

8

Time's Arrow and Boltzmann's Entropy

Joel L. Lebowitz

*Departments of Mathematics and Physics
Rutgers University, New Brunswick, NJ 08903, USA*

*Time past and time future
What might have been and what has been
Point to one end, which is always present.*

T.S. Eliot in Four Quartets

Dedicated to my teachers, Peter Bergmann and Melba Phillips, who taught me statistical mechanics and much, much more.

8.1 Introduction

Let me begin by declaring my premises: for the purpose of this article, my notion of time is essentially the Newtonian one – time is real and the basic laws of physics are time reversible, they connect the states of a physical system, possibly of the whole universe, at different instants of time. This of course does not take account of relativity, special or general, and is therefore certainly not the whole story. Still I believe that the phenomenon we wish to explain, namely the time asymmetric behavior of macroscopic objects, would be for all practical purposes the same in a non-relativistic universe. I will therefore focus here on idealized versions of the problem, in the simplest context, and then see how far the answers we get go towards its solution.

8.2 Acknowledgements

The analysis I present here is certainly not novel. It is based in large part on various sections in the books of R. Feynman [1], O. Penrose [2], R. Penrose [3], and D. Ruelle [4]. I have also benefitted greatly from discussions and arguments with many colleagues: those with Yakir Aharonov, Gregory Eyink, Oliver Penrose, Eugene Speer and especially Shelly Goldstein have been particularly useful to me. I have also learned much from the other participants at this conference. Since their contributions also appear in this volume I make no explicit reference to their lectures. Finally, I want to thank the organizers of the conference for a splendid meeting.

8.3 The Problem

Given these premises I will start by formulating the problem concerning the origin of the distinction, i.e. asymmetry, between past and future in a non-relativistic universe. Now this distinction is so obvious in all our immediate experiences that it is often quite hard to explain just what exactly is the problem which needs an explanation. On the other hand it is equally hard, once the question has been formulated, to answer it in a way that puts an end to the discussion once and for all. There appears to be no way to convince some people, including sometimes one's self, that the problem has really been resolved, once and for all, by Boltzmann and that there is no need to worry about it (and hold conferences about it) again and again. Let me quote Schrödinger (selectively) in one of his many discourses about this problem [5]: "the spontaneous transition from order to disorder is the quintessence of Boltzmann's theory ... This theory really grants an understanding and does not ... reason away the dissymmetry of things by means of an a priori sense of direction of time variables... No one who has once understood Boltzmann's theory will ever again have recourse to such expedients. It would be a scientific regression beside which a repudiation of Copernicus in favor of Ptolemy would seem trifling." Schrödinger continues however, "... nevertheless objections to the theory have been raised again and again in the course of the past decades and not (only) by fools but (also) by fine thinkers. If we ... eliminate the subtle misunderstandings ... we ... find ... a significant residue ... which needs exploring ..."

I will come back to the "significant residue" after I formulate the basic problem in an idealized setting: Consider an isolated macroscopic system evolving in time, as exemplified by the schematic snapshots of a gas in the four frames in Fig. 8.1. The dots in this figure represent schematically the density profile of the gas at different times during the undisturbed evolution of the system and the question is to identify the time order in which the sequence of snapshots were taken. The "obvious" answer, based on experience is: time increases from left to right - any other order is clearly impossible. Now it would be very simple and nice if this answer could be justified from the laws of nature. But this is not the case, for the laws of physics, as we know them, tell a different story: if the sequence going from left to right is a permissible one, so is the one going from right to left. This is most easily seen in classical mechanics and so I shall use this language for the present. I believe that the situation is similar in quantum mechanics and will discuss that later.

8.4 Mathematical Formulation

The complete microscopic (or micro) state of an isolated mechanical system of N particles can be represented by a point X in its phase space Γ , $X = (r_1, v_1, r_2, v_2, \dots, r_N, v_N) \in \Gamma$, r_i and v_i being the position and velocity of the i th particle. The Hamiltonian time evolution of this micro state is described by a flow T_t , i.e., as t

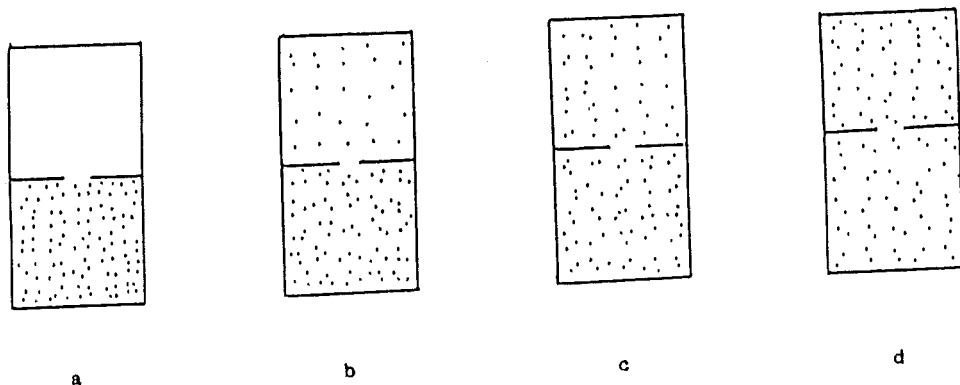


Fig. 8.1. Snapshots of macroscopic density profiles of an isolated container of gas at four different times.

varies between $-\infty$ and $+\infty$, $T_t X$ traces out the trajectory of a point $X \in \Gamma$, with $T_{t_1} T_{t_2} = T_{t_1+t_2}$ and T_0 the identity. Thus if $X(t_0)$ is the micro state at time t_0 then the state at time t_1 is given by

$$X(t_1) = T_{t_1-t_0} X(t_0).$$

Consider now the states $X(t_0)$ and $T_\tau X(t_0) = X(t_0 + \tau)$, $\tau > 0$. If we reverse (physically or mathematically) all velocities at time $t_0 + \tau$, we obtain a new microscopic state, which we denote by $RX(t_0 + \tau)$. We now follow the evolution for another interval τ to get $T_\tau R T_\tau X(t_0)$. The time reversible nature of the Hamiltonian dynamics then tells us that the new micro state at time $t_0 + 2\tau$ is just the state at $X(t_0)$ with all velocities reversed, i.e. $T_\tau R T_\tau X(t_0) = RX(t_0)$.

Let us return now to our identification of the sequence in Fig. 8.1. The snapshots clearly do not specify uniquely the microscopic state X of the system; rather they represent macroscopic states, which we denote by M . To each macro state M there corresponds a set of micro states making up a region Γ_M in the phase space Γ . Thus if we were to divide the box in Fig. 8.1 into say a million little cubes then the macro state M in each frame could simply specify the number N_j of particles in cube j , $j = 1, \dots, 10^6$. In order to make the volume of Γ_M finite let us assume that we are also given the total energy of this gas which, like the total particle number $\sum N_j = N$, does not change from frame to frame.

Clearly this specification of the macroscopic state contains some arbitrariness, but this need not concern us right now. What is important is that the snapshots shown in the figure contain no information about the velocities of the particles so that if $X \in \Gamma_M$ then also $RX \in \Gamma_M$. (The technical reader might worry that we have left out the velocity field from the macro description in Fig. 8.1. However, since the figure is meant to be solely illustrative I will continue to use only the particle density). Now we see the problem with our definite assignment of a time order to the snapshots in the figure: going from a macro state M_1 at time t_1 , to another

macro state M_2 at time $t_2 = t_1 + \tau$, $\tau > 0$, means that there is a micro state $X \in \Gamma_{M_1}$ for which $T_\tau X = Y \in \Gamma_{M_2}$, but then also $RY \in \Gamma_{M_2}$ and $T_\tau RY = RX \in \Gamma_{M_1}$. Hence the snapshots depicting M_α , $\alpha = a, b, c, d$, in Fig. 8.1 could as far as the laws of mechanics (which for the moment we take to be the laws of nature) go, correspond to time going in either direction.

It is thus clear that our judgement of the time order in Fig. 8.1 was not based on the dynamical laws of evolution alone; they permit either order. Rather it was based on experience: one direction is common and easily arranged, the other is never seen. *But why should this be so?*

8.5 Boltzmann's Answer

The answer given by Boltzmann's statistical theory starts by associating to each macroscopic state M and thus to each phase point X (through the $M(X)$ which it defines) a "Boltzmann entropy", defined (up to multiplicative and additive constants) as

$$S_B(M) = \log |\Gamma_M|$$

where $|\Gamma_M|$ is the phase space volume associated with the macro state M , i.e. $|\Gamma_M|$ is the integral of the Liouville volume element $\prod_{i=1}^N dr_i dv_i$ over Γ_M . Boltzmann's stroke of genius was to see the connection between this microscopically defined entropy $S_B(M)$ and the thermodynamic entropy S_{eq} , which is a macroscopically defined, operationally measurable (up to additive constants), extensive function of macroscopic systems in *equilibrium*. Thus when the gas in Fig. 8.1 is in equilibrium at a given energy E and volume V ,

$$S_{eq}(E, V, N) \equiv N s_{eq}(e, v) \simeq S_B(M_{eq}), \quad e = E/N, v = V/N, \quad (8.1)$$

where $M_{eq}(E, V, N)$ is the macro state observed when the system is in equilibrium at a given E and V . By \simeq we mean that for large N , when the system is really macroscopic, the equality holds up to negligible terms when both sides of (8.1) are divided by N and the additive constant, which is independent of e and v , is suitably fixed. We require here that the number of cells used to define M_{eq} should grow more slowly than N . For a lucid discussion of this point see Chapter V in Oliver Penrose's book [2].

Boltzmann's great insight at once gave a microscopic interpretation of the mysterious thermodynamic entropy of Clausius as well as a natural generalization of entropy to nonequilibrium macro states M . Even more important, it gave a plausible explanation of the origin of the second law of thermodynamics – the formal (if restricted) expression of the time asymmetric evolution of macroscopic states occurring in nature. It is certainly reasonable to expect that when a macroscopic constraint is lifted in a system in equilibrium, as when a seal which was confining the

gas to half the box in Fig. 8.1 is removed, the dynamical motion of the microscopic phase point will "more likely" wander into newly opened regions of Γ for which Γ_M is large than into those for which Γ_M is small. Thus if we monitor the time evolution of the macro state $M(t)$ (short for $M(X(t))$) we expect it to change in such a way that $S_B(M(t))$ will "typically" increase as time increases.

In particular the new macroscopic equilibrium state M_{eq} which will be reached by the system as it evolves under the new, less constrained, Hamiltonian in the bigger box can then be "expected" to be one for which Γ_M has the largest phase space volume, i.e. $S_B(M_{eq}) \geq S_B(M)$ for all M consistent with the remaining constraints. Note that when we take the system in Fig. 8.1 to be macroscopic, say one mole of gas in a one liter container, the ratio of $\Gamma_{M_{eq}}$ of the unconstrained system and the one constrained to the bottom half of the container (roughly $|\Gamma_{M_d}|/|\Gamma_{M_a}|$) is of order $10^{10^{20}}$. We thus have "almost" a mechanical derivation of the second law.

The eventual attainment of a macroscopic equilibrium state of the gas in the larger box is expressed by the zeroth law of thermodynamics: it certainly depends on the microscopic dynamics having some reasonable ergodic behavior. The precise requirements are unknown but are probably very mild since we are interested in systems with very large N . The latter is required in any case for the whole business of thermodynamic equilibrium to make sense (see discussion later) and should be generally sufficient for producing adequate ergodic behavior of real systems, including deterministic chaos with its attendant sensitive dependence on initial conditions and small perturbations. We shall assume this to be the case and not discuss it further. Very large N also takes care of the objection against Boltzmann involving Poincaré recurrence times.

8.6 Mathematical Elaboration

Boltzmann's ideas are, as Ruelle [4] says, at the same time simple and rather subtle. They introduce notions of probability as indicated by the use of such words as "likely", "expected", etc, into the "laws of nature" which, certainly at that time, were quite alien to the scientific outlook. Physical laws were supposed to hold without any exceptions, not just almost always. Thus it is no wonder that many well known objections were raised against Boltzmann's ideas by his contemporaries (Loschmidt, Zermelo, ..., see S. Brush [6] for a historical account) and that, as Schrödinger wrote in 1954, objections have continued to be raised in the past decades. It appears to me, in 1992, that although ideas of probability, unpredictability and chaos are now part of the general scientific and even popular world outlook, Boltzmann's seminal ideas are still not universally accepted or understood. Let me try to explain how I understand them. To do this I will formulate a bit more precisely the nature of the probability distribution to which the notions of "likely", "expected", etc. used in the previous paragraph refer. (One can actually make these ideas very precise

mathematically but that does not necessarily illuminate the physical reality – in fact it can obscure it.)

Let us first consider the gas in Fig. 8.1a to be in equilibrium in the bottom half of the box – excluded from the other half by a wall. Observations and analysis show that the exact values of certain types of phase functions $f(x)$ such as the number of particles contained in some subset of the volume, their kinetic energy in that region, or the force exerted by the gas on some region of the wall, will fluctuate in time about some stationary mean value. The single and multi-time statistics of such observations (obtained by independent repetitions of a specified experiment or situation) will be stationary in time – that is more or less what is meant by the system being in equilibrium. Furthermore the relative magnitude of these fluctuations will decrease as the size of the region increases. A quantitative theory of this behavior – both averages and fluctuations, including time statistics – can be obtained by the use of the Gibbs microcanonical ensemble [2]. This ensemble assigns probabilities to finding the micro state X in a phase space region Δ , *consistent with the specified constraints*, proportional to the volume of Δ .

These probabilities can be interpreted either subjectively or as a statement about empirical statistics. Whatever the interpretation it is important that such ensembles are, by Liouville's theorem, time invariant under the dynamics. For our purposes we shall regard these probabilities as representing the fraction of time (over a sufficiently long time period) which the system spends in Δ since, for the reasons discussed earlier, we can assume that macroscopic systems are effectively ergodic. For such systems the microcanonical ensemble is the only stationary measure for which the probability density is absolutely continuous with respect to the projection of Liouville measure on the energy surface in Γ [7]. (The central role played by Liouville measure is to be noted but, aside from its intuitive "obviousness" and experimental validation, it is only partially justified at present [even assuming ergodicity]. Perhaps the best justification is that in the classical limit of quantum statistical mechanics equal Liouville volume corresponds to equal weight for each quantum state. There are also some interesting stability arguments for singling out these Gibbs ensembles.)

The microcanonical ensemble thus provides a quantitative measure for the fraction of time a typical equilibrium trajectory will spend in untypical regions of the phase space, e.g. in regions where $S_B(M)$ differs significantly from its maximal value. The fractions of time S_B will be decreasing or increasing are the same – in fact the behavior of $M(t)$ is entirely symmetric around local minima of $S_B(t)$, c.f. [8], p. 249.

Suppose now that we compute the time evolution of a microscopic phase point, which is "typical" of such a microcanonical ensemble, when the constraining partition is removed. Then it can be proven in some model systems, and is "believed" to be true for systems with realistic potentials obeying Newtonian laws, that the time evolution of the momentum and energy density will be described to a "high degree of accuracy" by the appropriate time asymmetric macroscopic equations, e.g.

Navier-Stokes type equations of hydrodynamics [8,9]. A particular consequence of these equations is a detailed prediction of how $S_B(M(t))$ increases monotonically in time. This is of course more than is needed for just having S_B increase but I want to discuss it a bit here because it fleshes out Boltzmann's statistical ideas for more general systems than the dilute gases which can be treated by Boltzmann's kinetic equation.

The requirement that we start with an equilibrium state is actually too restrictive. We can also start with a nonuniform macroscopic density profile, such as the state M_b given in Fig. 8.1, and consider micro states typical of local equilibrium type ensembles consistent with M_b . We then find again evolution towards states like M_c for subsequent times.

In the above statement the words in quotes have the following meanings: "typical" behavior is that which occurs with large probability with respect to the given initial ensemble, i.e. the set of points X in the ensemble for which the statement is true comprise a region of the energy surface whose volume fraction is very close to one, for N very large; "believed" means that the basic ingredients of a mathematical proof are understood but an actual derivation is too difficult for our current mathematical technology; "high degree of accuracy" means that the hydrodynamic equations become exact when the ratio of microscopic to macroscopic scales goes to zero – that is in the so called hydrodynamic scaling limit, c.f. [8,9].

The main ingredient in this analysis is first and foremost the very large number of microscopic events contributing to the macroscopic evolution. Since the direct influence between the particles of the system takes place on a microscopic scale the macroscopic events satisfy, for realistic interactions (and here is where the gap in our mathematics is greatest), a "law of large numbers" which means that there is very little dispersion about specified deterministic behavior, i.e. that we are able to derive macroscopic laws not just for averages over ensembles but for individual systems – with probability approaching one when the micro-macro scale separation becomes large.

8.7 Typical versus Averaged Behavior

Having results for typical micro states rather than averages is not just a mathematical nicety but goes to the heart of the problem of deriving observed macroscopic behavior – we do not have ensembles when we carry out observations like those illustrated in Fig. 8.1. What we need and can expect to have is typical behavior. This also relates to the distinction (unfortunately frequently overlooked or misunderstood) between irreversible and chaotic behavior of Hamiltonian systems. The latter, which can be observed in systems consisting of only a few particles, will not have a unidirectional time behavior in any particular realization. Thus if we had only a few hard spheres in the box of Fig. 8.1, we would get plenty of chaotic dynamics and very good ergodic behavior (mixing, K-system, Bernoulli) but, we could not tell the

time order of any sequence of snapshots. To summarize: when a constraint is lifted from a macroscopic system in equilibrium at some time t_0 then in the overwhelming majority of cases, i.e. with probability approaching one with respect to the microcanonical ensemble, the micro state $X(t_0)$ will be such that the subsequent evolution of $M(t)$ will be governed by irreversible macroscopic laws.

We may regard the above (I certainly do) as the mathematical elaboration (and at least partial proof) of Boltzmann's original ideas that the observed behavior of macroscopic systems can be understood by combining dynamics with phase-space volume considerations.

This may be a good point to compare Boltzmann's entropy – defined for a micro state X of a macroscopic system – with the more usual entropy S_G of Gibbs, defined for an ensemble density $\rho(X)$ by

$$S_G(\{\rho\}) = - \int \rho(X) [\log \rho(X)] dX.$$

If we now take $\rho(X)$ to be the generalized microcanonical ensemble associated with a macro state M ,

$$\rho_M(X) \equiv \begin{cases} |\Gamma_M|^{-1}, & \text{if } X \in \Gamma \\ 0, & \text{otherwise} \end{cases}$$

then clearly,

$$S_G(\{\rho_M\}) = \log |\Gamma_M| = S_B(M).$$

It is a consequence of this equality that the two entropies agree with each other (and with the macroscopic thermodynamic entropy) for systems in local equilibrium, up to negligible terms in system size.

It is important to note however that the time evolutions of S_B and S_G are very different. As is well known, $S_G(\{\rho\})$ does not change in time when ρ evolves according to the T_t evolution, while $S_B(M)$ certainly does. In particular even if we start a system in a state of local thermal equilibrium, such as M_a in Fig. 8.1, S_G would equal S_B only at that initial time. Subsequently S_B would typically increase while S_G would not change with time. S_G would therefore not give any indication that the system is evolving towards equilibrium. This is connected with the fact discussed earlier that the micro state of the system does not remain typical of the local equilibrium state as it evolves under T_t . Clearly the relevant entropy for understanding the time evolution of macro systems is S_B and not S_G . Unfortunately this point is often missed in many discussions and leads to unnecessary confusion. The use of S_G in nonequilibrium situations is often a convenient technical tool but is not related directly to the behavior of an individual macroscopic system.

8.8 Irreversibility and Macroscopic Stability

Coming back now to the time ordering of the macro states in Fig. 8.1 we would say that the sequence going from left to right is typical for a phase point in Γ_{M_a} .

The sequence goes untypical for a phase point in Γ_{M_b} representing the time ordering of the state M_b is that of a "typical" time ordering of its state, by velocity increases and it seems effectively following observations the state M_b is typical of Γ_{M_b} is concerned by RY is concerned $t > 0$ but unstable equations describe are stable. In evolution is chaotic have increasing evolution of the perturbed in "chaotic" and backward one.

Mechanics its state, by velocity increases and it seems effectively following observations the state M_b is typical of Γ_{M_b} is concerned by RY is concerned $t > 0$ but unstable equations describe are stable. In evolution is chaotic have increasing evolution of the perturbed in "chaotic" and backward one.

This behavior of Y will tend to the unperturbed of RY is towards somewhat and to get a small right to get it macroscopic comparable to

The difference entropy-increase to understand the case in Γ is stable this predictions our ability type of velocity

The sequence going from right to left on the other hand while possible is highly untypical for a phase point in Γ_{M_a} . The same is true if we compare *any pair* of the *macro states* in that sequence: left to right is typical, right to left atypical for the Γ_M representing the *initial state*. Experience tells us that there is no "conspiracy" – if we do something to a macroscopic system and then leave it isolated, its future behavior is that of a "typical" phase point in the appropriate Γ_M . This then determines our time ordering of the snapshots.

Mechanics itself doesn't of course rule out deliberately creating an initial micro state, by velocity reversal or otherwise, for which $S_B(t)$ would be decreasing as t increases and thus make the sequence in Fig. 8.1 go from right to left – it just seems effectively impossible to do so in practice. This is presumably related to the following observations: the macroscopic behavior of a system with micro state Y in the state M_b coming from M_a which is typical with respect to Γ_{M_a} , i.e. such that $T_t Y$ is typical of Γ_{M_a} , $S_B(M_a) < S_B(M_b)$, is *stable* against perturbations as far as its future is concerned but very *unstable* as far as its *past* (and thus of the future behavior of RY) is concerned [10]. That is the *macro state* corresponding to $T_t Y$ is stable for $t > 0$ but unstable for $t < 0$. (I am thinking here primarily of situations where the equations describing the macroscopic evolution, e.g. the Navier-Stokes equations, are stable. In situations, such as the weather, where the forward macroscopic evolution is chaotic, i.e. sensitive to small perturbations, [4], all evolutions will still have increasing Boltzmann entropies in the forward direction. For the backward evolution of the micro states however the unperturbed one has decreasing S_B while the perturbed ones have [at least after a very short time] increasing S_B . So even in "chaotic" regimes the forward evolution of M is much more stable than the backward one.)

This behavior can be understood intuitively by noting that a random perturbation of Y will tend to make the micro state more typical and hence will not interfere with the unperturbed behavior of increasing S_B for all $t > 0$ while the forward evolutions of RY is towards smaller phase space volume which requires "perfect aiming". It is somewhat analogous to those pinball machine type puzzles where one is supposed to get a small metal ball into a particular small region. You have to do things just right to get it in but almost anything you do gets it out into larger regions. For the macroscopic systems we are considering the disparity between relative sizes of the comparable regions in the phase space is unimaginably larger.

The difference between the stability of the macroscopic evolution in the forward, entropy-increasing, direction and its instability in the reverse direction is very relevant to understanding the behavior of systems which are not completely isolated – as is the case in practice with all physical systems. In the direction in which the motion is stable this lack of complete isolation interferes very little with our ability to make predictions about macroscopic behavior. It however almost completely hampers our ability to actually observe "back motion" following the application of some type of velocity reversal as in the case of spin echo experiments. After a *very short*

straint is lifted
overwhelming
to the micro-
evolution

rotation (and
behavior of
phase-space

and for a micro
Gibbs, defined

associated with

each other (and
equilibrium, up

and S_G are very
when ρ evolves
similar even if we
in Fig. 8.1, S_G
typically increase
any indication
with the fact
typical of the
entropy for
Unfortunately
nary confusion.
technical tool but
system.

8.1 we would
point in Γ_{M_a} .

time in which S_B decreases the outside perturbations will make it increase again [4]. The same happens also in computer simulations where velocity reversal is easy to accomplish but where roundoff error plays the role of outside perturbations.

8.9 Remaining Problems

8.9.1 Significant Residue

I now turn to the "significant residue" of Schrödinger. As for the "subtle misunderstandings" I can only hope that they will be taken care of. The point is that when we consider a local equilibrium corresponding to a macroscopic state like M_b , and compute, via Newton's equations, the antecedent macro state of a typical micro state $X \in \Gamma_{M_b}$, we also get a macro state like M_c and not anything resembling M_a . This is of course obvious and inevitable: since the local equilibrium ensemble corresponding to the macro state M_b , at some time t_o , gives equal weight to micro states X and RX it must make the same prediction for $t = t_o - \tau$ as for $t = t_o + \tau$. (The situation would not be essentially changed if our macro state also included a macroscopic velocity field.)

We are thus apparently back to something akin to our old problem: Why can we use statistical arguments based on phase space volume (e.g. local equilibrium type ensemble) considerations to make predictions about the future behavior of macroscopic systems but not to make retrodictions? Now in the example of Fig. 8.1 if indeed the macro state M_b came from M_a , and we take its micro-state at that earlier time to be typical of equilibrium with a constraining wall, i.e. of Γ_{M_a} , then its micro state corresponding to M_b is *untypical* of points in Γ_{M_b} : by Liouville's theorem the set of all such phase points has at most volume $|\Gamma_{M_a}|$ which is much smaller than $|\Gamma_{M_b}|$. Nevertheless its *future* but not its *past* behavior, as far as macro states are concerned, will be similar to that of typical points taken from Γ_{M_b} . It is for this reason that we can use autonomous equations, like the diffusion equation, to predict *future* behavior of real macroscopic systems without worrying about whether their micro states are typical for their macro states. They will almost certainly not be so after the system has been isolated for some time - although in the real world the inevitable small outside perturbations might in fact push the system towards typicality - certainly if we wait long enough, i.e. we are in an equilibrium macro state.

The above analysis thus explains why, if shown only the two snapshots M_b and M_c and told that the system was isolated for some time interval which included the time between the two observations, our ordering would be M_b before M_c and not vice versa. This would in fact be based on there being an initial state like M_a , with even lower entropy than M_b , for which the micro state was typical. From such an initial state we get a monotone behavior of $S_B(t)$ with the time ordering M_a , M_b and M_c . If on the other hand we *knew* that the system in Fig. 8.1 had been "completely"

isolated for a very long time, compared to the hydrodynamic relaxation time of the system before the snapshots in Fig. 8.1 were taken then (in this *very very very* unlikely case) we would have no basis for assigning an order to the sequence since, as already mentioned, fluctuations from equilibrium are typically symmetric about times in which there is a local minimum of S_B . In the absence of any knowledge about the history of the system before and after the sequence we use our experience to deduce that the low entropy state M_a was the initial prepared state [11].

The origin of low entropy initial states poses no problem in "laboratory situations" such as the one depicted in Fig. 8.1. In such cases systems are prepared in states of low Boltzmann entropy by "experimentalists" who are themselves in low entropy states. Like other living beings they are born in such states and maintained there by eating low entropy foods which in turn are produced by plants using low entropy radiation coming from the sun, etc., etc. But what about events in which there is no human participation, e.g. if instead of Fig. 8.1 we are given snapshots of a meteor and the moon before and after their collision? Surely the time direction is just as obvious as in Fig. 8.1.

To answer this question along the Boltzmann chain of reasoning leads more or less inevitably (depending on considerations outside our domain of discourse) to a consistent picture with an initial "state of the universe" having a very small value of its Boltzmann entropy, i.e. an initial macro state M_o for which $|\Gamma_{M_o}|$ is a very small fraction of the "total available" phase space volume. Roger Penrose, in his excellent chapter on the subject of time asymmetry [3], takes that initial state, the macro state of the universe just after the "big bang", to be one in which the energy density is uniform. He then estimates that $|\Gamma_{M_o}|/|\Gamma_{M_f}| \sim 10^{-10^{123}}$, where M_f is in the state of the "final" crunch, with $|\Gamma_{M_f}| \sim$ total available volume. This is a sufficiently small number (in fact much smaller than necessary) to produce all we observe. The initial "micro state of the universe" can then be taken to be typical of Γ_{M_o} .

In R. Penrose's analysis the low value of $S_B(M_o)$, for a universe with a uniform density, compared to $S_B(M_f)$ is due to the vast amount of the phase space corresponding to macro states with black holes, in which the gravitational energy is very negative. I do not claim to understand the technical aspects of this estimate, which involves the Bekenstein-Hawking formula for the entropy of a black hole; it certainly goes beyond the realm of classical mechanics being considered here. The general idea, however, that the gravitational energy, which scales like N^2 for a star or galaxy, can overwhelm any non-gravitational terms, which scale like N , seems intuitively clear.

8.10 The Cosmological Initial State Problem

I hope that I have convinced you that, as Schrödinger says, "Boltzmann's theory ... really grants an understanding ...". It certainly gives a plausible and consistent picture of the evolution of the universe following some initial low entropy state M_o .

The question of how M_o came about is of course beyond my task (or ability) to answer. That would be, as Hawking puts it, "knowing the mind of God" [12]. Still, as R. Penrose has pointed out, it would be nice to have a theory which would force, or at least make plausible, an initial M_o so special that its phase space volume $|\Gamma_{M_o}|$ is infinitesimally small compared to the proverbial needle in the haystack, see Fig. 7.19 in [3]. He and others have searched, and continue to do so, for such a theory. While these theories properly belong to the, for me, esoteric domain of quantum cosmology there is, from a purely statistical mechanical or Boltzmannian point of view, a naturalness to a spatially homogeneous initial state M_o . Such an M_o would indeed be an equilibrium state in the absence of gravity. It is therefore tempting to speculate that "creation" or the big bang was "just" the turning on of gravity, but I am told by the more knowledgeable that this is quite unreasonable. The initial state problem is thus very much open. It is by far the oldest open problem.

Within the context of special (or singular) origin theories of which the big bang is a special example, widely accepted as the truth, there is nothing, not even time, before the initial state. There is an alternate suggestion, dating back to much before the advent of black holes or the big bang theory, in which one doesn't have to assume a special singular creation. Boltzmann speculated that a low entropy "initial state" may have arisen naturally as a fluctuation from an "equilibrium universe." This is in some ways a very appealing minimal hypothesis requiring no beginning or end or special creation. All you have to do is wait "long enough" and you will get any state you want, assuming that a microcanonical ensemble and some mild form of ergodicity exist for the universe as a whole. This requires, at the minimum, some short range regularization of gravity. We shall not worry however about such "technical details" since, as we shall argue next, such a hypothesis is very implausible for other entirely conceptual reasons.

While the obvious objection to this hypothesis, that such a fluctuation is enormously unlikely, can be countered by the argument that if indeed the history of the microstate of the universe is typical of trajectories in Γ then, without waiting for some huge fluctuation, we humans would not be here to discuss this problem, there remains a more serious objection. As pointed out by Schrödinger and others and particularly by Feynman [1], the actual "size" of the observed ordered universe is too large by orders and orders of magnitude for what is needed. A fluctuation producing a "universe" the size of our galaxy would seem to be sufficient for us to be around. In fact using *purely* phase space volume arguments the "most likely" fluctuation scenario of how I come to be here to discuss this problem is one where only "I" or even only my consciousness really exists, i.e. one in which the smallest region possible is out of equilibrium – and this happened just this instant. While irrefutable as an academic debating position this is, of course, even more in conflict with our observed macro state (e.g. our memories). Merely accepting that what we observe and deduce logically from our marvelous scientific instruments about the world is really there, the idea of a recent fluctuation seems "ridiculous" and therefore

makes the whole fluctuation from equilibrium scenaria seem highly implausible. In fact Feynman after discussing the problem in some detail concludes (in the tape of his lecture) that "...it is necessary to add to the physical laws the hypothesis that in the past the universe was more ordered, in the technical sense (smaller S_B), than it is today – to make sense, and to make an understanding of the irreversibility" [1].

I should say however that, even after rejecting the "fluctuation from equilibrium" scenario, the evidence or argument present for any particular, minimal S_B initial state of the universe, is not entirely without difficulty. Let me present the problem in the form of a question: given that $|\Gamma_{M_0}|$ is so extraordinarily small compared to the available $|\Gamma| \sim |\Gamma_{M_f}|$ and hence that *every* point in Γ_{M_0} is atypical of Γ , how can we rule out an initial micro state which is itself atypical of Γ_{M_0} – *whatever* the original M_0 was? This could correspond to a scenario in which S_B first decreased, reaching a minimum value at some long ago time. Since we do not assume this scenario that the trajectory of the universe is typical of the whole Γ – in fact we permit a singular initial condition as in big bang theories – some of the objections to a long ago fluctuation from equilibrium scenario are not so telling. Also in this type of scenario we need not assume any symmetry about the minimum entropy state.

The alternatives come down to this: if we accept that the Boltzmann entropy was minimal for the initial state M_0 , then the initial micro state can be assumed to be typical of Γ_{M_0} , while in a universe in which S_B first decreased and only then increased, the initial micro state would have to be *atypical* with respect to M_0 . It seems to me that there is a strong rationale for not accepting such an additional improbable beginning without being forced to it by some observational considerations. This seems to be the point which Schrödinger tried to illustrate with his prisoner story [13]. The concluding moral of that story in which the poor prisoner, who has (very probably) missed his chance for freedom by not being willing to trust probabilities in his favor after realizing that the initial state he was dealing with had to be an unlikely one, is "Never be afraid of dangers that have gone by! It is those ahead which matter."

8.11 Quantum Mechanics

The analysis given above in terms of classical mechanics can be rephrased, formally at least, in terms of quantum mechanics. We make the following correspondences:

- (i) micro state $X \Leftrightarrow$ wave function $\psi(r_1, \dots, r_N)$
- (ii) time evolution $T_t X \Leftrightarrow$ unitary Schrödinger evolution $U_t \psi$
- (iii) velocity reversal $RX \Leftrightarrow$ complex conjugation $\bar{\psi}$
- (iv) phase space volume of macro state $|\Gamma_M| \Leftrightarrow$ dimension of projector on macro state M .

This correspondence clearly preserves the time symmetry of classical mechanics. It does not however take into account the non-unitary or "wave function collapse" (measurement) part of quantum mechanics, which on the face of it appears time asymmetric. In fact this theory "is concerned exclusively with the prediction of probabilities of specific outcomes of future measurements on the basis of the results of earlier observations. Indeed the reduction of the wave packet has as its operational contents nothing but this probabilistic connection between successive observations." The above quote is taken from an old article by Aharonov, Bergmann and Lebowitz (ABL) [14] which to me still seems reasonable now. In fact I will now quote the whole abstract of that article:

"We examine the assertion that the "reduction of the wave packet," implicit in the quantum theory of measurement introduces into the foundations of quantum physics a time-asymmetric element, which in turn leads to irreversibility. We argue that this time asymmetry is actually related to the manner in which statistical ensembles are constructed. If we construct an ensemble time symmetrically by using both initial and final states of the system to delimit the sample, then the resulting probability distribution turns out to be time-symmetric as well. The conventional expressions for prediction as well as those for "retrodiction" may be recovered from the time-symmetric expressions formally by separating the final (or the initial) selection procedure from the measurements under consideration by sequences of "coherence destroying" manipulations. We can proceed from this situation, which resembles prediction, to true prediction (which does not involve any postselection) by adding to the time-symmetric theory a postulate which asserts that ensembles with unambiguous probability distributions may be constructed on the basis of preselection only. If, as we believe, the validity of this postulate and the falsity of its time reverse result from the macroscopic irreversibility of our universe as a whole, then the basic laws of quantum physics, including those referring to measurements, are as completely time symmetric as the laws of classical physics. As a by-product of our analysis, we also find that during the time interval between two noncommuting observations, we may assign to a system the quantum state corresponding to the observation that follows with as much justification as we assign, ordinarily, the state corresponding to the preceding measurement."

I interpret the ABL analysis as showing that one can conceptually and usefully separate the measurement formalism of conventional quantum theory into two parts, a time symmetric part and a second-law type asymmetric part - which can be traced back, using Boltzmann type reasoning, to the initial low entropy state of the universe. (Of course it is not clear how to discuss meaningfully the concept of measurement in the context of the evolution of the universe as a whole.)

I believe that my colleagues agree with this interpretation of our work. Aharonov in particular has emphasized and developed further the idea described in the last sentence of the Abstract. He assigns two wave functions to a system - one coming from the past and one from the future measurement. It is not clear to me whether

this will lead to new insights into the nature of time. Aharonov does think so and there are others too who feel that there are new fundamental discoveries to be made about the nature of time [11]. While this is certainly something interesting to think about, it definitely goes beyond my introductory premises so I will not pursue this further here.

8.12 Concluding Remarks

The reader who has gotten to this point will have noticed that my discussion has focused almost exclusively on what is usually referred to as the thermodynamic arrow of time and on its connection with the cosmological arrow. I did not discuss the asymmetry between advanced and retarded electromagnetic potentials or "causality" [11]. It is my general feeling that these and other arrows, like the one in the wave packet reduction discussed in the last section, are all manifestations of Boltzmann's general principle, and of the low entropy initial state of the universe. For this reason I also agree with most physicists that there would be no change in the monotone increase of entropy if and when the universe stops expanding and starts contracting.

Let me close by noting the existence of many well-known and some obscure connections between "entropy" and degree of order or organization in various physical and abstract systems far removed from the simple gas in Fig. 8.1. It is my feeling that, at least when dealing with physical objects containing many microscopic constituents, e.g. macroscopic or mesoscopic systems, the distinction between Boltzmannian and Gibbsian entropies, made earlier for simple systems, is always important and needs to be explored. I am therefore suggesting that there is interesting work to be done on obtaining more refined definitions of such concepts for complex systems like a Rembrandt painting, a beer can, or a human being. It is clear that the difference in S_B between a Rembrandt and a similar size canvas covered with the same amount and type of paint by some child is orders of magnitude smaller than the entropy differences we have been talking about earlier. The same is true, I am afraid, for the entropy difference, if at all definable, between a living and a dead person. We therefore need more refined, logically consistent and physically meaningful definitions of organization for a given complex system than those currently available in information or complexity theory.

Note added in proof: For a more extensive discussion of some of the points discussed here see, J. L. Lebowitz, *Physica A*194, 1 (1993).

References

- [1] R. Feynman, *The Character of Physical Law*, The MIT Press, Cambridge, Mass., 1967. Chapter 5.
- [2] O. Penrose, *Foundations of Statistical Mechanics*, Pergamon Press, 1970, Chapter 5.

- [3] R. Penrose, *The Emperor's New Mind*, Oxford University Press, Oxford 199, Chapter 7
- [4] D. Ruelle, *Chance and Chaos*, Princeton Univ. Press, 1991; chapters 17 and 18.
- [5] E. Schrödinger, *What is Life? and Other Scientific Essays*, Doubleday Anchor Books, Garden City, N.Y., 1965; The Spirit of Science, Sec. 6.
- [6] S. Brush, *The Kind of Motion we Call Heat*, Vol. VI in *Studies in Statistical Mechanics* North-Holland, 1976.
- [7] J. Lebowitz and O. Penrose, Modern Ergodic Theory, *Physics Today*, 26, 155 (1973).
- [8] H. Spohn, *Large Scale Dynamics of Interacting Particles*, Springer-Verlag 1991.
- [9] A. de Masi and E. Presutti, *Mathematical Methods for Hydrodynamic Limits*, Lect. Notes in Math. 1501, Springer-Verlag 1991.
- [10] Y. Aharonov has emphasized this point in lectures and private conversations.
- [11] A.J. Leggett, *The Arrow of Time in Quantum Mechanics*, The Encyclopedia of Ignorance, ed. R. Duncan and M. Weston-Smith, Pergamon Press (1977), reprinted in *The Enigma of Time*, Adam Hilger Ltd. (1982), ed. P.T. Landsberg.
- [12] S.W. Hawking, A Brief History of Time, Bantam Press. (1988).
- [13] E. Schrödinger, Irreversibility, *Proc. Roy. Irish Acad.*, 53, 22 (1950); reprinted in ref. [11].
- [14] Y. Aharonov, P.G. Bergmann and J.L. Lebowitz, Time Symmetry in the Quantum Process of measurement, *Phys. Rev.* B134, 1410 (1964).