodic solutions of (7), with periodic Hamiltonian given by (13). We call this imposing periodic (boundary) conditions. This idea has been introduced by many authors.[11,12]

It is physically clear from our discussion, and easy to prove from (7),

that if $f(\mathbf{X}, \mathbf{V})$ is any function also satisfying the periodic conditions (14), then

$$\int_\Gamma f(\mathbf{X}, \mathbf{V})\, \rho(\mathbf{X}, \mathbf{V}, t)\, d^{Ns}X\, d^{Ns}V \tag{15}$$

is constant.

4. FOURIER SERIES REPRESENTATION

We impose periodic conditions by expanding ρ and H in Fourier series of period $2l$ in each component of $\mathbf{X}$

$$\rho(\mathbf{X}, \mathbf{V}, t) = \sum_{\mathbf{K}} \rho_{\mathbf{K}}(\mathbf{V}, t) \exp i\mathbf{K} \cdot \mathbf{X} \tag{16}$$

$$H(\mathbf{X}) = \sum_{\mathbf{K}} H_{\mathbf{K}} \exp i\mathbf{K} \cdot \mathbf{X} \tag{17}$$

where $\mathbf{K}$ ranges over the Ns-dimensional lattice of vectors given by $K_{rj} = (\pi/l)\, n_{rj}$. Because $\rho(\mathbf{X}, \mathbf{V}, t)$ and $H(\mathbf{X})$ are even in all components of X, all the $\rho_{\mathbf{K}}$ and $H_{\mathbf{K}}$ are real, and

$$\rho_{\mathbf{K}'}(\mathbf{V}, t) = \rho_{\mathbf{K}}(\mathbf{V}, t); \qquad H_{\mathbf{K}} = H_{\mathbf{K}'} \tag{18}$$

where $\mathbf{K}'$ is obtained by changing the sign of any component of $\mathbf{K}$.

Equation (8) gives

$$0 = [\partial \rho_{\mathbf{K}}(\mathbf{V}, t)/\partial t] + i\mathbf{K} \cdot \mathbf{V}\rho_{\mathbf{K}} + \sum_{\mathbf{K}'} \mathbf{F}_{-\mathbf{K}'}[\partial \rho_{\mathbf{K}+\mathbf{K}'}(\mathbf{V}, t)/\partial \mathbf{V}] \tag{19}$$

where $\mathbf{F}_{\mathbf{K}} = -i\mathbf{K}H_{\mathbf{K}}$. We assume that the form of H corresponds to two-body central interactions, with no external field. For $\mathbf{X} \in \Gamma$,

$$H(\mathbf{X}) = \sum_{p=2}^{N} \sum_{q=1}^{p-1} V(|\, \mathbf{x}_p - \mathbf{x}_q \,|) \tag{20}$$

and so, since H is even,

$$H_{\mathbf{K}} = (1/l^{Ns}) \int_0^l dx_{11} \cdots \int_0^l dx_{Ns}\, H(\mathbf{X}) \cos(k_{11}x_{11}) \cdots \cos(k_{Ns}x_{Ns}) \tag{21}$$

$$= \sum_{p=2}^{N} \sum_{q=1}^{p-1} \delta_{\mathbf{k}_1} \cdots V_{\mathbf{k}_p, \mathbf{k}_q} \cdots \delta_{\mathbf{k}_N} \tag{22}$$

where

$$V_{\mathbf{k}_1\mathbf{k}_2} = (1/l^{2s}) \int_0^l dx_{11} \cdots \int_0^l dx_{2s}\, V(|\mathbf{x}_1 - \mathbf{x}_2|) \cos(k_{11}x_{11}) \cdots \cos(k_{2s}x_{2s}) \quad (23)$$

$$\begin{aligned} \delta_{\mathbf{k}} &= 1, \qquad \mathbf{k} = 0 \\ &= 0, \qquad \mathbf{k} \neq 0 \end{aligned} \quad (24)$$

In general, this is as much as we can say about $V_{\mathbf{k}_1,\mathbf{k}_2}$. For *any* $\mathbf{X}$,

$$H(\mathbf{X}) = \sum_{p=2}^{N} \sum_{q=1}^{p-1} V(\mathbf{x}_p, \mathbf{x}_q) \quad (25)$$

where our method (13) of extending the definition outside means that $V(\mathbf{x}_p, \mathbf{x}_q)$ is no longer a function of $|\mathbf{x}_p - \mathbf{x}_q|$, or even $\mathbf{x}_p - \mathbf{x}_q$, only. Thus while the former would give us the usual result for central interactions

$$V_{\mathbf{k}_1,\mathbf{k}_2} = V_{\mathbf{k}_1}\delta_{\mathbf{k}_1+\mathbf{k}_2} \quad (26)$$

we cannot expect (26) to be true here.

There are, however, certain simplifying assumptions we can often make. Let λ be the typical distance in space over which $V(r)$ may vary appreciably. By integrating (23) by parts, it may be shown that if k_1 or k_2 is much greater than $\pi/\lambda = k_0$, $V_{\mathbf{k}_1,\mathbf{k}_2}$ will be very small. Suppose now that $s > 1$, and that $\lambda \ll l$, i.e., the interaction potential varies considerably over a distance much smaller than the side of the vessel containing the system. There will be a large number, of the order $(l/\lambda)^{2s}$, of pairs $\mathbf{k}_1$, $\mathbf{k}_2$ such that k_1 and k_2 do not greatly exceed k_0, and of these, all but a small proportion, of order λ/l, have *all* components of $\mathbf{k}_1$, $\mathbf{k}_2$ of order k_0. In this sense, we can speak of k_0 as the typical magnitude of wave vectors $\mathbf{k}_1$, $\mathbf{k}_2$ such that $V_{\mathbf{k}_1,\mathbf{k}_2} \neq 0$.

We may also show that $V_{\mathbf{k}_1,\mathbf{k}_2}$ need be taken as nonzero only when either the sum or the difference of each pair of corresponding components of $\mathbf{k}_1$, $\mathbf{k}_2$ is of order π/l, while each component will generally be of order π/λ. The number of such $\mathbf{k}_1$ and $\mathbf{k}_2$ is of order $(l/\lambda)^s$.

5. THE TIME EVOLUTION OF THE PROBABILITY DENSITY

Let $\rho(\mathbf{X}, \mathbf{V}, t)$ satify (8) and (14), and $\rho(\mathbf{X}, \mathbf{V}, 0)$ be given. Then, $\rho(\mathbf{X}, \mathbf{V}, t)$ is determined for all time.

Suppose that $\rho(\mathbf{X}, \mathbf{V}, t)$ tended to a limit $\rho(\mathbf{X}, \mathbf{V}, \infty)$ as $t \to \infty$. This would be a time-independent solution of (8). It is a well-known result of Fermi's[2] that except in the trivial case of no interactions, any such solution

[2] See Ref. 1, Appendix C or Ref. 7, Chapter 14 for a proof.

depends on $\mathbf{X}$ and $\mathbf{V}$ through $\mathscr{H}(\mathbf{X}, \mathbf{V})$ only. The conservation law (15) then shows that $\rho(\mathbf{X}, \mathbf{V}, \infty)$ would have to equal the *phase average* of the initial conditions, as usually defined. If there are no interactions, $\rho(\mathbf{X}, \mathbf{V}, \infty)$ would have to be the spatially homogeneous component $\rho_0(\mathbf{V}, 0)$ of ρ (which, we note, is here invariant under the change in sign of any component of $\mathbf{V}$).

We prove at once from (8) that

$$(d/dt) \int_{\Gamma} [\rho(\mathbf{X}, \mathbf{V}, t) - \rho(\mathbf{X}, \mathbf{V}, \infty)]^2 = 0 \tag{27}$$

which implies that unless ρ were originally a time-independent solution of Liouville's equation (in which case it does not vary), it can never tend to one as $t \to \infty$ or as $t \to -\infty$.

Indeed, the physical interpretation of Liouville's equation (8) is that ρ is constant at each representative point moving along a system trajectory. Therefore, unless ρ is originally constant along each such trajectory, i.e., an equilibrium solution of Liouville's equation, it cannot become constant in time. We expect that the motion of the system will be very complicated, and that representative points that were originally far separated will sometimes approach each other closely. Thus it seems likely that the behavior of ρ as a function of $\mathbf{X}$ and $\mathbf{V}$ will become more and more wildly oscillatory as $t \to \pm\infty$; later, we shall see from examples that this is in fact so.

However, we have already remarked at the end of Section 1 that we are not directly interested in ρ. All that we can measure in any given series of experiments on a given ensemble with a given standard of accuracy is the value of the *coarse-grained density* $\tilde{\rho}(\Omega, t)$, as defined by (6), at various times t. Here, Ω is a subset of Γ of measure greater than a certain lower bound fixed by the ensemble size and the required accuracy.

We are usually particularly interested in the evolution of *functions of state*, linear functions of ρ

$$\mathscr{F}(t) = \int_{\Gamma} \beta(\mathbf{X}, \mathbf{V})\, \rho(\mathbf{X}, \mathbf{V}, t)\, d^{Ns}X\, d^{Ns}V \tag{28}$$

where $\beta(\mathbf{X}, \mathbf{V})$ is some function of position and velocity; for example, $\beta(\mathbf{X}, \mathbf{V}) = V^2/2$ gives the average kinetic energy of the ensemble (i.e., the kinetic energy of a particular system will be a random variable whose mean is this expression). We measure $\mathscr{F}$ by dividing Γ up into a number M of cells, such that β has an approximately constant value β_i in cell i, counting the number n_i of systems in cell i, and summing

$$\mathscr{F}(t) = (1/G) \sum_{i=1}^{M} n_i \beta_i \tag{29}$$

$G = \sum n_i$ being, as before, the total number of systems in the ensemble. That is, we try to measure

$$\sum_{i=1}^{M} \mu(\Omega_i)\, \rho(\Omega_i, t) \tag{30}$$

which may be written as an integral

$$\int_{\Gamma} \beta(\mathbf{X}, \mathbf{V})\, \rho(\mathbf{X}, \mathbf{V}, t)\, d^{Ns}X\, d^{Ns}V \tag{31}$$

where

$$\rho(\mathbf{X}, \mathbf{V}, t) = \rho(\Omega_i, t)\ [(\mathbf{X}, \mathbf{V}) \in \Omega_i] \tag{32}$$

We would like to prove that for any Ω of a physically sensible form, $\rho(\Omega, t)$ tends to the expression $\rho(\Omega, \infty)$ obtained by coarse-graining $\rho(\mathbf{X}, \mathbf{V}, \infty)$ as defined above either as $t \to \infty$ or as $t \to -\infty$. If we could do this, we will have explained completely the fact that the observed values, given by (31), of expressions such as (28) tend to limits as $t \to \pm\infty$, the phenomenon known as "the approach to thermal equilibrium" or "irreversibility"; furthermore, we would have justified the use of the *microcanonical distribution* in equilibrium statistical mechanics.

6. THE "TIME-REVERSAL" AND "RECURRENCE" "PARADOXES"

Before investigating further the behavior of ρ, we relate our ideas to two problems which have been much discussed over the last century.

(a) We have stressed that any conclusions reached about the limit $t \to \infty$ should also be valid when $t \to -\infty$. Our coarse-graining technique gives no definite direction to time; we merely follow the behavior of certain functionals of ρ as it evolves according to the Liouville equation, which is time-reversible, i.e., invariant under the transformation $(X, V, t) \to (X, -V, -t)$. In the next section, we shall derive an equation for the evolution of a type of smoothed density; this equation will be time-reversible. It might seem, therefore, that our hopes of the last section are thwarted byt the "time-reversal paradox."

This, not quite as originally stated toward the end of the nineteenth century (for a bibliography, see Ref. 13) runs as follows: At time zero, we set up an ensemble corresponding to density $\rho(\mathbf{X}, \mathbf{V}, 0)$; we allow a long time T to elapse. We now set up another ensemble, with density $\rho(\mathbf{X}, -\mathbf{V}, T)$; after a further time T, this must have evolved back to $\rho(\mathbf{X}, \mathbf{V}, 0)$. (The particles of a system retrace their paths exactly.) Thus the ensemble may just as well evolve away from equilibrium as toward it.

The fallacy in this argument is clearly exposed by our previous discussion. For such a procedure, we must stipulate an experimental technique for setting up the system at $t = 0$; then, $\rho \approx \bar{\rho}$. But after a long time T, ρ and $\bar{\rho}$ will differ considerably, and it will become impossible to set up an ensemble representing density $\rho(\mathbf{X}, -\mathbf{V}, t)$ by the same technique without a likely error so large as to make the experiment meaningless. *Whenever we set up an experiment afresh, we must start with* $\rho \approx \bar{\rho}$. For a description of just such an "experiment" simulated on a computer, see Ref. 14.

(b) We emphasize that we are considering a system of finite size, containing any number of particles. It was first shown by Poincaré[9,15] that almost any initial state of such a system is such that the system will, after a sufficiently long time, return to an arbitrarily small neighbourhood (in Γ) of it. In fact, given any subset Ω of Γ, we can prove (Kac's recurrence theorem) that the average time taken for a point originally in Ω to return to Ω, the Poincaré recurrence time, is finite. This time will be very large indeed, far longer than any physical time scale, if the system contains more than a few particles and Ω is small. Nevertheless, it is often said that the existence of Poincaré recurrences precludes any useful kind of irreversible behavior; to obtain true irreversibility, we have to make the system infinitely large (in size, number of degrees of freedom, or both); for finite systems, irreversibility can occur only over times short compared to the Poincaré recurrence time.

This statement arises from confusion between the *actual* behavior of *one* member of an ensemble and the *mean* behavior of the ensemble *as a whole.* The actual proportion of members of the ensemble in a given region Ω of Γ at time t is an integer-valued random variable whose mean is $\bar{\rho}(\Omega, t)$; this actual density will show random fluctuations which will occasionally be large enough to carry an individual member right back to its initial state. We shall see this happening in our example of Section 8.

7. THE SMOOTHED PROBABILITY DENSITY

To proceed further, we must be more specific about the mehtod of constructing $\bar{\rho}$ from ρ. It is reasonable to assume that if Ω is a small set containing $(\mathbf{X}, \mathbf{V})$, the existence of $\lim_{|t|\to\infty} \bar{\rho}(\Omega, t)$ does not depend on the exact shape of Ω, provided that this is not too pathological. Hence we are justified in considering the more specialized functional of ρ, the "smoothed density"

$$\bar{\rho}(\mathbf{X}, \mathbf{V}, t) = \int \alpha(\Delta\mathbf{X}, \Delta\mathbf{V})\, \rho(\mathbf{X} - \Delta\mathbf{X}, \mathbf{V} - \Delta\mathbf{V}, t)\, d^{Ns}(\Delta\mathbf{X})\, d^{Ns}(\Delta\mathbf{V}) \tag{33}$$

where α is a well-behaved positive function, with

$$\int \alpha(\Delta\mathbf{X}, \Delta\mathbf{V})\, d^{Ns}(\Delta\mathbf{X})\, d^{Ns}(\Delta\mathbf{V}) = 1 \tag{34}$$

We expect the "smoothing kernel" α to be very small except when $|\Delta\mathbf{X}|$ and $|\Delta\mathbf{V}|$ are small; thus the smoothing corresponds to taking a small cell Ω round each point of Γ.

In fact, we specialize further to

$$\bar{\rho}(\mathbf{X}, \mathbf{V}, t) = \int \alpha(\Delta\mathbf{V})\, \rho(\mathbf{X}, \mathbf{V} - \Delta\mathbf{V}, t)\, d^{Ns}(\Delta\mathbf{V}) \tag{35}$$

where

$$\int \alpha(\Delta\mathbf{V})\, d^{Ns}(\Delta\mathbf{V}) = 1 \tag{36}$$

We shall find that (35) has all the properties that we want of (33) [or (6)], and it is much easier to handle.

We want to be able to work entirely in terms of $\bar{\rho}$. In order to do this, we must be able to derive from Liouville's equation an equation in closed form for the evolution of $\bar{\rho}$. Let us investigate the possibility of doing this for some α. We have

$$\begin{aligned} 0 &= \int d^{Ns}(\Delta\mathbf{V})\, \alpha(\Delta\mathbf{V}) \\ &\quad \times \{(\partial/\partial t)\, \rho(\mathbf{X}, \mathbf{V} - \Delta\mathbf{V}, t) + (\mathbf{V} - \Delta\mathbf{V}) \cdot (\partial\rho/\partial\mathbf{X}) + \mathbf{F} \cdot (\partial\rho/\partial\mathbf{V})\} \\ &= \mathscr{L}\bar{\rho} - \int d^{Ns}(\Delta\mathbf{V})\, \alpha(\Delta\mathbf{V})\, \Delta\mathbf{V} \cdot (\partial\rho/\partial\mathbf{X})(\mathbf{X}, \mathbf{V} - \Delta\mathbf{V}, t) \end{aligned}$$

If, now,

$$\sigma^2\, \partial\alpha(\Delta\mathbf{V})/\partial(\Delta\mathbf{V}) = -(\Delta\mathbf{V})\, \alpha(\Delta\mathbf{V})$$

so that α is Gaussian, of the normalized form

$$\alpha(\Delta\mathbf{V}) = [1/(2\pi\sigma^2)^{Ns/2}] \exp[-(\Delta\mathbf{V})^2/2\sigma^2] \tag{37}$$

an integration by parts gives

$$0 = \mathscr{L}\bar{\rho} + \sigma^2(\partial/\partial\mathbf{V}) \cdot (\partial\bar{\rho}/\partial\mathbf{X}) \tag{38}$$

as required. Henceforth, we shall study the behavior of $\bar{\rho}$ using (38), and shall denote by α the function (37). The variance or "spread," σ^2, is an arbitrary real (positive) quantity. If $\sigma = 0$, $\bar{\rho} = \rho$ and (38) reduces to (8).

We note particularly that (38) is time-reversible. However, we should expect fundamental differences between the behavior of solutions of (6) and

those of $\mathscr{L}\rho = 0$ as $t \to \pm\infty$, for the latter is a first-order partial differential equation corresponding to propagation along its characteristics, the system paths, while the former is a second-order equation of "pseudoparabolic" type—the second-order operator is hyperbolic rather than elliptic. Unfortunately, there are no general theorems on the behavior of solutions of such equations as $t \to \pm\infty$. The change from $\sigma = 0$ to $\sigma \neq 0$ introduces a term of higher degree; it is a "singular perturbation."

We may prove that

$$\rho(\mathbf{X}, \mathbf{V}, t) = \int d^{Ns}(\Delta\mathbf{V})\, \alpha(\Delta\mathbf{V})\, \bar{\rho}(\mathbf{X}, \mathbf{V} - i\,\Delta\mathbf{V}, t) \tag{39}$$

provided that ρ and $\bar{\rho}$ are analytic functions of the complex variables $\mathbf{V}$, and that the integral converges. This restriction, which we shall make throughout our work, is physically reasonable since we can find an analytic function defined on the real line and taking values arbitrarily close to a given continuous function.

The boundary conditions (10) or the periodic conditions (14) apply also to $\bar{\rho}$. In view of the discussion of Section 6(a), we usually restrict our initial conditions to those satisfying

$$\rho(\mathbf{X}, \mathbf{V}, 0) \approx \bar{\rho}(\mathbf{X}, \mathbf{V}, 0) \tag{40}$$

that is, we insist that our initial ρ and $\bar{\rho}$ vary little on scale σ (i.e., as $\mathbf{V}$ varies by a quantity of magnitude σ). In other words,

$$\sigma/\mu \ll 1 \qquad \text{where} \quad \mu \sim (1/\rho)|\, \partial\rho(\mathbf{X}, \mathbf{V}, 0)/\partial\mathbf{V}\,| \tag{41}$$

This is what we shall mean by saying that "σ is small."

8. NONINTERACTING SYSTEMS

As a first simple illustration of our theory so far, consider a noninteracting system. N noninteracting particles are subject to no external forces, save those confining them to a cubic box with perfectly reflecting walls of side l. $\rho(\mathbf{X}, \mathbf{V}, t)$ will satisfy the Liouville equation with $\mathbf{F} = \mathbf{0}$,

$$(\partial\rho/\partial t) + \mathbf{V} \cdot (\partial\rho/\partial\mathbf{X}) = 0 \tag{42}$$

In terms of Fourier components, as (16),

$$[\partial\rho_{\mathbf{K}}(\mathbf{V}, t)/\partial t] + i\mathbf{K} \cdot \mathbf{V}\rho_{\mathbf{K}} = 0 \tag{43}$$

$$\rho_{\mathbf{K}}(\mathbf{V}, t) = \rho_{\mathbf{K}}(\mathbf{V}, 0) \exp - i\mathbf{K} \cdot \mathbf{V}t \tag{44}$$

This is the most general solution of the Liouville equation. The $\rho_{\mathbf{K}}(\mathbf{V}, 0)$ are arbitrary functions of $\mathbf{V}$ satisfying (18) and an indistinguishability condition, and also such that $\rho(\mathbf{X}, \mathbf{V}, 0)$ is positive and

$$l^{Ns} \int \rho_0(\mathbf{V})\, d^{Ns}V = 1$$

Clearly $\rho_{\mathbf{K}}(\mathbf{V}, t)$ does not tend to any limit as $t \to \infty$ or as $t \to -\infty$; it rather oscillates more and more rapidly as $\mathbf{V}$ varies, as we mentioned earlier. On the other hand, for any Ω,

$$\rho(\Omega, t) = [1/\mu(\Omega)] \int_{\Omega} d^{Ns}X\, d^{Ns}V \sum_{\mathbf{K}} \rho_{\mathbf{K}}(\mathbf{V}, 0) \exp i\mathbf{K} \cdot (\mathbf{X} - \mathbf{V}t)$$

Provided the integration over Ω includes some integration over $\mathbf{V}$, each term of this expression will tend to zero as $t \to \pm\infty$ except that with $\mathbf{K} = \mathbf{0}$. This is assured by the Riemann–Lebesgue lemma.[16] So,

$$\lim_{|t|\to\infty} \bar{\rho}(\Omega, t) = [1/\mu(\Omega)] \int_{\Omega} d^{Ns}X\, d^{Ns}V\, \rho_0(\mathbf{V}, 0) \tag{45}$$

Thus, for these systems, our "coarse-graining" procedure does give us the results we hoped for in Section 5. We note particularly that it is the integration over $\mathbf{V}$ components that is important; this supports our idea of specializing the definition of $\bar{\rho}$ to (44) from (42).

The quantity $\bar{\rho}$, as defined by (35) and (37), satisfies

$$[\partial\bar{\rho}(\mathbf{X}, \mathbf{V}, t)/\partial t] + \mathbf{V} \cdot (\partial\bar{\rho}/\partial\mathbf{X}) + \sigma^2(\partial/\partial\mathbf{X}) \cdot (\partial\bar{\rho}/\partial\mathbf{V}) = 0 \tag{46}$$

i.e.,

$$[\partial\bar{\rho}_{\mathbf{K}}(\mathbf{V}, t)/\partial t] + i\mathbf{K} \cdot \mathbf{V}\bar{\rho}_{\mathbf{K}} + i\sigma^2\mathbf{K} \cdot \partial\bar{\rho}_{\mathbf{K}}/\partial\mathbf{V} = 0$$

thus

$$(\partial/\partial t)\{[\exp(-V^2/2\sigma^2)]\, \bar{\rho}_{\mathbf{K}}\} + i\sigma^2\mathbf{K} \cdot (\partial/\partial\mathbf{V})\{[\exp(-V^2/2\sigma^2)]\, \bar{\rho}_{\mathbf{K}}\} = 0 \tag{47}$$

so $[\exp(-V^2/2\sigma^2)]\, \bar{\rho}_{\mathbf{K}} = g(\mathbf{V} - i\mathbf{K}\sigma^2 t)$ for some function g. In terms of the initial conditions,

$$\bar{\rho}_{\mathbf{K}}(\mathbf{V}, t) = \bar{\rho}_{\mathbf{K}}(\mathbf{V} - i\sigma^2\mathbf{K}t, 0) \exp(-i\mathbf{K} \cdot \mathbf{V}t - \tfrac{1}{2}\sigma^2K^2t^2) \tag{48}$$

If $\sigma = 0$, (48) reduces to (44). We may in fact derive (48) from (44) by integration, using (35).

As long as $\bar{\rho}$ corresponds through (35) to an analytic ρ, $\bar{\rho}_{\mathbf{K}}(\mathbf{V}, t) \to 0$ as $|t| \to \infty$ for $\mathbf{K} \neq 0$; although $\bar{\rho}_{\mathbf{K}}(\mathbf{V} - i\mathbf{K}\sigma^2 t, 0)$ may increase with t, it will do so less rapidly than $\exp(-\sigma^2K^2t^2/2)$ decreases. Provided the initial

$\bar{\rho}(\mathbf{X}, \mathbf{V}, 0)$ satisfies (41), we may neglect the $-i\mathbf{K}\sigma^2 t$ compared to $\mathbf{V}$ in (48) for t such that $\bar{\rho}_{\mathbf{K}}$ is nonnegligible. Then,

$$\bar{\rho}_{\mathbf{K}}(\mathbf{V}, t) = \bar{\rho}_{\mathbf{K}}(\mathbf{V}, 0) \exp(-i\mathbf{K} \cdot \mathbf{V}t - \tfrac{1}{2}\sigma^2 K^2 t^2) \tag{49}$$

and will become negligibly small once $| t | \gg (\sigma K)^{-1}$. This is the characteristic decay time of a single Fourier component of $\bar{\rho}$. All non-spatially homogeneous components must disappear after a time of order l/σ.

Usually, however, the time over which an initially non-spatially homogeneous $\bar{\rho}(\mathbf{X}, \mathbf{V}, 0)$ decays to its spatial average will be much shorter than this. We shall often assume that $\bar{\rho}(\mathbf{X}, \mathbf{V}, 0)$ varies substantially in space over a distance λ_l much smaller than l. In this case, an analysis very similar to that of Section 4 shows that the vast majority [all but a proportion $O(\lambda_l/l)$] of the nonnegligible $\bar{\rho}_{\mathbf{K}}(\mathbf{V}, 0)$ have $K \sim \lambda_l^{-1}$. All these will decay in a time of order λ_l/σ, and $\bar{\rho}(\mathbf{X}, \mathbf{V}, t)$ will be very near its limit after this time.

Furthermore, any expression containing $\bar{\rho}$ that involves an integration over $\mathbf{X}$, that is, a summation over $\mathbf{K}$, will reach its limiting value more quickly still. Consider, for example,

$$\int d^{Ns}X f(\mathbf{X})\, \bar{\rho}(\mathbf{X}, \mathbf{V}, t) = \sum_{\mathbf{K}} \bar{\rho}_{\mathbf{K}}(\mathbf{V}, 0) \exp(-i\mathbf{K} \cdot \mathbf{V}t - \tfrac{1}{2}\sigma^2 K^2 t^2) \tag{50}$$

The expression

$$\sum_{\mathbf{K}} f_{-\mathbf{K}} \bar{\rho}_{\mathbf{K}}(\mathbf{V}, 0) \exp -i\mathbf{K} \cdot \mathbf{V}t \tag{51}$$

that is, the corresponding expression with $\sigma = 0$, never tends to zero; in fact, it may be periodic in t. However, if it is periodic, the period is of order l/v_0, where v_0 is a typical value of a particle velocity component (generally $v_0 \approx \mu$, but we shall keep them separate for the time being). The mutual interference of terms with different values of $\mathbf{K} \cdot \mathbf{V}$ in (51) is likely to make the behavior of the whole sum roughly as shown in Fig. 1.

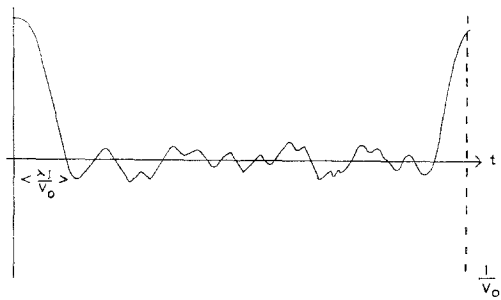

Fig. 1. Periodic fluctuations of expression (51).

After a time of order λ_I/v_0 has elapsed, the sum will have been reduced in magnitude to a tiny fraction of its initial value, and will execute fairly random oscillations about zero until time l/v_0, when the exponentials may once again be in phase. The average value of the sum during this period is seen, by considering the mean of its square, to be about $C^{-1/2}$ of its initial value, where C is the total number of terms in the sum. C will be some power of l/λ_I, depending on the complexity of the initial conditions. If we now assume that

$$l/v_0 \gg \lambda_I/\sigma \tag{52}$$

the extra exponential in (50) will have knocked out all the terms of the sum well before any periodicity can occur. Thus we seem to be justified in assuming functionals such as (50) to decay to zero in a time of order λ_I/v_0.

We thus have four time scales in our problem; in terms of increasing order of magnitude, these are as follows:

(a) λ_I/v_0: Over this time, *functions of state* reach their limiting values.

(b) λ_I/σ: Over this time, *individual Fourier components of* $\bar{\rho}$ reach their limiting values.

(c) l/v_0: A particle takes this time to *cross the containing vessel.*

(d) l/σ: Any *bulk inhomogenieties*, of spatial extent $\sim l$, take this time to be erased.

We have taken the side of the vessel containing the system to be so large that (1) the length scale of the inhomogenieties is much smaller than l; (2), the decay time of the inhomogenieties is much smaller than l/v_0.

In the next section, we shall show that the results we have obtained using these assumptions are very similar to those obtained by allowing l to become actually infinite.

We stress, however, that neither of these assumptions is necessary to get $\bar{\rho}$ tending to a limit; it must do this in a time of order l/σ, whatever the size of the vessel and the initial conditions.

9. THE INFINITE SYSTEM CASE

As we have seen, if $\sigma = 0$, expressions for functions of state such as (50) may be periodic. In fact, of course,

$$\rho(\mathbf{X}, \mathbf{V}, t) = \sum_{\mathbf{K}} \rho_{\mathbf{K}}(\mathbf{V}, 0) \exp i\mathbf{K} \cdot (\mathbf{X} - \mathbf{V}t) = \rho(\mathbf{X} - \mathbf{V}t, \mathbf{V}, 0) \tag{53}$$

which is periodic provided the components of $\mathbf{V}$ are mutually commensurable; this is just a consequence of our periodic conditions. To make this period

tend to infinity, we must make l tend to infinity; the Fourier spectra of $\tilde{\rho}$, H, etc. become continuous rather than discrete,

$$\rho(\mathbf{X}, \mathbf{V}, t) = \int d^{Ns}\mathbf{K}\, \rho^{+}(\mathbf{K}, \mathbf{V}, t) \exp i\mathbf{K} \cdot \mathbf{X} \tag{54}$$

etc.; ρ^{+} is the Fourier transform of ρ. The behavior of an individual $\rho^{+}(\mathbf{K})$ is exactly the same as in the discrete case, but any *wave packet* formed by integration over a continuous range of $\mathbf{K}$,

$$\int_{\Omega_k} d^{Ns}\mathbf{K}\, \rho^{+}(\mathbf{K}, \mathbf{V}, t) \tag{55}$$

will tend to a limit, similarly to the integral over $\mathbf{V}$ (45). In particular, since the Fourier transform of an analytic function $f(\mathbf{X})$ is

$$f^{+}(\mathbf{K}) = f_0\, \delta(\mathbf{K}) + f^{+\prime}(\mathbf{K}) \tag{56}$$

where $f^{+\prime}(\mathbf{K})$ is analytic, f_0 the spatial average of $f(\mathbf{X})$, and $\delta(\mathbf{K})$ the Dirac δ function, we shall have that

$$\int d^{Ns}\mathbf{X} f(\mathbf{X})\, \rho(\mathbf{X}, \mathbf{V}, t) = \int d^{Ns}\mathbf{K} f^{+}(-\mathbf{K})\, \rho^{+}(\mathbf{K}, \mathbf{V}, 0) \exp -i\mathbf{K} \cdot \mathbf{V}t \tag{57}$$

tends to a limiting value $f_0\rho_0(\mathbf{V}, 0)$ in a time of order λ_l/v_0 (using, once again, the Riemann–Lebesgue lemma). This is the same result as that of the last section, proved rather more rigorously. Provided we retain the condition (52) as we let $l \to \infty$ and $\sigma \to 0$ (as we have done in this section by letting $l \to \infty$ and *then* considering ρ), our discussion leading to (52) becomes steadily more exact. Three of our four time scales become infinite, but the smallest remains finite.

We notice also that (53) is trivially still true for an infinite system; if the disturbances to ρ are initially localized in space, $\rho(\mathbf{X}, \mathbf{V}, t)$ will eventually reach local spatial homogeniety for any given $\mathbf{X}$. The disturbances are "rejected to infinity."

It is the method of this section—to make the size of the system actually infinite and hence to smooth ρ over a continuous Fourier spectrum—that has been the basis of most work in nonequilibrium statistical mechanics over the last fifteen years, particularly that of the Brussels school.[7,8,17] It is essential to make the size of the system actually infinite; if it is finite, however large, the spatial Fourier spectrum of ρ will be discrete, and $\rho_{\mathbf{K}}(\mathbf{V}, t)$ may be periodic.

In order to make the system size infinite while keeping the average density of particles greater than zero, as is often desirable, the number N of particles in the system must be allowed to become infinite. Ordinary

theories of integration are then inadequate to give meaning to expressions of the form (55) or (57), and, indeed to the definition of ρ. Rigorous work on this subject becomes very difficult.

By building smoothing into our definition of $\bar{\rho}$, we hope to avoid having to make l become infinite rather than large and finite. We have seen how in the case of an interactionless system our hopes are fulfilled.

10. REDUCED PROBABILITY DENSITY FUNCTIONS

We are normally interested only in reduced probability densities specifying the likely positions and velocities of one or two particles, though it must be emphasized that there is no conceptual reason for this restriction, only a practical one. In theory, we can set up apparatus for measuring $\bar{\rho}(\Omega, t)$ defined by (6) for a quite arbitrary open subset Ω of Γ having nonzero measure. If we take a large enough ensemble of systems, we can make this measurement with any specified accuracy. What we cannot do is measure $\rho(\mathbf{X}, \mathbf{V}, t)$ itself. In practice, however, it is not worth building the complicated (but finite) apparatus or setting up the large (but finite) ensembles required for attempts to measure higher distribution fuctions.

The smoothed q-particle probability density function $\bar{f}_q$ may be defined either from the unsmoothed density f,

$$\begin{aligned}\bar{f}_q(\mathbf{x}_1,\ldots,\mathbf{x}_q;\mathbf{v}_1,\ldots,\mathbf{v}_q;t) &= \int d^s(\Delta v_1)\cdots\int d^s(\Delta v_q)\times(2\pi\sigma^2)^{-qs/2}\exp[-2\sigma^{-2}(\Delta v_1^2+\cdots+\Delta v_q^2)] \\ &\quad\times f_q(\mathbf{x}_1,\ldots,\mathbf{x}_q;\mathbf{v}_1-\Delta\mathbf{v}_1,\ldots,\mathbf{v}_q-\Delta\mathbf{v}_q;t) \qquad (58)\end{aligned}$$

or by replacing ρ by $\bar{\rho}$ in the usual definition of f. Integrating (38) over $\mathbf{x}_{q+1},\ldots,\mathbf{v}_N$, and substituting (20), we obtain, for $1 < q < N-1$,

$$\begin{aligned}&\frac{\partial\bar{f}_q}{\partial t}+\sum_{j=1}^{q}\mathbf{v}_j\cdot\frac{\partial\bar{f}_q}{\partial\mathbf{x}_j}-\sum_{\substack{i=1\\i\neq j}}^{q}\sum_{j=1}^{q}\frac{\partial V(\mathbf{x}_i,\mathbf{x}_j)}{\partial\mathbf{x}_i}\cdot\frac{\partial\bar{f}_q}{\partial\mathbf{v}_i}+\sigma^2\sum_{j=1}^{q}\frac{\partial}{\partial\mathbf{x}_j}\cdot\frac{\partial\bar{f}_q}{\partial\mathbf{v}_j} \\ &\quad=(N-q)\sum_{j=1}^{q}\int_{\mu}d^sx_{q+1}\,d^sv_{q+1}\frac{\partial V(\mathbf{x}_j,\mathbf{x}_{q+1})}{\partial\mathbf{x}_j}\cdot\frac{\partial\bar{f}_{q+1}}{\partial\mathbf{v}_j} \qquad (59)\end{aligned}$$

just the BBGKY hierarchy with one extra term.

Before going on, we note one important physical point. If we assume that we can describe the evolution of the average values of quantities we want to measure entirely in terms of f_1 or $\bar{f}_1$, we can discard entirely the idea of an ensemble of systems and say that for N large, an accurate measurement of $\bar{f}_1$ can be obtained by measuring the number of particles whose position and

velocity coordinates lie near $\mathbf{x}_1$ and $\mathbf{v}_1$. This is the viewpoint generally adopted in kinetic theory. But we stress once again that our discussion of Sections 1 and 5 is valid with only minor rewording: The number of particles of the system in a region ω of μ at a given time is a random variable whose mean is (from now on we shall assume that $s = 3$)

$$N \int_{\omega} f_1(\mathbf{x}_1, \mathbf{v}_1, t)\, d^3x_1\, d^3v_1 \tag{60}$$

Its variance will also be proportional to N, and so as N increases, the observed density is likely to lie proportionately closer and closer to (60). Thus we may regard a single large system as an ensemble of weakly coupled one-particle systems. This is the "μ-space viewpoint," in contrast to the "Γ-space viewpoint" that we shall generally use. While the former requires the system to be large to obtain meaningful results, the latter does not.

It may be shown[18] that for a large, weakly interacting system in equilibrium, f_1 is a Gaussian function of velocity only. In general, the distribution of the energy of any small component of a large, weakly interacting system will be of Maxwellian form (the *canonical ensemble* of equilibrium statistical mechanics). It is only in the μ-space viewpoint that the Maxwellian distribution has any special significance.

11. THE VLASOV EQUATION. ELECTRON PLASMA OSCILLATIONS

If we assume that the particles are uncorrelated,

$$\bar{f}_2(\mathbf{x}_1, \mathbf{x}_2; \mathbf{v}_1, \mathbf{v}_2; t) = \bar{f}_1(\mathbf{x}_1, \mathbf{v}_1, t)\bar{f}_1(\mathbf{x}_2, \mathbf{v}_2, t) \tag{61}$$

then (59) with $q = 1$ reduces to

$$\frac{\partial \bar{f}_1}{\partial t} + \mathbf{v}_1 \cdot \frac{\partial \bar{f}_1}{\partial x_1} + \sigma^2 \frac{\partial}{\partial \mathbf{x}_1} \cdot \frac{\partial \bar{f}_1}{\partial \mathbf{v}_1}$$
$$= (N-1) \int_{\mu} d^3x_2\, d^3v_2 \frac{\partial V(\mathbf{x}_1, \mathbf{x}_2)}{\partial \mathbf{x}_1} \cdot \frac{\partial \bar{f}_1(\mathbf{x}_1, \mathbf{v}_1, t)}{\partial \mathbf{v}_1} \bar{f}_1(\mathbf{x}_2, \mathbf{v}_2, t) \tag{62}$$

When $\sigma = 0$, this is Vlasov's equation. Its most important application is the study of electron plasma oscillations, where

$$V(\mathbf{x}_1, \mathbf{x}_2) = e^2/(4\pi m \epsilon_0 |\, \mathbf{x}_1 - \mathbf{x}_2 \,|) \quad [(\mathbf{x}_1, \mathbf{v}_1), (\mathbf{x}_2, \mathbf{v}_2) \in \mu] \tag{63}$$

[and the even periodic continuation for $\mathbf{x}_1$ or $\mathbf{x}_2$ outside this range, as in (3.6); thus we still write $V(\mathbf{x}_1, \mathbf{x}_2)$ and *not* $V(|\, \mathbf{x}_1 - \mathbf{x}_2 \,|)$].

If we write down the total electric field (excluding that due to the neutralizing effect of a positive ion background, which, to a first approxi-

mation, plays no part in the motion since ions are so much heavier than electrons) $\mathbf{E}$,

$$(-e/m)\,\mathbf{E}(x,t) = (N-1)\int\!\!\int_{\mu} d^3x'\,d^3v'\,[\partial V(\mathbf{x},\mathbf{x}')/\partial\mathbf{x}]\,f(\mathbf{x}',\mathbf{v}',t) \quad (64)$$

(changing $\mathbf{x}_1$ to $\mathbf{x}$, $\mathbf{x}_2$ to $\mathbf{x}'$, and dropping the suffix 1 on f from now on), we have

$$0 = \frac{\partial \bar{f}(\mathbf{x},\mathbf{v},t)}{\partial t} + \mathbf{v}\cdot\frac{\partial \bar{f}}{\partial \mathbf{x}} + \sigma^2\frac{\partial}{\partial \mathbf{x}}\cdot\frac{\partial \bar{f}}{\partial \mathbf{v}} - \frac{e}{m}\,\mathbf{E}\cdot\frac{\partial \bar{f}}{\partial \mathbf{v}} = 0 \quad (65)$$

Here, $\bar{f}$ satisfies the periodic conditions (14) with $\mathbf{x}$ for $\mathbf{X}$, and $\mathbf{E}$ satisfies

$$\mathbf{E}(\mathbf{x}',t) = \pm\mathbf{E}(\mathbf{x},t) \quad (66)$$

in such a way that it is an odd function of each component of $\mathbf{x}$. If we stipulate that $\bar{f}$ must be such that $\mathbf{E}(\mathbf{x},t) = 0$ whenever $\mathbf{x}$ lies on a boundary of μ, $\mathbf{E}$ will be a continuous function of $\mathbf{x}$ at these boundaries.

The pair (64)–(65) with $\sigma = 0$ has been the subject of much study. A solution of the linearised problem for small perturbations about equilibrium was first obtained by Landau.[4] The techniques used were justified by Backus.[19] Both these papers refer to an infinite plasma, as does the account of Ref. 20, whose notation we shall follow where possible. Boundary conditions like ours were introduced in Ref. 11.

It is well known that the linearization used breaks down after a long enough time, however small the initial perturbation (for numerical examples, see Ref. 19). We shall now, however, apply the method of Landau to (64) and (65) with $\sigma \neq 0$, and show that by working with $\bar{f}$ instead of f, we may recover the exponential decay of the electric field in a self-consistent way: the linearization does not break down.

As usual, we write

$$\bar{f}(\mathbf{x},\mathbf{v},t) = \bar{f}_0(\mathbf{v}) + \bar{f}_{(1)}(\mathbf{x},\mathbf{v},t) \quad (67)$$

$$\mathbf{E}(\mathbf{x},t) = \mathbf{E}_{(1)}(\mathbf{x},t) \quad (68)$$

where $\bar{f}_{(1)}$ and $\mathbf{E}_{(1)}$ are assumed "small." Then,

$$\frac{e}{m}\,\mathbf{E}_{(1)}(\mathbf{x},t) = (N-1)\int\!\!\int_{\mu} d^3x'\,d^3v'\,\frac{\partial V(\mathbf{x},\mathbf{x}')}{\partial \mathbf{x}}\,\bar{f}_{(1)}(\mathbf{x}',\mathbf{v}',t) \quad (69)$$

$$0 = \frac{\partial \bar{f}_{(1)}}{\partial t} + \mathbf{v}\cdot\frac{\partial \bar{f}_{(1)}}{\partial x} + \sigma^2\frac{\partial}{\partial \mathbf{v}}\cdot\frac{\partial \bar{f}_{(1)}}{\partial \mathbf{x}} - \frac{e}{m}\,\mathbf{E}_{(1)}\cdot\frac{\partial \bar{f}_0}{\partial \mathbf{v}} \quad (70)$$

where in (70) we have neglected a product of $\mathbf{E}_{(1)}$ and $\partial\bar{f}_{(1)}/\partial\mathbf{x}$; our approximation is that this term is small compared to the last of (70), i.e.,

$$|\,\partial\bar{f}_{(1)}/\partial\mathbf{v}\,| \ll |\,\partial\bar{f}_0/\partial\mathbf{v}\,| \quad (71)$$

Define Fourier components $E_{\mathbf{k}}$ and $\bar{f}_{\mathbf{k}}$ of $E_{(1)}$ and $\bar{f}_{(1)}$ as before. From (70),

$$0 = \frac{\partial \bar{f}_{\mathbf{k}}}{\partial t} + i\mathbf{k} \cdot \mathbf{v} f_{\mathbf{k}} + i\sigma^2 \mathbf{k} \cdot \frac{\partial \bar{f}_{k}}{\partial \mathbf{v}} - \frac{e}{m} \mathbf{E}_{\mathbf{k}} \cdot \frac{\partial \bar{f}_0}{\partial \mathbf{v}} \tag{72}$$

Equation (69) shows that $E_{\mathbf{k}}$ is parallel to $\mathbf{k}$, while from (69) using the continuity of $\mathbf{E}(\mathbf{x}, t)$,

$$i\mathbf{k} \cdot \mathbf{E}_{\mathbf{k}}(t) = [-(N-1)e/\epsilon_0] \int d^3v' \, \bar{f}_{\mathbf{k}}(\mathbf{v}', t) \tag{73}$$

For each $\mathbf{k}$, (72) and (73) are a coupled pair of equations for $\mathbf{E}_{\mathbf{k}}$ and $f_{\mathbf{k}}$; there is no coupling between different Fourier components in this approximation. If then we write $E_{\mathbf{k}}$ for the magnitude of $\mathbf{E}_{\mathbf{k}}$, and define

$$\bar{F}_0(u) = l^3 \int \bar{f}_0(\mathbf{v}) \, \delta(u - [\mathbf{k} \cdot \mathbf{v}/|\, \mathbf{k} \,|]) \, d^3v \tag{74}$$

(this depends on $\mathbf{k}$ unless $\bar{f}_0$ is isotropic in $\mathbf{v}$), and

$$\bar{F}_{\mathbf{k}}(u, t) = l^3 \int \bar{f}_{\mathbf{k}}(\mathbf{v}, t) \, \delta(u - [\mathbf{k} \cdot \mathbf{v}/|\, \mathbf{k} \,|]) \, d^3v \tag{75}$$

where $k = |\, \mathbf{k} \,|$, the two transverse velocity components may be eliminated, giving

$$[\partial \bar{F}_{\mathbf{k}}(u, t)/\partial t] + iku\bar{F}_{\mathbf{k}} + i\sigma^2 k(\partial \bar{F}_{\mathbf{k}}/\partial u) = (e/m) \, E_{\mathbf{k}} \, \partial \bar{F}_0/\partial u \tag{76}$$

$$E_{\mathbf{k}}(t) = [i(N-1)e/kl^3\epsilon_0] \int_{-\infty}^{\infty} \bar{F}_{\mathbf{k}}(u, t) \, du \tag{77}$$

We desire the solution of (76) and (77) given the initial value of $\bar{F}_{\mathbf{k}}$,

$$\bar{F}_{\mathbf{k}}(u, 0) = \bar{G}_{\mathbf{k}}(u) \tag{78}$$

The solution of (76) with (78) is

$$\begin{aligned} \bar{F}_{\mathbf{k}}(u, t) = {} & \bar{G}_{\mathbf{k}}(u - ik\sigma^2 t) \exp(-ikut - \tfrac{1}{2}\sigma^2 k^2 t^2) \\ & + (e/m) \int_0^t d\tau \, E_{\mathbf{k}}(t - \tau)[\exp(-iku\tau - \tfrac{1}{2}\sigma^2 k^2 \tau^2)] \, \bar{F}_0'(u - ik\sigma^2 \tau) \end{aligned} \tag{79}$$

To obtain this solution, we have assumed, as in the standard theory, that $\bar{F}_0(u)$ and $\bar{G}_{\mathbf{k}}(u)$ are analytic functions at all finite points of the complex u plane; they are entire functions. Substituting into (77),

$$\begin{aligned} E_{\mathbf{k}}(t) = {} & Q_{\mathbf{k}}(t) + [i(N-1) \, e^2/\epsilon_0 mkl^3] \int_0^t d\tau \, E_{\mathbf{k}}(t - \tau) \\ & \times \int_{-\infty}^{\infty} du[\exp(-iku\tau - \tfrac{1}{2}\sigma^2 k^2 \tau^2)] \, \bar{F}_0'(u - ik\sigma^2 \tau) \end{aligned} \tag{80}$$

where

$$Q_{\mathbf{k}}(t) = [i(N-1)e/\epsilon_0 k l^3] \int_{-\infty}^{\infty} \bar{G}_{\mathbf{k}}(u - ik\sigma^2 t) \exp(-ikut - \tfrac{1}{2}\sigma^2 k^2 t^2)\, du \tag{81}$$

Equation (80) is a convolution. Thus if $\bar{F}_{\mathbf{k}}$ and $E_{\mathbf{k}}$ have Laplace transforms

$$E_{\mathbf{k}}(p) = \int_0^{\infty} E_{\mathbf{k}}(t)\, e^{-pt}\, dt \tag{82}$$

etc., we have

$$\epsilon_0 k F_{\mathbf{k}}(p)\, l^3/i(N-1)e = \int_{-\infty}^{\infty} du\, L_1(\mathbf{k}, u, p) + (e/m)\, E_{\mathbf{k}}(p) \int_{-\infty}^{\infty} du\, L_2(\mathbf{k}, u, p) \tag{83}$$

where

$$L_1(\mathbf{k}, u, p) = \int_0^{\infty} dt\, \bar{G}_{\mathbf{k}}(u - ik\sigma^2 t) \exp[-(p + iku)t - \tfrac{1}{2}\sigma^2 k^2 t^2] \tag{84}$$

$$L_2(\mathbf{k}, u, p) = \int_0^{\infty} dt\, \bar{F}_0'(u - ik\sigma^2 t) \exp[-(p + iku)t - \tfrac{1}{2}\sigma^2 k^2 t^2] \tag{85}$$

That is,

$$E_{\mathbf{k}}(p) = [i(N-1)e/\epsilon_0 k D(\mathbf{k}, p)\, l^3] \int_{-\infty}^{\infty} du\, L_1(\mathbf{k}, u, p) \tag{86}$$

where

$$D(\mathbf{k}, p) = 1 - [i(N-1)\, e^2/\epsilon_0 m k l^3] \int_{-\infty}^{\infty} du\, L_2(\mathbf{k}, u, p) \tag{87}$$

If $\sigma = 0$, we recover the standard expression for D.[20]

12. DISCUSSION OF RESULTS. LANDAU DAMPING

By means of some rather heavy algebra, the following may be shown:

(a) The exponential rate of growth of $E_{\mathbf{k}}(t)$ as given by (76)–(78) is bounded above, so that the Laplace transform (82) must exist for and be an analytic function of p, for large enough p (the proof is very similar to that for the $\sigma = 0$ case[9]).

(b) $L_1(\mathbf{k}, u, p)$ and $L_2(\mathbf{k}, u, p)$ are analytic functions of u and p throughout the complex u and p planes.

As a consequence of (b), we can write

$$\int_{-\infty}^{\infty} L_2(\mathbf{k}, u, p)\, du = \int_0^{\infty} dt\, e^{-pt} \times \int_{-\infty}^{\infty} du\, \bar{F}_0'(u - ik\sigma^2 t) \exp(-ikut - \tfrac{1}{2}\sigma^2 k^2 t^2)$$

reversing the order of integration over u, p. But from the definitions of Section 7,

$$\bar{F}_0'(u - ik\sigma^2 t)\exp(-ikut - \tfrac{1}{2}\sigma^2 k^2 t^2)$$
$$= (2\pi\sigma^2)^{-1/2}\int_{-\infty}^{\infty} du'\, F_0(u')\exp\{-iku't - [(u-u')^2/2\sigma^2]\}$$

and so, integrating out the u,

$$\int_{-\infty}^{\infty} L_2(\mathbf{k}, u, p)\, du = \int_{-\infty}^{\infty} du \int_0^{\infty} dt\, e^{-(p+iku)t} F_0'(u)$$
$$= \int_{-\infty}^{\infty} \frac{du\, F_0'(u)}{p + iku} \tag{88}$$

This is all for $\operatorname{Re} p > 0$, when the integrals on the right converge; but as we well know, the expression (88) may be continued analytically to $\operatorname{Re} p < 0$. In other words, the integrals over u occurring in (86) and (87) are the same whatever the value of σ (provided the definition of $\bar{F}_0$ is adjusted). Thus we have the very sensible result that the expression derived for $E_{\mathbf{k}}(p)$ is the same whatever the value of σ taken in the initial approximation.

We may now discuss the behavior of $E_{\mathbf{k}}(t)$ as usual; (86) must give $E_{\mathbf{k}}(p)$ for large enough $\operatorname{Re} p$ and its analytic continuation elsewhere. This continuation is analytic at all points of the complex p plane except for poles at the zeros of $D(\mathbf{k}, p)$. The expression

$$E_{\mathbf{k}}(t) = (1/2\pi i)\int_{C_0} E_{\mathbf{k}}(p)\, e^{pt}\, dp \tag{89}$$

certainly exists if C_0 is the contour $-\infty < \operatorname{Im} p < \infty$, $\operatorname{Re} p$ large enough. Further, as $\operatorname{Im} p \to \infty$, the numerator of $E_{\mathbf{k}}(p)$ tends to zero and the denominator to unity. So $E_{\mathbf{k}}(p) \to 0$. Thus it is permissible to move the ends of the inversion contour (at infinity) an arbitrary finite distance leftward. So,

$$E_{\mathbf{k}}(t) = \sum_s R_s \exp[p_s(\mathbf{k})t] + (1/2\pi i)\int_C E_{\mathbf{k}}(p)\exp(pt)\, dp \tag{90}$$

where C is any contour running from $\operatorname{Im} P = -\infty$ to $\operatorname{Im} p = \infty$, with $\operatorname{Re} p$ bounded; $p_s(\mathbf{k})$ are the zeros of $\bar{D}_{\mathbf{k}}(p)$ that lie to the right of C; and R_s are the residues of $E_{\mathbf{k}}(p)$ at these points. [There can be only a finite number of them, as $E_{\mathbf{k}}(p) \to 0$ as $\operatorname{Im} p \to \infty$; we assume C does not go through a pole.] If C lies to the left of $\operatorname{Re} p = -\alpha$, $\alpha > 0$, the integral in (89) will decrease to zero faster than $e^{-\alpha|t|}$ as $t \to \infty$ (applying the usual damping theorems,

based on the Riemann–Lebesgue lemma). In most practical situations, we can expect that there is a zero p_1 of $D_{\mathbf{k}}(p)$ with a very small negative real part, and in this case, after a very long positive time has elapsed,

$$E_{\mathbf{k}}(t) \approx R_1 \exp[p_1(\mathbf{k})\, t] \tag{91}$$

This is the phenomenon of "Landau damping" of the electric field; $\mathrm{Re}\, p_1$ is called the "Landau damping decrement."

Let us now consider the behavior of $\bar{F}_{\mathbf{k}}$, using (79). The first term tends to zero as $|t| \to \infty$, while the second has Laplace transform

$$E_{\mathbf{k}}(p)\, L_2(\mathbf{k}, u, p) \tag{92}$$

L_2 is an entire function of p; it does *not* have a pole at $p + iku = 0$, as does the corresponding expression for $\sigma = 0$. So the behavior of $\bar{F}_{\mathbf{k}}$ is given by an expression similar to (91); it, together with any other quantity we can actually observe, will show Landau damping. In particular, $\partial \bar{F}_{\mathbf{k}}(u, t)/\partial u$ will show Landau damping.

In the case $\sigma = 0$, we have

$$\bar{F}_{\mathbf{k}}(u, t) = e^{-ikut} G_{\mathbf{k}}(u) + \text{Landau-damped terms} \tag{93}$$

so that $\partial F_{\mathbf{k}}/\partial u$ increases linearly with time; eventually, the fundamental linearizing approximation we have made, (71), must break down. When $\sigma \neq 0$, however,

$$\bar{F}_{\mathbf{k}}(u, t) = [\exp(-ikut - \tfrac{1}{2}\sigma^2 k^2 t^2)]\, G_{\mathbf{k}}(u - ik\sigma^2 t) + \text{Landau-damped terms} \tag{94}$$

The magnitude of $\partial \bar{F}_{\mathbf{k}}/\partial u$ will increase to a maximum at $t \sim 1/\sigma k$, the phase mixing time for waves of length $2\pi/k$, as in Section 8. The value of this maximum will be of order $G_{\mathbf{k}}(u)/\sigma$, and the linearization will not break down provided

$$|\, \partial \bar{F}_0(u)/\partial u \,| \gg (1/\sigma) |\, \bar{G}_{\mathbf{k}}(u)| \tag{95}$$

i.e., provided that the magnitude of the initial perturbation to $\bar{F}_0$ is much less than the amount by which $\bar{F}_0$ vaires on scale σ,

$$|\, \bar{G}_{\mathbf{k}}(u)| \ll (\sigma/\mu)\, \bar{F}_0 \tag{96}$$

We note particularly that throughout this chapter we have avoided assuming that σ is small, so (95) is not a strong restriction; provided that initially the perturbation is small compared to $\bar{F}_0$

$$|\, G_{\mathbf{k}} \,| \ll F_0 \tag{97}$$

we can find a σ such that (95) is true.

We have thus shown how a self-consistent derivation of the theory of Landau damping of electrostatic plasma waves may be obtained by working in terms of $\bar{f}$ rather than in terms of f.

ACKNOWLEDGMENTS

This work was done during the tenure of a Junior Research Fellowship at Churchill College, Cambridge, and of a Research Studentship awarded by the Central Electricity Generating Board. My thanks are due to Drs. J. P. Dougherty and B. D. Josephson for their advice and help.

REFERENCES

1. L. Boltzmann, *Vorlesungen über Gastheorie* (Leipzig) (1896).
2. S. Chapman and T. G. Cowling, *The Mathematical Theory of Nonuniform Gases* (Dover, New York, 1936).
3. A. I. Khinchin, *Mathematical Foundations of Statistical Mechanics* (Dover, New York, 1949).
4. L. D. Landau, *J. Phys. (USSR)* **10**:25 (1946).
5. N. N. Bogoliubov, *Problemi Dynamitcheski Theori i Statisticheskei* (Moscow, 1946).
6. R. Brout and I. Prigogine, *Physica* **22**:621 (1956).
7. I. Prigogine, *Nonequilibrium Statistical Mechanics* (Interscience, New York, 1963).
8. R. Balescu, *Statistical Mechanics of Charged Particles* (Interscience, New York, 1963).
9. I. E. Farquhar, *Ergodic Theory in Statistical Mechanics* (Interscience, New York, 1964).
10. R. C. Tolman, *The Principles of Statistical Mechanics* (Oxford, 1938).
11. D. Montgomery and D. Gorman, *Phys. Fluids* **5**:947 (1962).
12. S. Gartenhaus, *Phys. Fluids* **6**:451 (1963).
13. D. ter Haar, *Rev. Mod. Phys.* **27**:289 (1955).
14. A. Bellemans and J. Orban, *Phys. Lett.* **24A**:620 (1967).
15. R. Jancel, *Foundations of Classical and Quantum Statistical Mechanics* (Pergamon, 1969).
16. E. C. Titchmarsh, *The Theory of Fourier Integrals* (Oxford, 1958).
17. P. Résibois, in *Many Body Physics*, L. Meeron, ed. (Gordon and Breach, New York, 1967).
18. B. R. A. Nijboer, in *Fundamental Probems of Statistical Mechanics*, Vol. I, E. G. D. Cohen, ed. (North-Holland Publ. Co., Amsterdam, 1962).
19. G. Backus, *J. Math. Phys.* **1**:178 (1960).
20. D. Montgomery and D. A. Tidman, *Plasma Kinetic Theory* (McGraw-Hill, New York, 1964).
21. C. J. Myerscough, in *Proc. 1969 Cornell Symp. on Kinetic Equations*, R. Liboff and N. Rostoker, eds. (Gordon and Breach, New York, 1971).