

87

87

I. INTRODUCTION

ERGODIC THEORY AND STATISTICAL MECHANICS OF NON-EQUILIBRIUM PROCESSES*

by

Joel L. Lebowitz
Belfer Graduate School of Science
Yeshiva University
New York, New York 10033

II. EQUILIBRIUM ENSEMBLES

*Lecture given at the 10th Annual Eastern Theoretical Physics Conference, 1971, Schenectady, New York, and at the 1972 Summer Institute in Non-Linear Mathematics, Battelle Memorial Institute, Seattle, Washington.

†Research supported by the U.S.A.F.O.S.R. under Contract No. F44620-71-C-0013, modification P001.

I. INTRODUCTION

Statistical Mechanics of Non-Equilibrium Processes like other physical theories has two aspects: (1) the mathematical investigation of certain well-posed, that is, mathematically formulated problems, and (2) the proper mathematical formulation of physical phenomena. It is wise, I believe, to separate these two parts of the problem and this talk is primarily about some of the progress that has been made in recent years in the mathematical theory of measure preserving transformations. Since the dynamical flow in the phase space, which describes the time evolution of a Hamiltonian system, is an example of such a transformation this work has, in my opinion, much relevance to statistical mechanics and to the question of irreversibility. The progress in this field is the result of the work of many people: Hopf, Kolmogoroff, Sinai and others, most of whose names and work I shall not have a chance to mention. (A small bibliography is given at the end).

II. EQUILIBRIUM ENSEMBLES

Let me begin by recalling briefly the situation in equilibrium statistical mechanics: There is a purely macroscopic theory, thermodynamics, which states that a great variety of the properties of a large system, of a given quantity $N \ll N = \text{particle number} \sim 10^{26}$ contained in a volume V , which is in equilibrium, are determined once its energy E is known. This information about the system is contained in a function $S(E, N, V)$, the entropy, which is extensive, that is, $S(E, N, V) = V s(E/V, N/V) \equiv V s(e, \rho)$ and which has certain convexity properties that insure thermodynamic stability. From $s(e, \rho)$ the pressure, temperature, etc., can be found. (If, instead of specifying e ,

the energy per unit volume, we specify the temperature $T = \beta^{-1}$ (in units in which Boltzmann's constant is unity) the 'thermodynamics' is determined by the Helmholtz free energy $A(\beta, N, V) = Va(\beta, \rho)$, which is related to the entropy by a Legendre transformation.)

Now, according to statistical mechanics, as developed by Boltzmann, Gibbs, Einstein and others, these thermodynamic functions for macroscopic systems are obtainable from a knowledge of the microscopic structure of the system by well known formulae. The microscopic nature of the system is specified by its Hamiltonian which we assume to be the sum of a kinetic energy term and a potential energy term

$$H_N(x) = \sum_{i=1}^N \frac{1}{2m} p_i^2 + U(r_1, \dots, r_N), \quad (1)$$

where m is the mass of a particle, r_i, p_i are the position and momentum vectors of the i th particle, $r_i \in V, p_i \in \mathbb{R}^3$ and $x = (r_1, \dots, r_N, p_1, \dots, p_N)$ is a point in the phase space Γ of $M (=6N)$ dimensions. We shall say that x specifies the 'dynamical state' of the system. To obtain the entropy $S(E, N, V)$ of this (classical) system, we define the energy surface S_E by the relation $H_N(x) = E$ for $x \in S_E$ and equate the entropy to the logarithm of its 'surface area',

$$S(E, N, V) = \ln |S_E|, \quad |S_E| = \int_{S_E} \frac{d\sigma_E(x)}{|\nabla H|}, \quad (2)$$

where $d\sigma_E$ is the $M-1$ dimensional surface area element on S_E induced by the Euclidian metric on Γ and $|\nabla H|$ is the length of the gradient of H . (Similarly, $A(\beta, N, V) \sim \ln \{ \int \exp[-\beta H_N(x)] dx \}$.)

The statistical mechanics of Gibbs does not stop at giving formulae for the thermodynamic potentials $S(E, N, V)$ (or $A(\beta, N, V)$), it

interprets these formulas in terms of 'ensembles' or probabilities: i.e., if $f(x)$ is a real valued function (dynamical observable) of the dynamical state of the system which is accessible to macroscopic measurements, the results of measuring its values in an equilibrium system with energy E will have a probability distribution obtained from a normalized probability measure, given by

$$d\mu_0(x) = \begin{cases} |S_E|^{-1} d\sigma_E(x) / |\nabla H(x)| & , \quad x \in S_E \\ 0 & , \quad x \notin S_E \end{cases} \quad (3)$$

The probability density μ_0 given in (3) (which is left invariant by the dynamical flow in Γ) is usually called the Gibbs micro-canonical ensemble density. Thus, according to statistical mechanics, if we make many observations on one system or on a collection (ensemble) of equilibrium systems with energy E (having, of course, the same Hamiltonian) then the average and mean square deviation of f are given by

$$\langle f(x) \rangle = \int f(x) d\mu_0(x) , \quad \langle [f(x) - \langle f \rangle]^2 \rangle = \int [f(x) - \langle f \rangle]^2 d\mu_0(x) \quad (4)$$

For measurements which take a certain amount of time to perform the appropriate dynamical functions are of the form

$$\bar{f}_T(x) = \frac{1}{T} \int_0^T f(x_t) dt \quad (5)$$

and

$$\langle \bar{f}_T(x) \rangle = \frac{1}{T} \int_0^T \langle f(x_t) \rangle dt = \langle f \rangle \quad (6)$$

where x_t is the point in the phase space or dynamical state of the system at time t if x is its dynamical state at $t = 0$, and in (6) we have used the fact that $d\mu_0(x)$ is time invariant, $d\mu_0(x_t) = d\mu_0(x)$. x_t is obtained from x through the solution of the Hamiltonian equations of motion

$$\dot{r}_i = \frac{\partial H}{\partial p_i} = m^{-1} p_i = v_i, \quad \dot{p}_i = - \frac{\partial H}{\partial r_i}, \quad (7)$$

together with 'boundary conditions' on the surface of the region V in \mathbb{R}^3 . (The boundaries usually taken to be reflective, i.e., the component of the velocity v_i is reversed when the i th particle hits the boundary.)

In particular, thermodynamic quantities like the pressure are equated in statistical mechanics with the ensemble average of the corresponding dynamical observables.

The two aspects of equilibrium statistical mechanics are then: (1) the investigation, for a given Hamiltonian, of the actual form of the thermodynamic functions and other expectation values, such as those which enter in the scattering of X-rays by fluids or crystals, and (2) an understanding of why the predictions of statistical mechanics work as well as they do in relating the observed properties of an equilibrium macroscopic system to the corresponding ensemble averages.

Considerable success has been achieved in the first aspect and there is no doubt at all that statistical mechanics works for macroscopic systems. The fact that the systems are macroscopic is very important here. Indeed it is only in the so-called thermodynamic limit in which the size of the system formally becomes infinite, $N \rightarrow \infty$, $V \rightarrow \infty$, $N/V \rightarrow \rho$, $E/V \rightarrow e$, that the statistical mechanically computed entropy per unit volume (or free energy per unit volume), $s(e, \rho)$ (or $a(\beta, \rho)$) have the right thermodynamic stability

properties. It is also only in this limit that the most striking aspects of equilibrium phenomena, phase transitions, have an unambiguous qualitative meaning; they correspond to (mathematical) singularities of $s(e, \rho)$ or $a(\beta, \rho)$. One can also show, in the thermodynamic limit, the equivalence of various ensembles in predicting the results of macroscopic measurements.

The justification of the use of ensembles is in a much less satisfactory state at the present time. Some of the elements entering into an explanation are: (1) the special nature of the dynamical functions which are accessible to measurement in macroscopic systems: they all have small dispersions, e.g., $\{[\langle f^2 \rangle - \langle f \rangle^2] / \langle f \rangle^2\}$ is very small, (the dispersion actually goes to zero in the thermodynamic limit when $\langle f \rangle$ can be identified with a thermodynamic quantity), and (2) the ergodic hypothesis:

$\lim_{T \rightarrow \infty} \bar{f}_T(x) \equiv \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(x_t) dt \equiv \bar{f}$ exists and is independent of x for almost all x , $x \in S_E$ and $f(x) \in L_1$, i.e., $\int |f(x)| d\mu_0(x) < \infty$.

It is easy to show that when the system is ergodic, $\bar{f} = \langle f(x) \rangle$.

Hence, if the 'effective time' T , in macroscopic observables $\bar{f}_T(x)$, were comparable with the time involved in the ergodic statement, i.e., with the T necessary for $\bar{f}_T(x)$ to become approximately equal to \bar{f} for almost all dynamical states x , we would have an "explanation". This is, however, clearly not the case since if it were true we would never observe non-equilibrium phenomena (except possibly for dynamical states x which lie on trajectories whose measure is zero in which case the ergodicity property would be irrelevant). It is, however, true that if

a macroscopic system is observed for a 'very long time', then the time average of the macroscopic observables $f(x)$ agree with the predictions of equilibrium statistical mechanics. Hence some ergodicity property for macroscopic observables seems necessary, but not sufficient to justify the Gibbs formalism.

The property of small fluctuations of macroscopic observables also seems more of a necessary than a sufficient condition for justifying the Gibbs assumptions. It helps us in understanding how a macroscopic system whose dynamical state keeps on changing all the time can have macroscopic observables $f(x)$ or $\bar{f}_T(x)$ which appear to be constant in time, i.e., the system appears to reach and remain in an equilibrium state.

The property of small fluctuations, which can actually be proven quite generally, may also have some relevance to the question of ergodicity.

If the fluctuations in f are small then $f(x) \sim \langle f \rangle$ for most $x \in S_E$ which implies that starting with a 'typical' x , $f(x_t) \sim \langle f \rangle$ for most of the time t which in turn makes it plausible that

$$T^{-1} \int_0^T f(x, t) dt \rightarrow \langle f \rangle \text{ as } T \rightarrow \infty \text{ for almost all } x, \text{ i.e., that the system}$$
 is ergodic with respect to the 'relevant' $f(x)$. (Ergodicity itself does

not however require large systems and does not imply small fluctuations.

Consider a simple one-dimensional oscillator $H = (2m)^{-1} p^2 + 1/2 m \omega^2 q^2$.

This system is ergodic on each energy surface with $\langle p^2 \rangle = mE$, $\langle p^4 \rangle = 3/2 \langle p^2 \rangle^2$.

Ergodicity is thus not incompatible with large fluctuations.)

I feel therefore that much further work is necessary to explain the applicability of the ensemble method to equilibrium phenomena. This

is even more so when we come to a discussion of non-equilibrium phenomena where we shall again adopt the ensemble, or probability, method with or without a good 'explanation'.

III. NON-EQUILIBRIUM ENSEMBLES

There are two entirely equivalent ways in which to proceed. The first approach which builds directly upon what we have already discussed for equilibrium systems is as follows: We consider what Penrose calls a 'compound observation', measuring first the value of a dynamical function $f(x)$ and then at a time t later, measuring the value of a dynamical function $g(x)$. An appropriate dynamical function for measuring the correlations in this compound observation is the explicitly time-dependent dynamical function $f(x)g(x_t)$ whose expectation value, in an equilibrium ensemble, should be given by

$$\begin{aligned}\langle g(t)f \rangle &= \int g(x_t)f(x)\mu_0(x)dx = \int g(x)[f(x_{-t})\mu_0(x_{-t})]dx_{-t} \\ &= \int g(x)f(x_{-t})\mu_0(x)dx\end{aligned}\quad (8)$$

You will recognize $\langle g(t)f \rangle$ as a time dependent correlation function in an equilibrium ensemble.

An alternative way of looking at Eq. (8) is to think of $[f(x)\mu_0(x)]$ as a non-equilibrium ensemble density at time $t = 0$, i.e., we set $\mu(x, t = 0) = K\mu_0(x)f(x)$ where K is a normalization constant $f(x)$ can 'always' be made positive by adding a suitable constant). The

ensemble density at time t , $\mu(x,t)$, then satisfies the Liouville equation

$$\frac{\partial \mu(x,t)}{\partial t} = (H, \mu) \equiv -iL\mu, \quad (9)$$

where (H, μ) is the Poisson bracket with the Hamiltonian, and L is the 'Liouville operator'. The solution of (9) is

$$\mu(x,t) = \exp[-itL]\mu(x,0) = U_{-t}\mu(x,0) = \mu(x_{-t},0) = Kf(x_{-t})\mu_0(x) \quad (10)$$

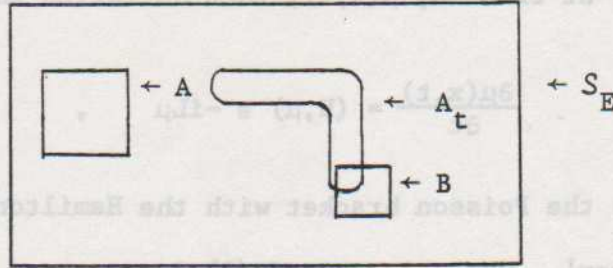
U_t is a unitary operator $U_t\varphi(x) = \varphi(x_t)$ on functions in $L_2(\Gamma)$, $\int |\varphi(x)|^2 d\mu_0 < \infty$. $\langle g(t)f \rangle$ in (8) is then the expectation value at time t of the dynamical function $K^{-1}g(x)$ in a system represented by an ensemble density $\mu(x,t)$. The interpretation of $\langle g(t)f \rangle$ in terms of non-equilibrium ensemble densities is particularly clear if $f(x)$ and $g(x)$ are characteristic functions of some sets

$$A, B \subset S_E, \quad f(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}, \quad g(x) = \begin{cases} 1, & x \in B \\ 0, & x \notin B \end{cases}, \quad (11)$$

(A and B could correspond to regions in S_E where some dynamical functions $F(x)$ and $G(x)$ have values a and b respectively, or are in the ranges (a_1, a_2) , (b_1, b_2) .) Calling $\mu_0(A)$ the measure (volume) of A with respect to the equilibrium measure $d\mu_0$

$$\mu_0(A) = \int_A d\mu_0(x), \quad \mu_0(B) = \int_B d\mu_0(x),$$

then $\mu(x,0) = f(x)\mu_0(x)/\mu_0(A)$ represents a normalized ensemble density which is concentrated 'uniformly' in A . Calling $U_t A = A_t$, with



A_t the set of points $\{x: x_{-t} \in A\}$, then $\mu(x, t) = f(x_{-t})\mu_0(x)/\mu_0(A)$ is an ensemble density concentrated uniformly in A_t . The expectation value of $g(x)$ in this non-equilibrium ensemble is then

$$\int g(x)\mu(x, t)dx = [\mu_0(A)]^{-1} \int g(x_t)f(x)d\mu_0(x) = \mu_0(A_t \cap B)/\mu_0(A), \quad (12)$$

which is simply the measure of the overlap of the set A_t with B , i.e., it is the 'fraction' of systems in the ensemble (all of which were originally in A) that are in B at time t .

A. Mixing Flows

Pursuing this analysis further, it seems reasonable to say that a necessary condition for a system to 'approach equilibrium' is that, after a 'long time', the ensemble density $\mu(x, t)$ becomes 'spread out' over the whole energy surface S_E , i.e.,

$$\frac{\mu_0(A_t \cap B)}{\mu_0(A)} \xrightarrow{t \rightarrow \infty} \mu_0(B), \text{ for all sets } A, B, \text{ with finite measure.} \quad (13a)$$

When the Hamiltonian flow has the property (13a) then the system is called mixing, a notion introduced by Hopf in 1934. The property of mixing may

also be expressed directly in terms of the time dependent correlation functions; it can be shown that a system is mixing if and only if

$$\langle g(t)f \rangle \xrightarrow{t \rightarrow \infty} \langle f \rangle \langle g \rangle, \quad (13b)$$

whenever f and g are in L_2 ,

$$\int |f|^2 d\mu_0 < \infty, \quad \int |g|^2 d\mu_0 < \infty.$$

Mixing is related in an inverse way to the notion of 'dynamical stability' of a flow. Suppose we start out with two phase points $x, y \in S_E$ and follow their trajectories x_t, y_t . We can then ask the question of 'how close' x_t and y_t will be if x and y are close, i.e., given an $\varepsilon > 0$ does there exist a $\delta > 0$ such that $|x_t - y_t| \leq \varepsilon$ for all t , if $|x - y| \leq \delta$? If the answer is yes, then the system is said to be dynamically stable. Clearly, a system which is mixing is dynamically unstable since an $M-1$ dimensional ball of radius δ centered on x will spread over the whole energy surface S_E as $t \rightarrow \infty$.

The property of mixing is stronger than, and implies, ergodicity. (It follows from (13a) that $A_t = A \Rightarrow \mu_0(A) = [\mu_0(A)]^2$ or $\mu_0(A) = 0$ or 1 which is true iff the system is ergodic.) It represents the kind of 'loss of memory' which Gibbs expected that 'coarse graining' in some way or another would bring about. Note however that a Hamiltonian system which is mixing in the forward time direction is also mixing in the backward time direction, i.e., (13) holds also with $t \rightarrow -t$. Mixing, or still stronger properties on the flow, which we shall discuss later, may thus

be necessary, but are (by themselves) certainly not sufficient to give time a direction. What these mixing properties show is that initial non-equilibrium ensemble densities of a certain type (those which are 'smooth' with respect to $d\mu_0$) will approach, in a certain well defined sense, the equilibrium ensemble density. (More on this later.)

Until recently there was no example of a dynamical system which is mixing. Recently, however, Sinai was able to prove that a system consisting of a finite number N , $N \geq 2$ hard spheres (or hard disks in two dimensions) confined to a cubical box is a mixing system. (Only part of the proof has been published so far.) It follows, therefore, in particular from Sinai's work, that for a finite system of two or more hard spheres in a box, in two or three dimensions, the velocity autocorrelation function of any particle, say particle one, approaches zero as $t \rightarrow \infty$,

$$\langle v_1(t)v_1 \rangle \xrightarrow{t \rightarrow \infty} \langle v_1 \rangle \langle v_1 \rangle = 0 \quad (14)$$

This is indeed remarkable and contrary to some folklore opinion which holds that it is necessary to go to an infinite size system in order to obtain a true decay of the correlation functions when $t \rightarrow \infty$. Here, on the other hand, this is shown to be true for a system consisting of N , $N \geq 2$, particles, as long as N is finite. (It is presumably true also for an infinite system, but this is far from proven.) The usual reason for the belief in the necessity of going to an infinite system is that for a finite system, one always has a finite Poincare recurrence time for each

dynamical state x and hence it is thought that the time correlation functions too will have such 'recurrences'. Note however that the mixing definition is meaningful only for sets of positive measure and that the equivalent decay of correlation functions definition applies only to square integrable functions (with respect to $d\mu_0$). What mixing therefore implies is that the times when different systems, which were initially together in the same set A (which can be as small as desired as long as $\mu_0(A) > 0$), return to the neighborhood of A' are so different from each other that eventually the set A_t spreads out uniformly over all of S_E . This is the important property of mixing flows which does not follow at all from ergodicity alone, and is in particularly striking contrast to what occurs in assemblies of oscillators such as harmonic crystals where all the phase points have the same periodicities. (Finite quantum systems behave like oscillators, c.f., remark at the end of Section C.)

B. Transport Coefficients

The study of time correlation functions, such as the velocity auto-correlation function in (14), plays a central role in the statistical mechanical theory of non-equilibrium phenomena. Their importance stems from the fact that linear transport coefficients, such as heat conductivity, viscosity, etc., may be expressed as integrals over time (from $t = 0$ to $t = \infty$) of the time correlation of appropriate dynamical functions (Einstein-Green-Kubo). These functions represent the 'fluxes' associated with the transport processes in question. A well known example of such a 'formula' is the Einstein relation between the self-diffusion constant D and the integral of the velocity auto-correlation function.

It might appear from (13b) that for mixing systems these transport coefficients could be defined meaningfully, without going to the thermodynamic limit of an infinite size system, as long as $[\langle g(t)f \rangle - \langle g \rangle \langle f \rangle]$ approached zero sufficiently rapidly to be integrable. Such is, unfortunately, not the case since the flux functions whose time correlations are of interest for transport coefficients can generally be written as Poisson brackets with the Hamiltonian H , i.e., $f = (F, H)$, $g = (G, H)$, and for mixing systems it can be shown that when f, g, F, G are all square integrable with respect to $d\mu_0$ then,

$$\lim_{T \rightarrow \infty} \int_0^T \langle f(t)g \rangle dt = \langle (F, H)G \rangle, \quad (15)$$

$$\lim_{T \rightarrow \infty} \int_0^T \langle f(t)f \rangle dt = 0.$$

Thus for a finite mixing system confined by a wall,

$$\lim_{T \rightarrow \infty} \int_0^T \langle v_1(t)v_1 \rangle dt = \lim_{T \rightarrow \infty} \langle q_1(T)v_1 \rangle = 0, \quad (16)$$

since

$$v_1 = (q_1, H) \quad \text{and} \quad \int v_1^2 d\mu_0 < \infty, \quad \int q_1^2 d\mu_0 < \infty.$$

Note that when q_1 is an angle variable, e.g., in the case of periodic boundary conditions, then $v_1 \neq (q_1, H)$ and (16) need not hold. We would still have, however, $\langle v_1(t)v_1 \rangle \rightarrow 0$ if the system is mixing.

When the system is not mixing, the limit $T \rightarrow \infty$ in the above integrals need not exist. It is still true however that for any finite system

$$\lim_{T \rightarrow \infty} \int \langle f(t)f \rangle dt = \lim_{T \rightarrow \infty} - \frac{d}{dT} \langle F(T)F \rangle = 0, \text{ if it exists.} \quad (17)$$

This is so since,

$$\langle F(T)F \rangle \leq \langle F(T)F(T) \rangle^{\frac{1}{2}} \langle F^2 \rangle^{\frac{1}{2}} = \langle F^2 \rangle, \quad (18)$$

so that when F is square integrable $d/dT \langle F(T)F \rangle$ can either oscillate or approach zero.

These time correlation integrals will therefore, if they exist at all, be equal to zero in any finite system. (The interesting fact is that they do exist for mixing systems). The Einstein type formulae for transport coefficients can therefore be mathematically meaningful only in the thermodynamic limit.

One of the most important problems in 'rigorous' statistical mechanics, at the present time, is therefore to investigate, and hopefully establish, the existence in the thermodynamic limit of the time integrals used in the Kubo formulae. Unfortunately, it seems impossible to even tackle this problem before one proves the existence of a time evolution in the thermodynamic limit. This has been established so far, for a general class of systems, only in one dimension (Lanford).

C. Spectrum of Liouville Operator

There is an intimate relation between the mixing properties of the flow and the spectrum of the Liouville operator L (iL is the generator of the unitary operator U_t). It can be readily shown that if the spectrum of L is absolutely continuous (except for the eigenvalue

zero) then the flow is mixing. The space on which L acts here is the Hilbert space of complex valued square integrable functions with the measure $d\mu_0(x)$.

Ergodicity, mixing and absolutely continuous spectrum are members of a hierarchy of increasingly stronger conditions on the flow (or the Hamiltonian H which generates it). It has been shown that

absolutely continuous spectrum \Rightarrow mixing \Rightarrow ergodicity, but not the converse. Mixing does however imply that L has no discrete eigenvalues other than zero which is a simple eigenvalue. Such a property of L implies in turn that the system is at least weakly mixing (the converse is also true). A weakly mixing system is one in which

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T |\mu_0(A_t \cap B) - \mu_0(A)\mu_0(B)| dt = 0 \quad (19)$$

An eigenfunction corresponding to the eigenvalue zero is of course, any constant on S_E . It follows already from ergodicity that the eigenvalue zero is simple, i.e., constants are the only eigenfunctions of L with eigenvalue zero. The converse is also true, i.e., if zero is a simple eigenvalue of L then the flow is ergodic.

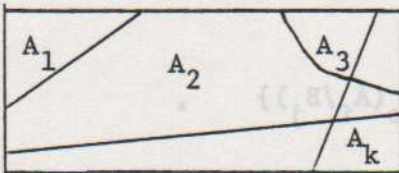
This may be a good place to note that, due to the discrete nature of the energy spectrum for finite quantum systems confined to a bounded domain V , there will be no mixing (decay of correlations) in such a system. For such quantum systems we do not therefore gain anything from the use of ensembles and we are forced to look at the infinite volume limit for signs of long time irreversibility. The remarkable thing about Sinai's

result is that it shows that finite classical systems can and do have purely continuous spectra. (Note that when Planck's constant $h \rightarrow 0$ the number of energy levels between some fixed E and $E + \Delta E$ becomes infinite.)

D. K-Systems

Sinai's method of proof that a system of hard spheres is mixing is based not on the study of the spectrum of the Liouville operator but on showing that the 'flow' of the hard sphere system on S_E is similar to the 'flow' of a free particle (geodesic flow) on a surface of negative curvature. Such flows are known to be very unstable and were shown by Sinai to be K-flows. Here K stands for Kolmogoroff who in the mid-fifties introduced the notion of a K-flow or a K-system. It can be shown that a K-system is also mixing. Indeed, K-systems seem in some way to have the right kind of 'randomness' which might lead to irreversible kinetic equations like the Boltzmann Equation or the hydrodynamic equations. I shall therefore try to explain what they are.

Imagine the energy surface S_E divided up into k disjoint cells A_i , $i = 1, \dots, k$,



$$\sum \mu_o(A_i) = 1, \mu_o(A_i \cap A_j) = 0.$$

This collection of sets $\{A_i\}$ is called a partition A , $A = \{A_i\}$; the

A_i are the 'atoms' of A . Since $\mu_o(A_i)$ is the probability (in the micro-canonical ensemble) of finding the system in A_i , Kolmogoroff defined the 'entropy' (not to be confused with the thermodynamic entropy) of this partition $h(A)$, in analogy with information theory entropy, as

$$h(A) = - \sum \mu_o(A_i) \ln \mu_o(A_i) \quad , \quad (20)$$

Clearly, $h(A) \geq 0$, with the equality holding if and only if $\mu_o(A_j) = 1$, for some j , i.e., there is complete certainty that $x \in A_j$. (We shall generally ignore sets of measure zero, setting $\mu_o(C) \ln \mu_o(C) = 0$ if $\mu_o(C) = 0$, and writing $A_j = S_E$ when $\mu_o(A_j) = 1$.) The maximum value which $h(A)$ can take is $\ln k$ corresponding to $\mu_o(A_i) = k^{-1}$ for all $i = 1, \dots, k$.

Given two partitions $A = \{A_i\}$, $i = 1, \dots, k$ and $B = \{B_j\}$, $j = 1, \dots, m$, we denote the 'sum' of the partitions A and B by $A \vee B$; $A \vee B$ is the partition whose atoms are all (non-zero measure) sets $A_i \cap B_j$. The entropy of $A \vee B$ is,

$$h(A \vee B) = \sum_{i,j} \mu_o(A_i \cap B_j) \ln \mu_o(A_i \cap B_j) \quad . \quad (21)$$

The 'conditional entropy' of a partition A , relative to a partition B is defined as

$$h(A/B) = \sum_j \mu_o(B_j) \{ \sum_i \mu_o(A_i/B_j) \ln \mu_o(A_i/B_j) \} \quad , \quad (22)$$

where

$$\mu_o(A_i/B_j) \equiv \mu_o(A_i \cap B_j) / \mu_o(B_j) \quad .$$

For a given flow operator U_t , and some fixed time interval τ , we construct the sets $U_{\tau}A_1, U_{2\tau}A_1, \dots$ and define $U_{\tau}A$ as the partition whose atoms are the sets $\{U_{\tau}A_i\}$. Kolmogoroff then sets

$$h(A, U_{\tau}) = \lim_{n \rightarrow \infty} \frac{1}{n} H\left(\bigvee_{j=0}^{n-1} U_{j\tau}A\right). \quad (23)$$

It is readily shown that $h(A, U_{j\tau}) = jh(A, U_{\tau})$. The K-S entropy of the flow U_{τ} is defined as (S for Sinai)

$$h(U_{\tau}) = \sup_A h(A, U_{\tau}) = h\tau, \quad ,$$

where h is now an intrinsic property of the flow. It was shown by Sinai that a system is a K-system iff $h(A, U_{\tau}) > 0$ for all nontrivial partitions A , i.e., for partitions whose atoms are not all of measure zero or one.

K-systems are in some sense 'random' even when the flow is entirely deterministic. To see this suppose that the atoms of the partition $A, \{A_i\}$, $i = 1, \dots, k$, correspond to different possible outcomes of the measurement of some dynamical function $f(x)$, i.e., if $x \in A_i$ then the result of the measurement will be α_i , etc. (Since the set of outcomes of the measurement is finite, being equal to k , $k < \infty$, the measurement is a 'gross' one. It need not however be restricted to measuring just one property of the system; we can replace $f(x)$ by a finite set of functions.) The probability (in the micro-canonical ensemble) of an outcome α_i is $\mu_0(A_i) \equiv p(\alpha_i)$. Now if these dynamical functions were measured first at $t = -\tau$, and then at $t = 0$, the joint probability that the result of the first measurement is α_j and the result of the second is α_i , is equal to the probability that the dynamical state of the system x at the time of the present measurement $t = 0$, is in the set $A_i \cap U_{\tau}A_j$, i.e., $p(\alpha_i, \alpha_j) = \mu_0(A_i \cap U_{\tau}A_j)$. The

conditional probability of finding the value α_i , if the result of the previous measurement was α_j , is $p(\alpha_i/\alpha_j) = \mu_o(A_i \cap U_\tau A_j) / \mu_o(U_\tau A_j) = \mu_o(A_i \cap U_\tau A_j) / \mu_o(A_j)$. In a similar way the probability of finding the result α_i at $t = 0$, given that the results of the previous measurements at times $-\tau, -2\tau, \dots, -n\tau$ were $\alpha_{i_1}, \alpha_{i_2}, \dots, \alpha_{i_n}$,

$$p(\alpha_i/\alpha_{i_1}, \dots, \alpha_{i_n}) = \mu_o(A_i \cap U_\tau A_{i_1} \dots \cap U_{n\tau} A_{i_n}) / \mu_o(U_\tau A_{i_1} \cap U_{2\tau} A_{i_2} \dots \cap U_{n\tau} A_{i_n}). \quad (24)$$

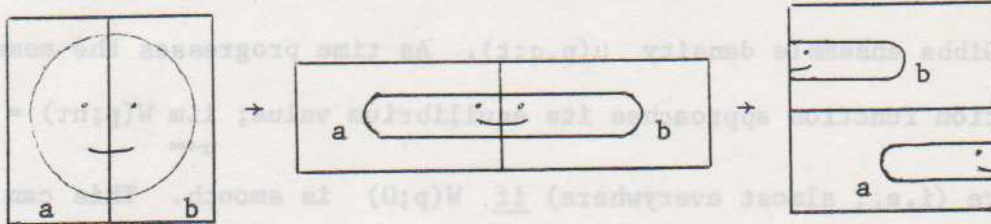
It can be readily shown that

$$h(A, U_\tau) = \lim_{n \rightarrow \infty} h(A / \bigvee_{k=1}^n U_{k\tau} A) = \lim_{n \rightarrow \infty} \{-\sum p(\alpha_{i_1}, \alpha_{i_2}, \dots, \alpha_{i_n}) \times [\sum_{i=1}^k p(\alpha_i/\alpha_{i_1}, \dots, \alpha_{i_n}) \ln p(\alpha_i/\alpha_{i_1}, \dots, \alpha_{i_n})]\} \quad (25)$$

Hence $h(A, U_\tau) > 0$ for all non-trivial partitions implies that no matter how many measurements of the values of $f(x)$ we make on a system at times, $-\tau, \dots, -n\tau$, the outcome of the next measurement is still uncertain. (N.B. the measurements are 'coarse' since $\mu_o(A_i) > 0$).

To see the kind of 'irreversibility' associated with K-systems, I shall use as an example one of the simplest kind of K-systems; the transformation of a two-dimensional square known as the baker's transformation. This is a discrete transformation, repeated at time intervals τ , on the points $x = (p, q)$ in the unit square (p and q are just labels which are meant to suggest, but have nothing to do with momentum and coordinates),

$$U_\tau x = U_\tau(p, q) = \begin{cases} (2p, \frac{1}{2}q) & \text{if } 0 \leq p < \frac{1}{2} \\ (2p-1, \frac{1}{2}(q+1)) & \text{if } \frac{1}{2} \leq p \leq 1 \end{cases} = x_\tau \quad (26a)$$



The baker's transformation

Using as our equilibrium ensemble density $d\mu_o(x) = dp dq$, $U_{-\tau}$ is a measure preserving, reversible, one to one transformation

$$U_{-\tau}x = U_{-\tau}(p, q) = \begin{cases} (\frac{1}{2}p, 2q) & , 0 \leq q < \frac{1}{2} \\ (\frac{1+p}{2}, 2q - 1) & , \frac{1}{2} \leq q \leq 1 \end{cases} = x_{-\tau} \quad (26b)$$

The K-S entropy for this transformation is $h(U_{-\tau}) = \ln 2$.

Following Penrose let us consider now a Gibbs' ensemble density which at $t = 0$, is $\mu(x, 0)$; $\int_0^1 \int_0^1 \mu(p, q; 0) dp dq = 1$. We have, as in the dynamical flow case, $\mu(x; n\tau) = \mu(x_{-n\tau}; 0)$, or

$$\mu(p, q; (n+1)\tau) = \mu(U_{-\tau}(p, q); n\tau) = \begin{cases} \mu(\frac{1}{2}p, 2q; n\tau) & , \text{ if } 0 \leq q < \frac{1}{2} \\ \mu(\frac{1+p}{2}, 2q - 1; n\tau) & , \text{ if } \frac{1}{2} \leq q \leq 1 \end{cases} \quad (27)$$

We now define a reduced ensemble density (momentum distribution function) $W(p; j\tau) \equiv \int_0^1 \mu(p, q; j\tau) dq$. It is readily verified that W satisfies the relation

$$W(p; (n+1)\tau) = \frac{1}{2} [W(\frac{1}{2}p; n\tau) + W(\frac{1+p}{2}; n\tau)] \quad (28)$$

Eq. (28) is an exact, irreversible, 'kinetic' equation, for the momentum distribution function, i.e., $W(p; t)$ determines $W(p; t')$ for

$t' > t$, but not for $t' < t$, coming from a reversible Liouville equation for the Gibbs ensemble density $\mu(p,q;t)$. As time progresses the momentum distribution function approaches its equilibrium value; $\lim_{n \rightarrow \infty} W(p;n\tau) = 1$, in measure (i.e., almost everywhere) if $W(p;0)$ is smooth. This can be seen from (28) by considering the derivative of W ,

$$\frac{d}{dp} W(p; (n+1)\tau) = \frac{1}{4} \left[\frac{d}{d\eta} W(\eta;n\tau) \Big|_{\eta=\frac{1}{2}p} + \frac{d}{d\eta} W(\eta;n\tau) \Big|_{\eta=\frac{1}{2}p+\frac{1}{2}} \right],$$

so that $\sup_p \frac{d}{dp} W(p;j\tau)$ is a monotonically decreasing function of j .

If we consider on the other hand the projection of $\mu(p,q;n\tau)$ on the q -axis, $K(q;n\tau) = \int_0^1 \mu(p,q;n\tau) dp$, then $K(q;n\tau)$ will be determined by $K(q;(n+1)\tau)$, but not by $K(q;(n-1)\tau)$ and $\sup_p dK(p;n\tau)/dp$ will be non-decreasing as $n \rightarrow \infty$. The roles of q and p are reversed if one looks at negative values of n for a given $\mu(x;0)$. It is therefore clearly necessary to choose properly the reduced description of the system if one is to get irreversible behavior. (Alternatively, one has to consider properly restricted initial distributions; Penrose, private communication.)

It should also be pointed out here that any kinetic equation for some projection of the full ensemble density which, like Eq. (28), holds for arbitrary initial densities $\mu(x;0)$, even when these and the corresponding measures $\mu(dx;0)$ are singular, must have the property of leaving singular parts of the initial reduced distribution (here $W(dp;0)$) singular. This means that it is impossible to have a kinetic equation which is valid for all initial distributions and at the same time acts (like the Boltzmann or diffusion equation) to smooth out initial singular distributions.

Acknowledgments: I would like to thank Oliver Penrose and Oscar Lanford for valuable discussions.

BIBLIOGRAPHY

1. V. I. Arnold and A. Avez, *Ergodic Problems of Classical Mechanics*, (Benjamin, 1968).
2. P. Billingsley, *Ergodic Theory and Information*, (Wiley, 1965).
3. J. L. Lebowitz, *Hamiltonian Flows and Rigorous Results in Non-Equilibrium Statistical Mechanics*, to appear in the Proceedings of I.U.P.A.P. Conference on Statistical Mechanics, Chicago, 1971.
4. D. S. Ornstein, *Measure-Preserving Transformations and Random Processes*, *American Mathematical Monthly*, V78, (1971).
5. O. Penrose, *Foundations of Statistical Mechanics*, (Pergamon, 1970).
6. Ja. Sinai, in *Statistical Mechanics Foundations and Applications*; Proceedings of the I.U.P.A.P. Meeting, Copenhagen, 1966, T. A. Bak, Editor (Benjamin, 1967), p. 559, *Russian Math. Sur.*, 25, 137 (1970).
7. A. S. Wightman, in *Statistical Mechanics at the Turn of the Decade*, E. D. Cohen, Ed. (Dekker, 1971).
8. J. Ford, *The Transition from Analytic Mechanics to Statistical Mechanics*, to appear in *Advances in Chemical Physics* (1972).