OFFPRINTS FROM LONG-TIME PREDICTION IN DYNAMICS
EDITED BY C. W. HORTON, JR., L. E. REICHL, AND A. G. SZEBEHELY

MICROSCOPIC DYNAMICS AND MACROSCOPIC LAWS

JOEL L. LEBOWITZ
Institute for Advanced Study, Princeton, New Jersey

Abstract. *Some thoughts about the relation between reversible microscopic dynamics and irreversible macroscopic laws are presented in somewhat sketchy form. It is argued that not only are they compatible but that one may (eventually) be able to derive the latter from the former. The problems and possible resolutions are illustrated (and hopefully illuminated) by means of a "gedanken-experiment" that is analyzed both heuristically and from the point of view of Lanford's rigorous derivation of the Boltzmann equation in a certain well-defined (Boltzmann–Grad) limit.*

1. INTRODUCTION

Time present and time past
Are both perhaps present in time future,
And time future contained in time past.
If all time is eternally present
All time is unredeemable.

T. S. Eliot, *Four Quartets*

I. Prigogine[1] has mentioned the perception of time in a way that recalls the opening lines of one of my favorite poems, quoted above. These lines should remind us that we do not comprehend the universe by logic alone. The poet appears to grapple with the paradox of simultaneous coexistence of past and future, as in a trajectory, despite the asymmetry of time's arrow. This *apparent* paradox is my subject (cf. Refs. 1–6).

There have been many references to coin tossings, roulette wheels, and other games of chance. It is well known that these games require differences of

Work supported in part by National Science Foundation grant PHY 78-15920.
Permanent address: Department of Mathematics and Physics
Rutgers University
New Brunswick, New Jersey

opinion—there would not be much betting at a horse race if everyone agreed on the horses. I do not know if anyone is making bets on the eventual resolution of the apparent paradoxes relating to the coexistence, in the description of the same phenomena, of both determinism and randomness, reversibility and time asymmetry, and so on. If there are people betting, however, I would be very happy to be the banker and keep the money until everyone has agreed on the answer.

Let me lay my cards on the table: I agree entirely with Prigogine's statement that it seems inconceivable that the deterministic, reversible microscopic laws (classical or quantum) do not hold for the evolution of the local density in an initially nonuniform fluid. Similarly, I too do not believe that irreversibility is due to approximations; that is, it is not just some small term missing from the diffusion equation that makes it irreversible. There remains, therefore, the question of how to derive the heat equation or the Boltzmann equation from microscopic dynamics. I say "derive" rather than "reconcile" because I do not believe that there is any contradiction, but there certainly is a need for a convincing mathematical derivation. Such a derivation would also help dispel some of the confusion surrounding the subject.

To illustrate, let us consider a somewhat idealized version of a typical time-asymmetric macroscopic event. It contains parts that may not be easy or even possible to achieve in practice (I consider an isolated system, use a classical description, etc.). The relevant question, however, is not simply what idealization can actually be carried out practically, but rather what is the right idealization for understanding rationally and being able to predict what will be observed at a particular level of precision under given circumstances. Misunderstanding of this question leads to statements like one that appeared some years ago in a respected popular magazine, namely, that Aristotle was right (hence Galileo wrong) in asserting that heavy objects fall faster than light ones under the action of gravity—just drop a feather and a penny together.

2. EXAMPLE

Figure 1 illustrates a box Λ, 10 cm on each side, divided into two equal parts connected by a channel. There are altogether $N \simeq 10^{21}$ atoms in the box, and the macroscopic experiment or observation consists in determining the numbers N_1 and N_2 of particles in the left and right parts, Λ_1 and Λ_2, $N_1 + N_2 = N$. Consider first just the left-hand side of the figure. The sequence represents qualitatively a macroscopic experiment carried out as follows. We start with the box with a plug in the hole and fill the left side Λ_1 with a gas, say helium, at room temperature and 0.01 atmospheric pressure. We then wait for a few minutes and at time $t = t_1$, the beginnings of our observations, we remove the plug. The configurations x_j are my imagined ones at the observations times t_j, $j = 1, 2, 3, 4, \ldots$. No surprise here as long as I tell you that $t_1 < t_2 < t_3 < t_4$. The system goes from a highly nonuniform density to a uniform one: precisely the kind of behavior we are used to seeing in such experiments. It is surely consistent with Hamilton's equations of motion for a system of particles interacting with a Lennard–Jones or hard sphere pair potentials. It is also true, and consistent with the microscopic equations, that the evolution of the density, kinetic energy, and other similar quantities can be described *for the times observed* very accurately by an irreversible kinetic equation (e.g., the Boltzmann equation). (If this does not seem credible, try a computer experiment.)

What *is inconsistent* is to say that the behavior of the system will be accurately described by the Boltzmann equation in *all* situations that can be imagined. Thus, if we imagine that at time t_2 we somehow put this system in the microscopic state y_2 obtained from x_2 by reversing all velocities, $y_2 = \bar{x}_2$, then for the time interval (t_2, t_3) we would not observe the density getting more uniform, as would be predicted by the Boltzmann equation. What should happen to such a system is shown on the right-hand side of Fig. 1 including, on top, the microscopic state at t_1, which would give y_2 at t_2 without further intervention. The appropriate question, then, is: Why is it typically or essentially always the case that Boltzmann's equation (or other irreversible equations) makes the right predictions? This has to do with what is going to happen (or is likely to happen) when we carry out experiments or observations of certain types on a *macroscopic* system prepared explicitly or implicitly in a certain way.

In terms of the example given above, why is it that when we observe a gas at t_2 looking as it does in the row x_2 we can quite safely predict that it will follow the course on the left rather than the one on the right? The apparent answer is that we have a prescription for constructing the pictures on the left, but we do not know how to construct the state y_2 or y_1. There are, of course, situations (e.g., the spin-echo experiments of Hahn[2]) in which a spin state "similar" to our y_2 is produced. The spin reversal is, however, considered to be something special. Also the "isolation" of the spin system is relatively low, and interactions with other degrees of freedom in the system, which have not been reversed, means that the effect is limited and of short duration. This has permitted us to accept this occasional "antiuniformization" behavior in spin systems without modifying our predictions about the course of events in general macroscopic systems such as in our example following the observation at time t_2.

Unfortunately, or fortunately, no one has succeeded in reversing all velocities in a macroscopic system, and all our experience corresponds to seeing the sequence on the left rather than the one on the right. We explain this behavior intuitively by saying that the amount of "phase space volume" consistent with a macroscopic observation of N_1 and N_2 and (energy E) increases in the left-hand sequence. But since *each* trajectory on the left has a counterpart trajectory on the right, the preparation of the system at t_1 must somehow be relevant in assigning appropriate probabilities to different configurations, and this is where we need more understanding.

3. SYSTEM ISOLATION

Our failure thus far to take into account outside perturbations makes the example of Section 2 incomplete. However, this is not very relevant for the time scales considered; that is, even if there were an isolated macroscopic system I do not believe that it would behave differently for the times considered. Certainly neither the Boltzmann nor other kinetic equations includes any terms due to walls, cosmic rays, or other outside interactions. It seems, therefore, inappropriate to invoke such outside perturbations in justifying them.

However, Poincaré's famous theorem about isolated systems should be mentioned here (see Ref. 3). The theorem states that if we surround the trajectory of an isolated Hamiltonian system by a tube of "diameter" δ then—since the energy surface has finite area—any point on the trajectory will be inside the tube infinitely often, no matter how small δ. This means if we keep *on* observing the system in Fig. 1 we will see it returning again and again to configurations "close" to x_1, to x_2, and so on. Boltzmann's equation or indeed any equation predicting an approach to uniformity without reversal cannot therefore be an even approximately valid description of an isolated system for *all* times. These "recurrence times" are, however, likely to increase very rapidly with the size of the system—being probably longer than the age of the universe for $N \simeq 10^{21}$. As Boltzmann is supposed to have told Zermelo, who raised this objection to Boltzmann's equation, "You should live so long."[4]

This extrapolation to *arbitrary times* is, therefore, invalid for isolated systems. But since no experiment lasts for a *very long* time, this need not worry us. In this case the idealization of an isolated system is not a useful one. Even small interactions with the "outside" world are likely to destroy completely this recurrence for dynamically unstable systems. Indeed, when we consider ensembles, this is not even a problem for an isolated system if it is at least mixing. For such systems ensembles can and do approach a uniform (coarse-grained) state.[5]

It should also be noted that the dynamical instability of a system's trajectory implies that if the reversal of velocities is not absolutely precise, the subsequent motion might behave very differently from the exactly reversed one. Doing a nonexact velocity reversal might then not have much observable effect on a real system. I shall come back to this point.

4. STATISTICAL MECHANICS

The discussion above is clearly far from conclusive, and to continue in this vein would be likely to add to the confusion. Therefore, we proceed to a more concrete analysis of the problem. The central question is how to describe, in a way amenable (at least in principle) to quantitative study, the observed "typical" behavior of macroscopic systems. The answer to this lies in our beloved statistical mechanics: the study of dynamics combined with probability. Here we replace the study of the microscopic trajectory of a macroscopic system (prepared initially by some macroscopic means) by the study of the time evolution of an initial ensemble. The question is now simpler: How are the appropriate ensembles to be characterized? This question is discussed in detail, although not entirely resolved, in a recent article by Penrose,[6] which I recommend highly. The article also contains a discussion of various approaches to the problem of irreversibility and a very extensive list of references.

Instead of considering the general problem, however, I discuss in more detail a further, more drastic idealization of the foregoing example for which one can actually *prove* some results. This involves consideration of a well-defined limit introduced by Harold Grad (see Ref. 7). It is the appropriate idealization of a dilute classical gas for which the Boltzmann equation ought to hold exactly and, therefore, *might* perhaps be proved rigorously. Grad's program was carried out brilliantly by Oscar Lanford[8] with certain limitations.

I now use the Lanford theorem to make the discussion above more precise, hence, I hope, more clear. The material that follows is from a joint paper with van Beijern, Lanford, and Spohn[9] in which the Lanford theorem is extended to all times for certain initial states corresponding to the motion of a test particle in a dilute gas in equilibrium. In the considerations here, however, I deal with the original theorem, as explicated in King's thesis,[10] so that to be strictly applicable the observation times t_j would have to be much sooner than is indicated in Fig. 1. (Truth is a necessary but not sufficient condition for a mathematical proof.)

5. LANFORD'S THEOREM

We consider a system of hard spheres of diameter ε and unit mass inside a box Λ. The spheres are elastically reflected among themselves and at the boundary of Λ. Let the state of the system be specified by the absolutely continuous distribution functions $\{\rho_n^\varepsilon | n > 0\}$. These satisfy the Bogolubov–Born–Green–Kirkwood–Yvon (BBGKY) equation for hard spheres.[11]

$$\frac{\partial}{\partial t}\rho_n^\varepsilon(x_1,\ldots,x_n,t) = H_n^\varepsilon \rho_n^\varepsilon(x_1,\ldots,x_n,t) + 2\sum_{j=1}^{n}\int_{R^3} dp_{n+1}\int_{S^2} d\omega\, \omega$$
$$\times\big((p_{n+1}-p_j)\rho_{n+1}^\varepsilon(x_1,\ldots,x_n)q_j+\varepsilon\omega,\, p_{n+1},\, t\big) \qquad (1)$$

Here

$$x_i = (q_i, p_i) \in \Lambda \times R^3$$

ω is a unit vector in R^3 and $d\omega$ is the surface measure of the unit sphere S^2 in three dimensions; H_n^ε describes the evolution of n hard spheres of diameter ε

inside Λ. The solutions of the BBGKY hierarchy are denoted by

$$\rho_n^\varepsilon(x_1,\ldots,x_n,t) = (V_t^\varepsilon \rho^\varepsilon)_n(x_1,\ldots,x_n) \tag{2}$$

for the initial vector of distribution functions

$$\rho^\varepsilon = (\rho_1^\varepsilon, \rho_2^\varepsilon,\ldots)$$

We want to study the low density, Boltzmann–Grad limit of the solutions of the BBGKY hierarchy. This limit is obtained by letting the fraction of volume occupied by the particles $\sim \rho\varepsilon^3$, with ρ the average density, go to zero while keeping the mean free path of the hard spheres, $\sim 1/\varepsilon^2\rho$, constant. This requires that as ε approaches zero, the density is increased as ε^{-2}. Therefore for each value of the diameter ε one chooses an initial state with distribution functions ρ_n^ε such that $\rho_n^\varepsilon \sim \varepsilon^{-2n}$. With this in mind we define the rescaled distribution functions

$$r_n^\varepsilon(x_1,\ldots,x_n) = \varepsilon^{2n}\rho_n^\varepsilon(x_1,\ldots,x_n) \tag{3}$$

Regarding the sequence $\{r_n^\varepsilon | n \geqslant 0\}$ as the vector r^ε, one can write Eq. (1) compactly as

$$\frac{d}{dt}r^\varepsilon(t) = H^\varepsilon r^\varepsilon(t) + C^\varepsilon r^\varepsilon(t) \tag{4}$$

where H^ε is a diagonal matrix with entries H_n^ε, and C^ε is a matrix with entries $C_{n,n+1}^\varepsilon$ and zero otherwise.

For $t > 0$ the time evolution of $r_n^\varepsilon(t)$ is determined by backward streaming. Therefore it seems natural to replace, for a collision, the phase point

$$(x_1,\ldots,q_j,\, p_j,\ldots,q_j + \varepsilon\omega,\, p_{n+1})$$

with outgoing momenta by the phase point

$$(x_1,\ldots,q_j,\, p_j',\ldots,q_j + \varepsilon\omega,\, p_{n+1}')$$

with incoming momenta. (These are just two different representations of the same phase point.) This leads to

$$\begin{aligned}\frac{\partial}{\partial t}r_n^\varepsilon(x_1,\ldots,x_n,t) = {}& H_n^\varepsilon r_n^\varepsilon(x_1,\ldots,x_n,t) \\ & + \sum_{j=1}^{n}\int_+ dp_{n+1}\,d\omega\;\omega\cdot(p_j - p_{n+1}) \\ & \times\big\{r_{n+1}^\varepsilon(x_1,\ldots,q_j,\,p_j',\ldots,q_j' - \varepsilon\omega,\,p_{n+1}',\,t) \\ & \quad - r_{n+1}^\varepsilon(x_1,\ldots,q_j,\,p_j,\ldots,q_j + \varepsilon\omega,\,p_{n+1},\,t)\big\}\end{aligned} \tag{5}$$

where $\int_+$ indicates that the integration over ω is restricted to the upper hemisphere $\omega \cdot (p_j - p_{n+1}) \geqslant 0$.

Formally, the limiting form of Eq. (5) that might be satisfied for $t \geqslant 0$ by the distribution functions $r(t) = \lim_{\varepsilon \to 0} r^\varepsilon(t)$ is obtained by simply setting $\varepsilon = 0$ in Eq. (5),

$$\begin{aligned}\frac{\partial}{\partial t} r_n(x_1, \ldots, x_n, t) = \sum_{j=1}^{n} p_j \frac{\partial}{\partial q_j} r_n(x_1, \ldots, x_1, t) \\ + \sum_{j=1}^{n} \int_+ dp_{n+1}\, d\omega\, \omega \cdot (p_j - p_{n+1}) \\ \times \{ r_{n+1}(x_1, \ldots, q_j, p_j', \ldots, q_j, p_{n+1}', t) \\ - r_{n+1}(x_1, \ldots, q_j, p_j, \ldots, q_j, p_{n+1}, t) \}\end{aligned} \tag{6}$$

(Implicitly, the free motion $-\Sigma_{j=1}^{n} p_j \partial / \partial q_j$ includes the specular reflection at the boundaries of Λ.)

For $t < 0$ the time evolution of $r_n^\varepsilon(t)$ is determined by forward streaming. In that case, for a collision, the phase point with incoming momenta should be replaced by the phase point with outgoing momenta. The formal limit of the resulting equation is then again Eq. (6) but with the sign of the collision term reversed.

Equation (6) for $t \geqslant 0$ (and with the sign of the collision term reversed for $t \leqslant 0$) is *called* the Boltzmann hierarchy, which can be written in the form

$$\frac{d}{dt} r(t) = Hr(t) + Cr(t) \tag{7}$$

Let $r^\varepsilon(t) \equiv V^\varepsilon(t) r^\varepsilon(0)$ and $r(t) \equiv V(t) r(0)$ be the solution of Eqs. (4) and (7), respectively, as defined, for example, by the Dyson series with $r(0) \equiv \lim_{\varepsilon \to 0} r^\varepsilon(0)$.

To prove that $r^\varepsilon(t)$ converges to $r(t)$, for $t \neq 0$, as $\varepsilon \to 0$, we need two conditions.

First, the initial distributions $r^\varepsilon(0)$ must be uniformly bounded in ε. This guarantees the uniform convergence of the Dyson series solution for some interval $|t| < t_0$. If h_β denotes the normalized Maxwellian at inverse temperature β, a suitable choice for this bound is as follows:

Condition 1. There exist a pair (z, β) such that

$$r_n^\varepsilon(x_1, \ldots, x_n) \leqslant Mz^n \prod_{j=1}^{n} h_\beta(p_j) \tag{8}$$

for all $\varepsilon < \varepsilon_0$ with a positive constant M independent of ε.

Second, $r_n^\epsilon(0)$ must converge to $r_n(0)$ in such a way that the Dyson series for $r^\epsilon(t)$ converges term by term to the series for $r(t)$. For the initial phase point $x^{(n)} = (x_1,\ldots,x_n) \in (\Lambda \times R^3)^n$, let $q_j(t, x^{(n)})$, $j = 1,\ldots,n$, be the position of the jth point particle at time t under the free motion. Then

$$\Gamma_n(t) = \left\{x^{(n)} = x_1,\ldots,x_n \in (\Lambda \times R^3)^n \,|\, q_i(s, x^{(n)}) \neq q_j(s, x^{(n)})\right\} \quad (9)$$

for $i \neq j = 1,\ldots,n$ and $-t \leqslant s \leqslant 0$ if $t \geqslant 0$, $0 \leqslant s \leqslant -t$ if $t \leqslant 0$.

In words, $\Gamma_n(t)$ is the restriction of the n-particle phase space to the set of phase points that under free backward streaming over a time t, if t is positive (or free forward streaming over a time $|t|$, if t is negative) do not lead to a collision between any pair of particles, regarded as point particles. By this restriction only a set of Lebesgue measure zero is excluded from $(\Lambda \times R^3)^n$.

Note that (i) $\Gamma_n(t)$ depends only on the free motion, (ii) $\Gamma_n(t)\ \Gamma_n(t')$ for $t' = \alpha t$, $\alpha \leqslant 1$, (iii) $\Gamma_n(t) \neq \Gamma_n(-t)$, and (iv) $x^{(n)} \in \Gamma_n(t)$ is equivalent to $\bar{x}^{(n)} \in \Gamma_n(-t)$, where $\bar{x}^{(n)} \equiv Rx^{(n)}$ is the phase point obtained from $x^{(n)}$ under the reversal $p_j \to -p_j$. In particular $\Gamma_n(t)$ is not invariant under reversal of velocities.

The suitable choice of convergence is then as follows:

Condition 2. There exists a continuous function r_n on $(\Lambda \times R^3)^n$ such that

$$\lim_{\epsilon\to 0} \epsilon^{2n}\rho_n^\epsilon = \lim_{\epsilon \to 0} r_n^\epsilon = r_n \quad (10)$$

uniformly on all compact sets of $\Gamma_n(s)$ for some $s > 0$.

Theorem (Lanford). Let $\{\rho_n^\epsilon | n \geqslant 0\}$ be a sequence of initial distribution functions of a fluid of hard spheres of diameter ϵ inside a region Λ and let the sequence $\{r_n^\epsilon | n \geqslant 0\}$ of rescaled distribution functions satisfy Conditions 1 and 2. Let $r_n^\epsilon(t)$ be the solution of the BBGKY hierarchy with initial conditions r_n^ϵ, and let $r_n(t)$ be the solution of the Boltzmann hierarchy with initial conditions r_n.

Then there exists a $t_0(z, \beta) > 0$ such that for $0 \leqslant t \leqslant t_0(z, \beta)$ the Dyson series for Eqs. (4) and (7) converge and such that $r_n(t)$ satisfies a bound of the form (Condition 1) with $z' > z$ and $\beta' < \beta$. Furthermore,

$$\lim_{\epsilon \to 0} r_n^\epsilon(t) = r_n(t) \quad (11)$$

uniformly on compact sets of $\Gamma_n(s + t)$.

For $-t_0(x, \beta) \leqslant t \leqslant 0$, Eq. (11) holds, provided in Condition 2 $s \leqslant 0$ and in the Boltzmann hierarchy the collision term $C_{n,n+1}$ is replaced by $-C_{n,n+1}$.

Remark. An interesting property of the Boltzmann hierarchy is the well-known "propagation of chaos": if the initial conditions of the Boltzmann hierarchy factorize,

$$r_n(x_1,\ldots,x_n) = \prod_{j=1}^{n} f(x_j) \tag{12}$$

then the solutions with this initial condition stay factorized,

$$r_n(x_1,\ldots,x_n,t) = \prod_{j=1}^{n} f(x_j,t) \tag{13}$$

$f(x,t)$ is the solution of the Boltzmann equation

$$\frac{\partial}{\partial t}f(q,p,t) = -p\frac{\partial}{\partial q}f(q,p,t) + \int_{+} dp_1\, d\omega\, \omega\cdot(p-p_1)$$
$$\times\{f(q,p',t)f(q,p'_1,t) - f(q,p,t)f(q,p_1,t)\} \tag{14}$$

with initial condition $f(q,p)$.

Note, however, that even when Eq. (12) is not satisfied, the solutions $r_n(t)$ of the Boltzmann hierarchy are *not* reversible, whereas the $r_n^\varepsilon(t)$ are—so what has happened?

The answer lies in the fact that the set on which the $r_n^\varepsilon(t)$ converge to $r_n(t)$ gets smaller and smaller as t increases, $\Gamma_n(s+t_2) < \Gamma_n(s+t_1)$ for $t_2 > t_1 \geqslant 0$ and $\Gamma(t) \neq \Gamma(-t)$.

When I get confused at this point (this happens at least four out of five times), the following picture is sometimes helpful.

Let $\Gamma(0) = U_{n>0}\Gamma_n(0)$ be represented schematically by the upper quadrant of the plane $x \geqslant 0$, $y \geqslant 0$. Let $\Gamma(t) = U_n\Gamma_n(t)$ correspond to the set $x \geqslant t$, $y \geqslant 0$ for $t > 0$ and to the set $x \geqslant 0$, $y \geqslant t$ for $t < 0$. Then for the case of Condition 2 satisfied on $\Gamma_n(0)$, the convergence at $t = t_2 > 0$ holds on the set $\Gamma(t_2) = \{x > t_2, y > 0\}$ but may not hold on the complementary set $\Gamma^c(t_2) = \{x, y | x < t_2, y > 0\}$. If we now reverse *all* velocities at $t = t_2$, the convergence of the *new* r_n^ε at t_2 will be on $\Gamma(-t_2)$ but will no longer hold on $\Gamma(s) = \{x, y | x > s, y > 0\}$, for *any* $s \geqslant 0$. Therefore since Condition 2 is no longer satisfied for any $s > 0$, the Lanford theorem need not hold for any time $t_2 + \tau$, $\tau > 0$.

Armed with these mathematical weapons, let us now return to our example.

6. EXAMPLE REVISITED

Let us now consider Fig. 1 from the point of view of ensembles. The initial state, at $t = 0$, corresponds to a canonical Gibbs state of N hard spheres of diameter ε all in the left-hand half-box Λ_1. It is clear that since the initial state is invariant to reversal of velocities, its distribution functions $\rho^\varepsilon = (\rho_1^\varepsilon, \varepsilon_2^\varepsilon, \ldots)$

satisfy the equality

$$V_t^\varepsilon \rho^\varepsilon = R V_{-t}^\varepsilon \rho^\varepsilon \tag{15}$$

where

$$(R\rho)_n(q_1, p_1, \ldots, q_n, p_n) = \rho_n(q_1, -p_1, \ldots, q_n, -p_n) \tag{16}$$

Furthermore,

$$V_t^\varepsilon(R V_t^\varepsilon \rho^\varepsilon) = \rho^\varepsilon \tag{17}$$

while

$$V_t^\varepsilon(V_t^\varepsilon \rho^\varepsilon) = V_{2t}^\varepsilon \rho^\varepsilon \tag{18}$$

Equation (17) states that if at time t we reverse all velocities, the system, after another time interval t, will return to its initial state in which all the particles are in Λ_1.

Consider now the sequence of initial states with distribution functions ρ^ε in which as ε approaches 0 the number of particles inside Λ_1 increases with fixed $N\varepsilon^2 = z$. Then

$$\begin{aligned}\lim_{\varepsilon\to 0} \varepsilon^{2n} \rho_n^\varepsilon(x_1, \ldots, x_n) &= \lim_{\varepsilon\to 0} r_n^\varepsilon(x_1, \ldots, x_n) = r_n(x_1, \ldots, x_n) \\ &= \prod_{j=1}^{n} \{x_{\Lambda_1}(q_j) z h_\beta(p_j)\}\end{aligned} \tag{19}$$

on $\Gamma_n(0)$, where X_{Λ_1} is the characteristic function of the set Λ_1, and, since Conditions 1 and 2 are satisfied, by Lanford's theorem

$$\lim_{\varepsilon\to 0} \varepsilon^{2n}(V_t^\varepsilon \rho^\varepsilon)_n(x_1, \ldots, x_n) = (V_t r)_n(x_1, \ldots, x_n) = \prod_{j=1}^{n} \{f(x_j, p_j, t)\} \tag{20}$$

on $\Gamma_n(t)$ for $|t| < t_0(z, \beta)$, where $f(x, t)$ is the solution of the Boltzmann equation with initial conditions $f(q, p) = X_{\Lambda_1}(q) z h_\beta(p)$.

Let us now reverse the velocities at time t, $0 < t < t_0/2$, and let us consider $R V_t^\varepsilon \rho^\varepsilon$ as the new initial state. Clearly

$$V_t(R V_t r) \neq r = \lim_{\varepsilon\to 0} V_t^\varepsilon(R V_t^\varepsilon r^\varepsilon) \tag{21}$$

according to Eq. (17), so the limiting r do not have the time reversibility of the r^ε. Indeed, the Boltzmann H-function decreases up to t, remains unchanged by R, and continues to decrease as $R V_t r$ is evolved for a time interval t.

At first sight this seems to contradict Lanford's theorem, which appears to assert that the right-hand side of Eq. (21) should indeed equal the left-hand side. There is, however, no such contradiction, for although

$$\lim_{\varepsilon\to 0} \varepsilon^{2n}(V_t^\varepsilon \rho^\varepsilon)_n = (V_t r)_n \qquad \text{on} \quad \Gamma_n(t) \tag{22}$$

we also have

$$\lim_{\varepsilon\to 0} \varepsilon^{2n}(RV_t^\varepsilon \rho^\varepsilon)_n = (RV_t r)_n \quad \text{on} \quad \Gamma_n(-t) \neq \Gamma_n(t+s), \qquad \forall s \geqslant 0 \tag{23}$$

Therefore, continuing in the same time direction as before, the reversal of velocities, $RV_t^\varepsilon\rho^\varepsilon$ no longer satisfies Condition 2 of Lanford's theorem. The theorem asserts nothing about the convergence of $\varepsilon^{2n}(V_t^\varepsilon(RV_t^\varepsilon\rho^\varepsilon))_n$ as $\varepsilon \to 0$. Of course, by Eq. (17) we can say something about this limit. The point is that we cannot conclude from Lanford's theorem that the limit is $(V_t(RV_tr))_n$, since Condition 2 is violated. For the theorem to be applicable with the initial condition at time t, one must consider either $V_t^\varepsilon(V_t^\varepsilon\rho^\varepsilon)$ or $V_{-t}^\varepsilon(RV_t^\varepsilon\rho^\varepsilon)$. In both cases the system evolves further toward equilibrium (e.g., in Fig. 1*a*, right, or Fig. 1*c*, left).

7. CONCLUDING REMARKS

1. Lanford's theorem deals with correlations that are absolutely continuous with respect to Lebesgue measure; see Eqs. (8) and (10). This singling out of Lebesgue measure, though "intuitively" very reasonable, cannot be justified on mathematical grounds alone. However, it may not be essential—there may be other physical conditions that rule out "bad" initial configurations of macroscopic systems prepared in the laboratory or found in nature.

2. The irreversible Boltzmann hierarchy is consistent with the reversible BBGKY hierarchy, since the approximation by the Boltzmann hierarchy is valid only for a particular class of initial states. Condition 2 excludes initial states such as the one just constructed by reversal of velocities. Although the question how to generally characterize good initial states for systems other than very dilute gases remains, this example does illustrate what form such an answer might take, that is, Condition 2. This is time asymmetric in just the right way[6]: the property is preserved under *forward* time evolution V_t^ε but is not invariant under R.

More precisely if Condition 2 is satisfied for some $s > 0$ (say 1 hour). the state evolves for some t, $t > 0$, Condition 2 is still satisfied on the smaller set $\Gamma(s + t)$. If we, however, do a reflection at t then Condition 2 is no longer satisfied for forward times. The same statements are true for $s < 0$, $s < t < 0$. The initial state considered in the example consists of (1) is symmetric under R and (2) has Condition 2 satisfied with $s = 0$. We can, therefore, derive either the forward or backward Boltzmann equation (both leading to identical uniformization of the density as $|t|$ increases), *but* we cannot go first in one direction and then "backtrack" by using R. It is this restriction that permits derivation of irreversible equations from reversible dynamics and symmetric initial conditions. Which way we actually go physically is determined by the fact that we *first* prepare the system (i.e., select the state), *then* observe it.

3. How would our conception of the "arrow of time" change if a method were found for actually reversing velocities in a fluid (or changing appropriate quantum phases)? It is not suggested that all the velocities in the universe be reversed, but only in systems like that of our example—something modest that would permit the right-hand side of Fig. 1 to represent the result of an actual experiment in which the system was isolated beginning with t_1. Would this be just like the spin-echo experiment,[2] which is now almost forgotten, or would this substantially change our concept of time's arrow. Put differently, do the laws of nature, as we understand them at present, exclude the possibility of ever observing the right-hand side of Fig. 1 in a real-life experiment? The instability of trajectories may be relevant here—precluding sufficiently exact

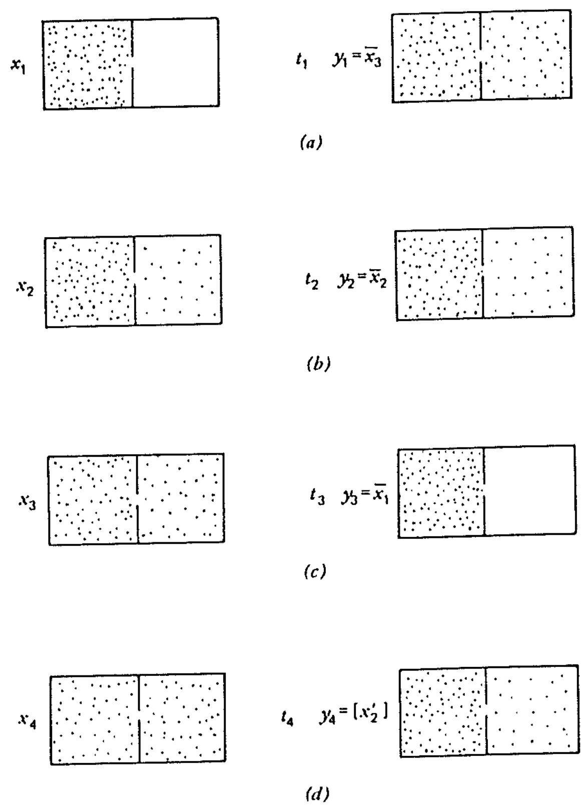


Figure 1.

reversal of velocities to give a macroscopically observable effect—like Maxwell's demon, it would cost more (in entropy) than it would gain.

This instability of trajectories relates to the good ergodic properties that physical systems are believed to possess. Thus despite having played down the role of ergodicity in providing the key to understanding irreversible macroscopic behavior, I nevertheless believe that macroscopic systems generally have good ergodic properties. Their absence would, I think, lead to effects that have not been observed.

Note however that Lanford's theorem never uses the ergodic properties of hard spheres—it is equally valid for cubes with the appropriate modification of the collision kernel. It is thus macroscopic size (remember $N \to \infty$ in the limit) that is essential for Lanford's derivation of the Boltzmann equation.

4. To emphasize the importance of the macroscopic size of the system in the observation of irreversible behavior, note that the example in Fig. 1 would not make much sense if there were only three particles in the system. Systems with few degrees of freedom can certainly exhibit instabilities in their trajectories. They can even be Bernoulli systems, like the baker's transformation, or the point particle moving among fixed convex scatterers—Sinai's billiard. In this case *certain types of initial* ensembles will behave irreversibly but a single trajectory *will not* exhibit irreversible behavior. For a macroscopic system, however, a simple trajectory can give observational results, which we would call irreversible as in the example above.

Thus the "stochastic-type" behavior of trajectories of nonlinear dynamical systems with a few degrees of freedom can serve as only one ingredient in the derivation of kinetic equations describing the time evolution of real macroscopic variables. This is true even though the study of the consequences of good ergodic properties on the behavior of measures absolutely continuous to a given stationary measure may be relevant directly to the behavior of ensembles for macroscopic systems. The situation here is similar to, but much less well understood than, the situation in equilibrium. The use of equilibrium Gibbs ensembles is formally similar for systems of few or many particles, but the relation between ensemble averages and observations is quite different in the two cases; it is only for macroscopic size systems that these can be expected to (approximately) coincide. Also, certain interesting behavior (e.g., phase transitions) shows up in large systems only.

5. Consider again the second row in Fig. 1. Suppose we measure, at time t_2, the numbers N_1, N_2 and also the total energy E of the system. We then want to *predict the future* behavior of this system without knowing "anything else" about its past or future history. Being statistical mechanicians, we would construct an ensemble to represent the initial state at t_2 and use Liouville's equation (which is equivalent to the Hamiltonian equations of motion) for the time evolution of this ensemble. It would seem appropriate to use an initial ensemble that is symmetric under velocity reversal. It is also reasonable that the ensemble be a "smooth" function on the energy surface E.

One such ensemble is $\mu(dx) = \chi(x|N_1, N_2)\,dx$, where dx is the Liouville measure projected on the energy surface S_E, $H(x) = E$ on S_E, and $\chi(x|N_1, N_2)$ is the characteristic function of the set in which there are N_1 particles on the left-hand side and N_2 particles on the right-hand side [i.e., $\chi(x|N_1, N_2)$ is 1 or 0 depending on whether the phase point is consistent with the observation]. This is the so-called generalized microcanonical ensemble, for the use of which (or of its relatives, the generalized canonical or grand canonical ensembles) many "justifications" have been given.[6] None of the arguments is entirely convincing on logical grounds *alone*—but then perhaps neither are the arguments for equilibrium Gibbs ensembles. The important question is how well this will predict the outcome of measurements. It seems clear on the basis of phase space volume arguments that the predictions would favor the left-hand sequence over the right-hand one, at least in a qualitative way. I, furthermore, think that for simple macroscopic systems (e.g., an inert fluid), the prescription would work also quantitatively, after some "short" transient time necessary for the system to establish its own quasi-steady state consistent with the macroscopic constraints. In our example this would presumably be something close to a product state with a one-particle distribution given by the Chapman–Enskog solution of the Boltzmann equation. What the appropriate quasi-steady state ensemble looks like for a more general system, even just a dense gas or an anharmonic crystal, is an open question. It is the *big* question in nonequilibrium statistical mechanics at the present time.[12]

6. Finally, the Boltzmann hierarchy, Eq. (7), does not have underlying it any flow in the phase space. Thus unlike the BBGKY hierarchy, Eq. (4), where the evolution of the states having these correlations can be implemented via the evolution of the phase points, there is no point transformation in the phase space that yields the time evolution of the correlation functions given by Eq. (7). This is yet another manifestation of the "loss of information" resulting from the use of the Boltzmann–Grad limit—a loss necessary to make dissipative macroscopic laws consistent with reversible microscopic dynamics.

ACKNOWLEDGMENTS

It is a great pleasure to thank P. G. Bergmann, S. Goldstein, O. Lanford, O. Penrose, I. Prigogine, the late P. Résibois, and H. Spohn for many useful discussion and arguments about the subject of irreversibility. A paper closely related to this work has appeared in Annals of the New York Academy of Sciences, **373**, 220 (1981).

APPENDIX

The main theme of the Workshop on Long-Time Prediction in Nonlinear Conservative Dynamical Systems appears to have been that in most cases such

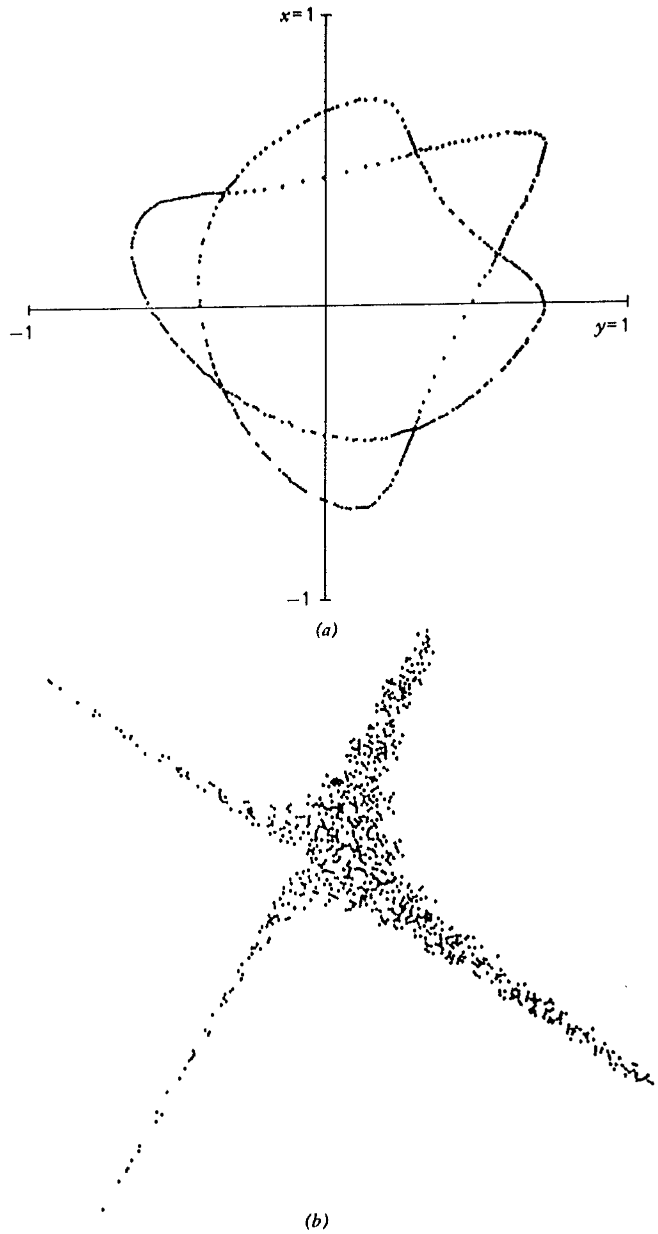


Figure A1.

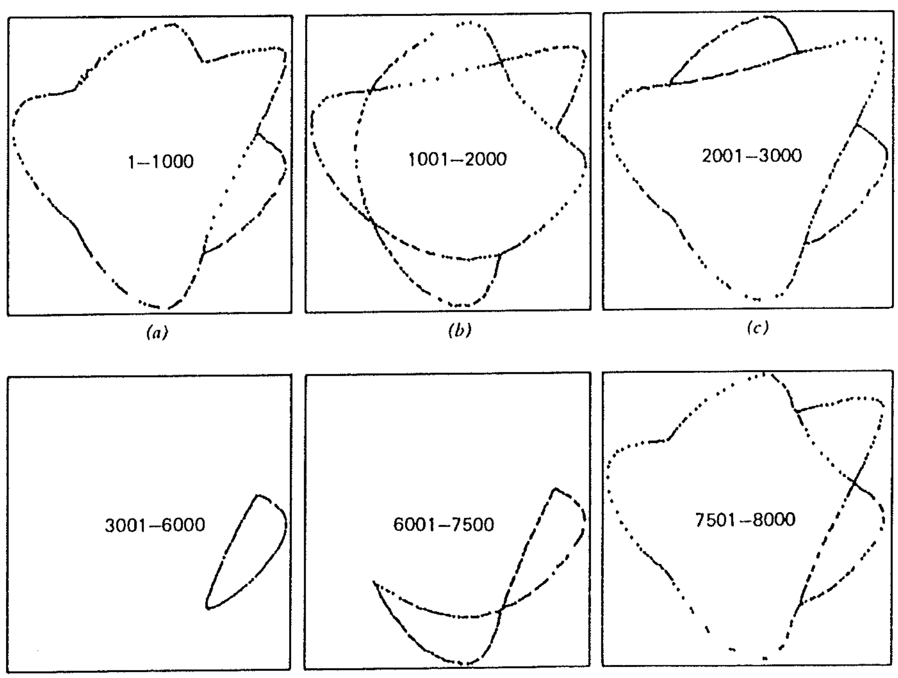
prediction is impossible. Unlike the questions about irreversibility, there is general consensus here. The reasons for, or the names given to, this phenomenon are sometimes "instabilities," at other times "sensitivity to initial conditions." Often the condition is simply referred to as "stochasticity of the motion." Therefore, I mention here briefly (or simply record in pictures) a particularly striking example of such an erratic trajectory astutely tracked by Channon[13] in his study of stochasticity in the Hénon area preserving quadratic map.[14] The still life picture of regular and stochastic trajectories of this system, familiar to statistical mechanicians, is presented again in Fig. A1. What I want to show, however, is a "moving" picture, which I think is quite impressive.

The discrete time area preserving evolution in the plane with phase point $x = (q, p)$ is $x_{n+1} = Tx_n$. It is given by

$$q_{n+1} = (\cos\alpha)q_n - (\sin\alpha)q_n + q_n^2(\sin\alpha)$$

$$p_{n+1} = (\sin\alpha)q_n + (\cos\alpha)p_n - q_n^2(\cos\alpha)$$

that is, just a rotation with some shear. Figure A2 shows the trajectory, in steps of T^5, passing through the point $q_0 = 0.718$, $p_0 = 0$ for $\cos\alpha = 0.24$, $\alpha \simeq 2\pi/5$ (the "canonical" stochastic value). The calculations were done with a precision



(a) (b) (c) (d) (e) (f)

Figure A2.

of 358 decimal digits (this provides reliability for about 8000 steps of T^5). Surely this trajectory is erratic, stochastic, and difficult to predict. More information about this exciting system is provided in the thesis of Channon.[15]

REFERENCES AND NOTES

1. I. Prigogine, lecture at The 1981 Workshop on Long-Time Predictions in Nonlinear Conservative Dynamical Systems (see chapter with B. Misra, this volume); see also *From Being to Becoming* (Freeman, San Francisco: 1980).
2. E. L. Hahn, *Phys. Rev.* **80**, 580 (1950); S. Hartmann and E. L. Hahn, *Phys. Rev.* **128**, 2024 (1962).
3. P. Ehrenfest and T. Ehrenfest, *Encykl. Mat. Wiss.* IV 2.II (Taubner, Leipzig, 1911); English translation, *The Conceptual Approach in Mechanics* (Cornell University Press, Ithaca, NY 1959).
4. For a history of the subject, see S. G. Brush, *The Kind of Motion We Call Heat*, Vols. 1 and 2 (North-Holland, New York, 1976).
5. J. L. Lebowitz and O. Penrose, *Phys. Today*, p. 23 (February 1973).
6. O. Penrose, *Rep. Prog. Phys.* **42**, 1937 (1979).
7. H. Grad, *Principles of the Kinetic Theory of Gases*, in *Handbuch der Physik*, Vol. 12, S. Flugge, Ed. (Springer-Verlag, Berlin, 1958).
8. O. E. Lanford, "Time Evolution of Large Classical Systems," in *Dynamical Systems, Theory and Applications*, J. Moser, Ed., Lecture Notes on Physics, Vol. 38 (Springer-Verlag. New York, 1975).
9. H. van Beyern, O. E. Lanford III, J. L. Lebowitz, and H. Spohn, *J. Stat. Phys.* 22, 237 (1980).
10. F. King, Ph.D. thesis, Department of Mathematics, University of California, Berkeley (1975).
11. The derivation of Eq. (1) requires some careful analysis, cf. Refs. 8 and 10.
12. Cf. J. L. Lebowitz, *Prog. Theor. Phys. Suppl.* **64**, 35 (1978).
13. S. Channon and J. L. Lebowitz, *Ann. N.Y. Acad. Sci.*, **357**, 108 (1980).
14. M. Henon, *Q. Appl. Math.* **27**, 291 (1969).
15. S. Channon, Ph.D. thesis, Rutgers University, New Brunswick, NJ, 1982.