8. The Feathers on the Arrow of Time

The fundamental laws of Physics appear to be essentially time reversible at the microscopic level: anything that can happen can equally well happen backwards. The "laws" applying at the everyday macroscopic level seem quite different, describing irreversible changes inevitably leading to a decrease in order. In particular, the Second Law of Thermodynamics says that the "entropy" of a closed system (a measure of its disorder) never decreases, and (almost) all changes to the macroscopic state of the system with the passage of time lead to an increase of entropy. Thus, the macroscopic laws define the so-called "Thermodynamic Arrow of Time", which points unambiguously from a more ordered past to a less ordered future. The emergence of this directed Arrow at the macro level from fundamental micro laws which are symmetric with respect to time has been well discussed. It has been shown to follow naturally from the micro laws and the way "order" is defined in our descriptions of the world we see. It is not the paradox which it first appears. I will briefly rehearse these arguments, but add little to them: they are no longer matters of controversy. However, there remains an aspect of the Arrow which does not appear to have been much discussed, and which seems worth some exploration.

If we observe a closed system at present, and find it to be partially ordered, there are indeed excellent reasons to suppose that some little time in the future it will be less ordered than at present. It has also been well noted that these same reasons would also lead us to suppose that, at some little time in the past, it was also less ordered than at present. That is, the standard (and valid) arguments showing that entropy, or disorder, of the closed system will almost certainly increase in the future also show that its entropy will almost certainly increase as we peer further into its past. Thus, while our arguments yield us the point of the Arrow pointing to a less ordered future, they also imply a point at the other end of the Arrow pointing to a less ordered past. However, as a matter of common sense, if we see a closed system to be partially ordered at present, we will usually infer that it was more ordered in the past, consistently with the Second Law of Thermodynamics which tells us that the system's orderliness should have decreased from the past time to the present.

For instance, if at present we open a well insulated box and find within it a metal bar, one end of which is hotter than the other, we feel entitled to predict that if we quickly shut the box, in ten minutes' time the two ends will be closer in temperature (a less ordered state) than they are now. But we also feel entitled to infer that ten minutes ago, the hot end was even hotter, and the cold end colder, than we now observe. This latter, commonsense, inference of a more ordered past is the kind we invariably make in discussing the physical history of closed systems. We assume that the Arrow does not have a point at both ends, but rather has a point at the Future end and feathers at the Past end. The point points to higher entropy in the future, the feathers trail from lower entropy in the past.

The substantive question I wish to address is why the Arrow has feathers at its rear rather than another point? In other words, why do we accept the arguments that entropy will increase in the future, but reject their equal implication that entropy decreased in the past? Why do we use and accept a certain, valid, logical apparatus in reasoning about the future destiny of a system, yet reject it in favour of something else in reasoning about the past history of the same system? And if we do use some other reasoning apparatus for the past, what is it and what validity does it have? I will attempt to show that we do indeed reason about future and past in different ways. In essence, I will conclude that our reasoning about the future, when rational, is a deductive process using the standard logic of statistics, but that our reasoning about the past is usually an inductive process not deducible from the standard logic of statistics. Note that I am not criticising this inductive process, nor the "commonsense" equivalent we informally and habitually employ. Rather, I will attempt to show that inductive reasoning about the past is well motivated, and, given certain assumptions about the real world, emerges from Minimum Message Length principles.

Most previous discussion of the asymmetry of the Arrow have treated the problem in different terms. Here, I draw heavily on a recent work by Huw Price [30].

Stefan Boltzmann "proved" that in an ideal gas, whatever its current distribution of particle velocities, a collision between particles would be expected to change the velocity distribution towards one of higher entropy. He noted that the proof did not rely on the direction of time, and hence that an entropy increase is to be expected when going from the time after the collision back to the time before it, just as much as in going from before the collision to after it. There is a technically valid objection to this argument. Boltzmann assumed that the velocities of his colliding particles were uncorrelated before the collision. But after the collision, they will inevitably be correlated. When retracing the history back in time, the final velocities become (when negated) the "initial" velocities of the reversed collision, and since they are correlated, Boltzmann's assumption is violated, and we cannot conclude from his "proof" that as the history is retraced into the past, entropy will increase into the past.

Actually, the objection to Boltzmann's proof applies in both directions. If a body of gas has a past, there will have been ample opportunity for its molecules to interact, both directly and via chains of collisions, so if two molecules collide at the present instant, even if they have never met before, their initial velocities cannot validly be assumed to be uncorrelated (or more precisely, statistically independent). Thus, the "proof" fails to prove rigorously that entropy will probably increase into the future. But in fact the conclusions of Boltzmann's argument can be shown to hold by more robust arguments, which again imply probable entropy increase into both the future and the past.

Price goes on to say that, even if Boltzmann's statistical argument is correct, it is overridden by the fact that we *know* the past to have been more ordered, and this acts as a "boundary condition" enforcing an unbroken entropy increase from past to present and beyond. For Price, therefore, the only question is why the past was so ordered. But (as Boltzmann already knew) we cannot rely on any memory or record of the past, of the body of gas or anything else, to deduce the probability, let alone certainty, of a low entropy past. Only *after* we have established some reason to accept a low entropy past can we consider a low entropy past as a boundary condition on system behaviour, or more weakly, a prior on the past favouring low entropy.

So, rather than accepting a more ordered past as a given, I have attempted to start from what I must take as given, namely my *present* senses and mental state, and try to elucidate why Price, I, and most others conclude that the past was more ordered than now.

8.1 Closed Systems and Their States

To simplify discussion, the question of what we can deduce about the future and past of some part of the world will start with the treatment of "closed systems". By a closed system I mean a physical, real world collection of matter and energy which is totally isolated from any outside influence. No matter or energy may escape from the system, nor may any new matter or energy enter it. While no known totally isolated system exists (except maybe the Universe), the concept of a closed system is a standard tool of Physics, as for practical purposes many real systems of interest are sufficiently isolated to be treated as being closed. For this discussion, I make the simplifying assumptions that the closed system is of fixed, finite spatial extent. In effect, our closed systems will be some matter and energy confined within a fixed, impenetrable box. The amounts of matter and energy are assumed to be known. (In one example, energy will be allowed to enter or leave the system, but discussion of this is deferred till that example. Until then, assume the energy of the system cannot change.) The "state" of a system at some instant, e.g., the present, is a collection of numbers which together completely specify all of the information about the system which is relevant to its future behaviour. For instance, in the case of a closed system comprising a box containing some inert gas, the present state would comprise specification of the position and velocity of every atom of the gas. As time goes by, the state of the system will change, the changes being governed by some "laws of physics".

A "view" of a system at some instant is a usually incomplete and/or imprecise specification of its state at that instant. For our box of gas, a view might specify every atom's velocity exactly, but its position only with an accuracy of one millimetre. A view could be much broader: it might only specify the temperature of the gas, and the observation that the gas appeared to be in equilibrium.

The "state space" of the system is the set of all states which can be specified, subject to the given confinement and the amounts of matter and energy. For simplicity, I will assume that these givens, or *invariants*, are known exactly (and of course do not change).

The "possible states" of a system is the set of all those states into which any present state consistent with the invariants and the present view could evolve at any time in the future, or could have obtained at any time in the past. Clearly, given a present view of the system, any state rationally asserted to have obtained in the past, or to obtain in the future, must lie in the set of possible states.

I make the further simplifying assumption that the state space of the system is discrete. That is, no numbers with infinitely many decimal places are needed for an exact specification of a state. This assumption has some justification in quantum mechanics, where the state space of a confined system is indeed a discrete set.

Finally, I assume that time changes in discrete, equal steps. Thus, the history of the system during some period is fully specified by the sequence of states it reaches at the discrete instants within the period. The discrete instants are indexed with the integers, and unless otherwise noted, the present instant is $t = 0$. This discretization of time is of course unrealistic, but I hope its assumption is not critical to the argument. With these assumptions, the "Laws of Physics" obeyed by the system are described by a mapping of the state space into itself.

8.2 Reversible Laws

I assume the Laws of Physics to obey the conservation of energy, momentum and angular momentum and, like the real Laws as we know them, to be essentially reversible. Let $s(t)$ denote the state of the system at time t. If the system is deterministic, the Laws form a one-to-one mapping $M()$ where $s(t + 1) = M(s)$. Time reversibility requires that if q and r are states, and $r = M(q)$, then $q' = M(r')$ where s' signifies a state identical to s, save that all velocities (and any other state variables which are odd derivatives with respect to time) are negated. Then the laws of physics appear exactly the same whether going forwards or backwards in time. I assume time reversibility.

If the system is non-deterministic, the mapping is a probabilistic mapping described by the function $M(r|q)$ where r and q are states and $M(r|q)$ is the probability that $s(t+1) = r$, given that $s(t) = q$. In this case, we require for reversibility of the laws that $M(r|q) = M(q'|r')$. Here and elsewhere, all probabilities and conditional probabilities are implicitly further conditioned on the known invariants of the system, that is, those properties of the system which are known and constant, e.g., the amount and type of matter, the spatial confinement and the total energy.

8.3 Entropy as a Measure of Disorder

In general, the calculation of the entropy of a closed system given its invariants and some view of its macroscopic situation can be complicated. In the special case, typified by a confined ideal gas, of a system of N identical tiny particles the state of each being fully determined by just a position and a velocity, a simple expression applies. Let z denote a (discrete) position-velocity pair which a particle may have, and Z the set of all possible such particle states. I use z to emphasize that a particle state is not to be confused with a system state. In this simple case, the system state s of the gas is the N-tuple of the particle states of all its N particles.

The known invariants and view may suffice to allow us to estimate the fraction of particles in each particle state z, or equivalently, the probability $\Pr(z)$ that a certain arbitrarily chosen particle has particle-state z. Then the entropy, here denoted by the symbol H, is classically defined by the equation

$$H = -\sum_{z \in Z} [\Pr(z) \log \Pr(z)]$$

The probability distribution $\Pr(z)$ is conditioned only on the known invariants and view. H measures the expected amount of information needed fully to specify the instantaneous state of every particle, and hence the system state, given only the invariants and whatever information is contained in the view. (Note that the particle-state distribution is constrained by the requirement that the sum of the particle energies equal the known invariant total system energy.)

Informally, a system is regarded as being in *equilibrium* if its present view is expected to persist forever with no significant change (assuming no external intervention). More formally, the system is in equilibrium if its entropy, as calculated from the present view, is close to the maximum value it could obtain if constrained only by the invariants. That is, the information in the present view does not imply an entropy lower than the maximum possible. As the probability distribution $\Pr(z)$ which maximizes H subject to the invariants is unique, the system appears to be in equilibrium if the present view is consistent with the maximum entropy particle-state probability distribution.

If the present view implies that $\Pr(z)$ differs from the maximum entropy distribution, the system is not in equilibrium, and its entropy is less than the maximum. Since the entropy is a measure of the amount of information needed fully to specify the system state given the present view, a non-equilibrium view conveys some information about the present state, and the system may be said to be partially ordered. The nature of its order is described by the non-equilibrium features of the present view.

Strictly speaking, for a partially ordered system, the above expression for its entropy should have an additional term, showing the amount of information needed to describe the view features which condition the probability distribution $\Pr(z)$. However, this term is negligible compared with the main term except for very small systems and those views including highly detailed features, and is usually ignored.

The classical definition of entropy defines the entropy of a view, i.e., a collection of macroscopically similar states. Since it shows the amount of information needed to specify the exact system state given the view, and there is no information available about the state save what is contained in the view and invariants, the entropy of a view is essentially the log of the number of system states consistent with the view. It is the amount of information needed to specify one of this number. If instead of a view we consider a state, no further information is needed to define it: the "collection" has only one member. Hence, the definition is not directly applicable to a single state. However, the definition can be extended to apply to a state if we allow the particle-state probability distribution $\Pr(z)$ to be re-interpreted, in the case of a single system state, as the frequency distribution of particle states in the system state. Then a "high entropy" state is one in which the particle-state distribution approximates the particle-state probability distribution of an "equilibrium" or non-informative view. Some low entropy states would present a low entropy macroscopic view to an observer whose measurements revealed the state's departures from the equilibrium particle-state distribution, but for some low entropy states, the departures might be too subtle to show up in macroscopic view having a realistic level of detail. For example, a state in which only every second possible particle-state was occupied (according to some arbitrary enumeration of particle states) would have low entropy according to this definition, but would probably show no macroscopic evidence of order. This difficulty in the definition of the entropy of a state fortunately is of no consequence to our arguments, as almost certainly partial order visible only at a microscopic level would be destroyed in a very short time.

It follows from this definition of the entropy of a state that the number of states with entropy H increases exponentially with H. The great majority of states in the state space have entropies close to the maximum possible value. If the system is in equilibrium (no partial order and no informative current macroscopic view) we may expect that as time goes by the state entropy will jitter around just a little below its maximum value.

8.4 Why Entropy Will Increase

If we are presented with a closed system which appears to be in equilibrium, there is nothing which can be usefully predicted of its future, save that it will probably remain in equilibrium. We can hope to make interesting predictions about the future of a closed system only if it is now partially ordered, i.e., has entropy significantly below the maximum. I will now argue that we may confidently predict that its entropy will increase. Several cases will be treated.

(a) Deterministic Laws, Exact View.
Assume the Laws of Physics are deterministic, and that we have exact knowledge of the present state $s(0)$ at time $t = 0$. Suppose $s(0)$ is a state of low entropy H_0. The deterministic laws allow us to deduce the state $s(f)$ at some future time $f > 0$. The repeated application of the time-step mapping $s(t+1) = M(s(t))$ will no doubt have changed the state and its entropy. Since the mapping is based on the microscopic laws, which make no reference to entropy and are reversible, there is nothing inherent in the mapping which can lead us to expect that the entropy of a state will be related in any obvious way to the entropy of its successor, although, since a single time step will usually produce only a small change in the macroscopic view of the system, we may expect the resulting change in entropy to be small. After the f steps leading to time f, therefore, we can expect the initial entropy to have changed by some amount, but otherwise to be almost a random selection from the set of possible entropy values. Recall that in a closed system, the number of states of entropy H is of order $\exp(H)$. There are therefore far more high entropy states than low entropy ones, so the chances are that $s(f)$ will very probably have an entropy greater than the initial state $s(0)$.
In a deterministic closed system of constant energy, the sequence of states from $t = 0$ into the future is a well-defined trajectory through the state space. The trajectory may eventually visit all states in the space, or only some subset. Given that the state space is discrete, it can only contain a finite, albeit huge, number of states. The trajectory from state $s(0)$ thus must eventually return to state $s(0)$, and since the mapping is one-to-one, must form a simple cycle in which no state is visited more than once before the return to $s(0)$. Since it is cyclic and produced by time reversible laws, the trajectory must show just as many steps which decrease entropy as increase it. However, because the trajectory can be expected to visit far more high entropy states than low entropy ones, if we start from some

low entropy state and advance for some arbitrary number of steps, we must expect to arrive at a state of higher entropy. If this argument is accepted, the Second Law will in general apply even given deterministic, time reversible laws. No paradox is involved.

(b) Deterministic Laws, Inexact View.
If, rather than knowing the precise initial partially ordered state $s(0)$, we have only an imprecise view of the initial state, the argument for expecting entropy to increase can only be strengthened. We may well feel more justified in regarding the entropy at future time f as being in some sense a random selection from a range of values, since we no longer have any certainty about the future state $s(f)$. Given this uncertainty, the larger number of high entropy than low entropy states becomes perhaps a sounder basis for expecting $H_f > H_0$.
Because each time step has but a small effect on the macroscopic view, we also expect the macroscopic behaviour of the system to be insensitive to the fine detail of the initial state, and to be well predicted by a macroscopic, or at least inexact, initial view.

(c) Nondeterministic Laws
If we suppose the fundamental microscopic laws to be nondeterministic (but still reversible) the state-to-state mapping $M()$ is replaced by a probabilistic law

$$\Pr\left[s(t+1) = r\right] = M(r|s(t))$$

Given $s(0)$, we can at best compute a probability distribution over the possible states at time $f > t$, and are the more justified in expecting $s(f)$ to have higher entropy than $s(0)$.

8.5 A Paradox?

The preceding arguments, backed up by the results from simulations of a small but, I hope, plausible thermodynamic closed system described in a later section, show that macroscopically irreversible behaviour not only can, but must be expected to, emerge from the operation of time reversible microscopic laws on an initially partially ordered system, especially but not only if these laws are non-deterministic. In particular, we can confidently expect the entropy of the system to increase. The arrow of time unambiguously points to a future less ordered than the present.

But a paradox of sorts remains. Re-reading the arguments of cases (a), (b) and (c), it can be seen that, although it was assumed that the predicted situation occurs at a time f in the future ($f > 0$), this assumption actually played no part in the argument. If f is replaced by a time b in the past ($b < 0$), all that was argued to the effect that $H_f > H_0$ becomes an equally valid argument that $H_b > H_0$. Thus, I conclude that while simple statistical deductions show that entropy will very probably increase in future, these same deductions show that entropy very probably decreased from the past. The entropy shown by the present state or view of the system is therefore almost certainly close to the lowest it will ever have, or ever did have, at least for times not hugely remote from the present. This is precisely the conclusion reached by Boltzmann. However, it is not what we would normally conclude about the past of a presently partially ordered closed system, and many would like to avoid it.

8.6 Deducing the Past

Before dismissing this deduction of past disorder as nonsense, let us consider what means might be used in the real world to deduce something about the past state T (for "Then") from exact or partial knowledge of the system's present state N (for "Now"). As before, I assume the microscopic laws (whether deterministic or probabilistic), the exact system energy, other invariants, and the state space all to be fully known. I will consider three cases: non-deterministic laws given an inexact view N^V of N; deterministic laws given full knowledge of N; and deterministic laws given an inexact view N^V. There are several approaches which might be taken to the deductive task.

8.6.1 Macroscopic Deduction

The most direct approach would be to assume the universal validity of the macroscopic laws of motion, Thermodynamics, Chemistry, etc., and attempt to use them to deduce the past state from the present state. While sometimes successful, this approach has severe limitations, as will appear.

These macro laws appear, as normally presented, to be deterministic: they are expressed in equations with no random terms. Some of them, e.g., the second law of Thermodynamics and the laws of viscous flow and diffusion, are not time symmetric, and describe "dissipative" behaviour which reduces order and increases entropy as time advances. Today, the dissipative "laws" are not seen as fundamental, but rather as emerging from the large-sample statistics of reversible micro processes. However, they have been well tested and found to be excellent models and predictors of macroscopic behaviour, so it is not unreasonable to assume that they held in the past period from state T to state N. The empirical evidence supporting them is so strong that we may be prepared to treat these laws as fundamental, and to ignore the argument that the large-sample statistics of reversible micro processes do not support them except for the prediction of future behaviour. If so, we may proceed to set up the differential equations embodying these laws, set N (or some macroscopic view N^V of the present) as boundary conditions at $t = 0$, and integrate the equations back to $t = -k$.

This approach may work adequately for small k, i.e., for the recovery of the very recent past, but in general is likely to fail. For instance, let our system be a long thin pipe closed at both ends and filled with water, with the present view showing some dye colouration in a short stretch of the pipe. If there is no bulk motion of the water and no significant temperature or gravity gradient, one of the applicable macro equations will be the equation for one-dimension diffusion of the dye concentration $v(x,t)$:

$$\frac{\partial v}{\partial t} = K \frac{\partial^2 v}{(\partial x)^2}$$

Taking the present-view dye distribution $v(x,0)$ as boundary condition, this differential equation can be integrated to negative times. Unfortunately, if the integration is carried too far back, it is very likely to yield negative values of dye concentration at some points of the pipe, which are of course impossible. For any negative time t, the function $v(x,0)$ is the convolution of $v(x,t)$ with a Gaussian function of x, so recovering $v(x,t)$ from $v(x,0)$ is a deconvolution calculation which is very ill-conditioned and may have no positive solution.

Another sort of possible failure arises from the fact that a purely deductive process based on backward integration of the macro laws has no room for our background knowledge. For example, suppose our system is contained in a strong, insulated box, and our present view shows the box to contain a hot mixture of nitrogen, various gaseous oxides, some sooty deposits of carbon and potassium compounds on the walls of the box, a small wooden stick charred at one end, and some fragments of discoloured paper. I cannot believe that integration of the macro laws could lead us to a past state where the box contained a firecracker and a lighted match.

More generally, if in the original state T some part of the system was in a low entropy metastable state, and the actions of some other part of the system triggered the collapse of the metastability, it is hard to see how deduction from the present N using the macro laws can lead to a recovery of the metastable state. In summary, the use of macro laws to deduce the past seems to have limited utility, and be unlikely to reach many “commonsense” inferences that we make about the past, except in cases such as the motions of the major bodies in the solar system, where dissipative processes and hence entropy increases are very small.

8.6.2 Deduction with Deterministic Laws, Exact View

If the laws are deterministic and we have exact knowledge of the state N, there is no problem. We simply trace back the history leading to N by using the inverse of the one-to-one mapping of states $M()$. Of course, we recover T without error. I will term this backwards tracing “devolution”, and say that N devolves to T.

The case is wholly unrealistic. In the real world, we never have exact knowledge of the present state of a closed system.

8.6.3 Deduction with Deterministic Laws, Inexact View

Assuming our knowledge of the present state N is an inexact view N^V, we can at best enumerate the collection of states compatible with N^V and trace each back to its only possible ancestor at the "Then" time. This will give us a set T^P of possible ancestors exactly as numerous as the collection of possible present states, and one of this set must be the true ancestor T. Deduction can at best lead to a probability distribution over the members of T^P. If all "Then" states were considered equally likely *a priori*, the distribution over T^P will be uniform, since each state in T^P has the same probability (one) of evolving into a "Now" state in N^V. We consider a non-uniform prior over "Then" states in the next section, but find the idea unsound.

We are left with the problem of deducing some conclusion about T knowing only that it is some member of T^P. One possible approach is to form a view of the past by averaging over the members of T^P. This seems not to be very helpful. The true "Now" state N devolves to the true "Then" state T, which may well have lower entropy than N. However, the precise configuration of N which implies its ancestor has lower entropy is extremely fragile. A state N^* differing only microscopically from N will very probably devolve to an ancestor of higher entropy. If we imagine the state space to be arranged so that states which are microscopically similar are close together, we find that in general, two trajectories which pass through neighbouring states in state space will diverge rapidly (whether traced forwards or backwards). While the (past) trajectory through N may pass through a lower entropy T, it is probable that (past) trajectories through close neighbours of N will diverge to higher entropy "Then" states.

Simulation experiments support this assertion. If a simulation is run from a low entropy state T, it will almost always reach a higher entropy state N which, providing the elapsed time is not too great, will still be partially ordered. In many hundreds of trials, it was found that if N was then perturbed by the smallest possible displacement of just one of the typically 10^5 particles involved, or by the reversal of a single particle velocity component, the resulting perturbed state devolved to a "Then" state with higher entropy than N and of course T. Further, the "Then" state reached presented a macroscopic view very different from T. It appears that for any realistic view N^V of the present state, the vast majority of states in T^P differ greatly from T, and in particular have entropies greater than N, not less.

It follows that trying to deduce the nature of T by averaging over T^P will lead to the conclusion that the past was less ordered than the present, and very different from what common sense would suggest. Simply averaging T^P to estimate some property of the past state corresponds to the (well justified) Bayesian "minimum loss" process if the "loss function" assumed is the square of the difference between the true and estimated value. One might consider whether some other loss function might yield estimates more in line with commonsense inferences about the past.

The only candidate loss function I have thought of which might work to discriminate among the members of T^P is a function which shows how much the evolution of some member of T^P to a member of N^V does violence to the usual (but not fundamental) irreversible physical laws such as the laws of diffusion, viscosity, friction, and in particular the Second Law. It is highly likely that the true evolution of T to N has pretty much conformed to these laws, but the evolution of a high entropy past state to some state in N^V will be in criminal violation. There are, unfortunately, two flaws in this suggestion. First, we are likely to infer a past state of *lower* entropy than T, showing fine-scale macroscopic structure which dutifully is attenuated in strict accordance to the irreversible laws until it matches the faint irregularities visible in the present view which common sense would consider the result of statistical fluctuations such as thermal noise, Brownian motion, etc., and in no way informative of the past. Second, and more importantly, we have no warrant for assuming such a loss function. We have as yet found no deduction proving that these "laws" held sway in the past, and hence no reason to suppose that their past violation is in any way a "loss".

8.6.4 Deduction with Non-deterministic Laws

I will assume an inexact view N^V of the present state N. Since the laws are non-deterministic, every state in N^V may have several possible immediate ancestors at time $t = -1$, and each of these several possible ancestors at $t = -2$, and so on. In general, we expect the set T^P of possible ancestral states at $t = -k$ to have cardinality much greater than N^V, and in some cases T^P may include the entire state space. Clearly, any deduction must deal with the inherent uncertainty of the past by using the logic of probability. The reversibility of the microscopic laws is easily shown to lead to the following equality, where r is a state occurring at some time t_1 and s is a state occurring at some later time $t_1 + k$.

The evolution probability $P_e(s|r, k)$ that state s will occur at time $t_1 + k$, given that state r occurred at time t_1, equals the probability $P_e(r'|s', k)$ that state r' will occur at time $t_2 + k$, given that state s' occurred at some time t_2.

(Recall that s' is a state identical to s save that all velocities and other odd derivatives with respect to time are negated.) In what follows, r and s will always refer to states separated by k time steps. We can therefore abbreviate the equality to the symbolic form:

$$\text{THE RETRACE EQUALITY:} \quad P_e(s|r) = P_e(r'|s')$$

The probability that r evolves into s after k time steps along some particular trajectory from r to s is the product of the probabilities of all the random choices made along the trajectory. For every such trajectory from r to s there exists a trajectory from s' to r' along which exactly the same random choices are made, in reverse order. Thus, s' has just the same probability of evolving into r' along this trajectory as r has of evolving into s along its trajectory. Summing probabilities over all trajectories from r to s to get $P_e(s|r)$, and over all trajectories from s' to r' to get $P_e(r'|s')$ yields the retrace equality.

The retrace equality gives us a simple way to investigate the Bayesian posterior probability distribution over T^P given a present view N^V. Let s be some state in N^V. The posterior distribution over states in T^P given present state s is given by Bayes' theorem as:

$$P_f(r|s) = P_e(s|r)P_p(r)/P_m(s)$$

where r is a state in T^P at time $-k$, and $P_p(r)$ is the "prior" probability that the system would be in state r at time $-k$. The prior probability is the probability we would ascribe to r given only knowledge of the system invariants, i.e., before seeing the present view.

$P_m(s)$ is the implied marginal distribution over present states given by:

$$P_m(s) = \sum_r [P_p(r)P_e(s|r)]$$

Let us assume that $P_p(r)$ is the same for all states in the state space. This is a reasonable assumption given that in a closed system, all states in the space have equal probability unless some non-equilibrium feature of the state is known, which is of course not the case before we have observed the present view. If $P_p(r)$ is uniform over the state space, so is the marginal distribution $P_m(s)$. In fact, for every state in the state space, its marginal probability at time $t = 0$ equals its prior probability at time $t = -k$, and $P_e(s|r) = P_f(r|s)$.

Given this uniform prior, a Monte Carlo approximation to the posterior could be obtained by repeatedly selecting a state s in N^V at random and using the retrace equality to trace a possible trajectory backwards from that state to time $-k$. The tracing backwards is performed by first replacing s by s', then letting s' evolve for k steps according to the microscopic laws, and finally replacing the resulting state r' by the time reversed state $r = (r')'$. The retrace equality guarantees that the probability that a state r at time $-k$ would evolve to a state s at time zero equals the probability that tracing backwards from s will give r at $t = -k$. I will say that s devolves to r.

More generally, we may say that a state r at an early time evolves to a probability distribution over states at some later time, and that a state s at the later time devolves to a probability distribution over states at the earlier time. The collection of states found by the repeated devolution of states randomly selected from N^V is an unbiased sample from the posterior over T^P. States in T^P in general will not have equal likelihoods of evolving into states in N^V. Hence, assuming equal prior probabilities for all states at $t = -k$, the posterior probabilities given N^V of states in T^P will be unequal. The devolution of s in N^V to r in T^P is equivalent to the evolution of s' for k time steps into r'. As argued, and as confirmed by simulation, any evolution of a partially ordered state is expected with very high probability to conform to the normal macroscopic laws of thermodynamics, and in particular, to result in a state of higher entropy than the original. Almost certainly, $H(r') > H(s')$. Since time reversal does not alter the entropy of a state, almost certainly, $H(r) > H(s)$. That is, almost all states in T^P will have higher entropy than those in N^V, and the Monte Carlo sampling of states from the posterior probability distribution over T^P will contain a great majority of states with higher entropy than N.

Both evolution and devolution of a partially ordered state yield states of higher entropy with probability close to one. The Bayesian expectation is that the ancestor at $t = -k$ of N had more entropy than N, and this expectation is as well-founded as the expectation that N will evolve to a state of higher entropy at $t = +k$.

Moreover, the great majority of possible ancestor states, possessing collectively the bulk of the posterior probability in T^P, have little chance of evolving into states in N^V, being of higher entropy than N. Evolution of a high entropy state in T^P into a lower entropy state in N^V is possible but very improbable, as it implies an evolution running counter to the normal thermodynamic laws.

The posterior distribution over T^P summarizes all we can deduce about T from the present view N^V. Further deduction can only tell us things implicit in this distribution. As in the previous Section 8.6.3, with a uniform prior over T^P, Bayesian minimum loss deduction with a quadratic loss function is equivalent to posterior weighted averaging over T^P, which could be done by Monte Carlo sampling. As just shown, this will give the minimum loss value for the past state entropy as greater than the present entropy. The argument of Section 8.6.3 against the validity of more biased loss functions still applies.

As the minimum-quadratic-loss estimate of the entropy of T is so poor, we must doubt the quality of similar estimates of other properties of T (a doubt confirmed by many simulation results: see Figure 8.1 and Figure 8.2).

8.6.5 Alternative Priors

The counter-intuitive behaviour of Bayesian deduction noted above was obtained assuming a uniform prior $P_p(r)$ over the past states. Although there are arguments for the uniform prior, it is worth asking if a different prior could produce Bayesian posteriors more in accord with commonsense inferences of the past. For instance, one could consider a prior which placed more weight on past states of low entropy, in order to make the mean posterior entropy lower than in the observed present.

A sufficiently strong prior preference for low entropy past states might indeed bring the Bayesian deductions about T into line with common sense. However, it seems hard to provide a rational basis for such a "prior". For Bayes' theorem to be validly applied, it must be the case that the prior

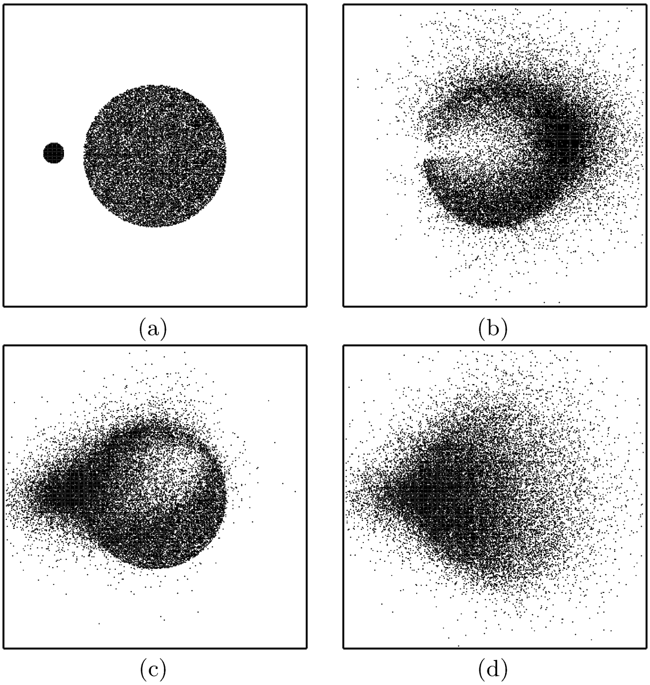


Fig. 8.1. (a) "Then" state of "Disc Collision" model; two-thirds scale; 31,623 atoms. (b) "Now" state of "Disc Collision" model; 35 time steps after "Then". (c) Reconstruction of Disc Collision "Then" from "Now" with one atom perturbed; deterministic devolution. (d) Reconstruction of Disc Collision "Then" from exact "Now" state; non-deterministic devolution.

probability distribution reflects the probabilities which might rationally be assessed on the basis of knowledge held *before* the present view is observed. In our case, this knowledge comprises only the microscopic laws of Physics, the system invariants, and the belief that the system has been isolated since before the time $t = -k$. That is, the prior knowledge is only what we have before we open the box to take the present view. There seems nothing in this knowledge to suggest we should believe that, had we opened the box at $t = -k$, the chances of our finding the system in some particular state q would depend on the entropy (or any other property) of q.

Further, adoption of a non-uniform prior $P_p(r)$ at time $-k$ leads to a non-uniform marginal distribution $P_m(s)$ for the present state, but $P_m(\cdot)$ is closer to uniformity than $P_p(\cdot)$, being smeared out by the variance of the evolution distribution $P_e(\cdot|\cdot)$. Thus, we would be supposing that the probability of finding the system in some particular state q depended on when we looked. Since we have no knowledge of when the system became isolated, or what its state was at that time, this supposition can have no basis.

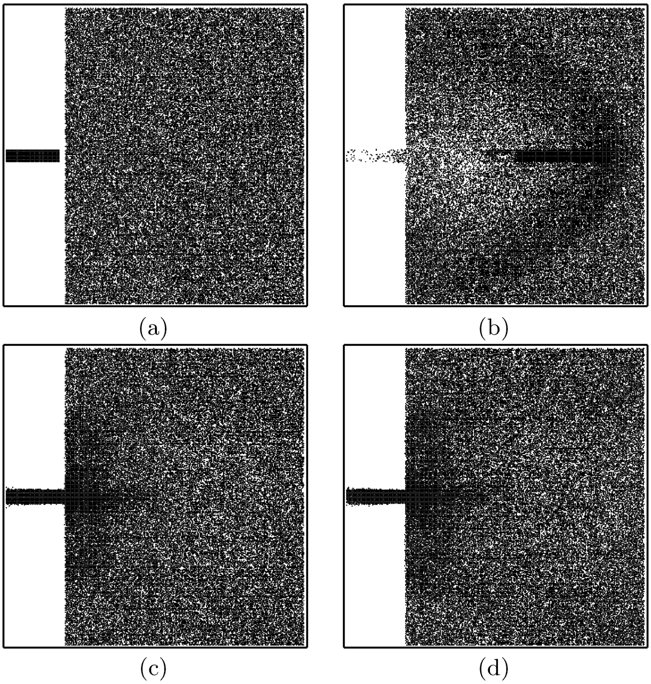


Fig. 8.2. (a) "Then" state of "Stick Collision" model; two-thirds scale. (b) "Now" state of "Stick Collision" model; 190 time steps later than "Then". (c) Reconstruction of "Then" state of "Stick" model from "Now" state with 100 atoms perturbed; deterministic laws. (d) Reconstruction of "Then" state of "Stick" model from exact "Now" state; non-deterministic laws.

The only prior which seems a defensible representation of our ignorance of T before observing the system is a prior which does not change with time. The uniform prior uniquely has this property. Price suggests that since we *know* the past to be more ordered, we can use this "fact" to modify the deduction of the past somehow to recover a lower entropy state from a present state. He does not say how, and in any case, we have so far seen no reason to accept his "fact".

Finally, it is not clear that a Bayesian prior is even applicable to the deduction of a probability distribution over possible "Then" states. We are happy to base predictions of future states on our knowledge (be it partial or exact) of the current state N, without reference to any "prior" beliefs about the future. The present state, by definition of a state, embodies all information relevant to the future behaviour of the system, so our calculations of the probabilities of possible future states need refer only to N, or our view N^V of N, unclouded by any distraction. Why should we lose our confidence in the sufficiency of the N state when computing the probabilities of possible past states? I see no compelling reason. If the sufficiency of N is accepted, we can make no better estimate of the probability of a possible past state T, given N, than that given by the Retrace Equality, and conclude

$$\Pr(T|N) = P_e(T'|N')$$

The result is of course equivalent to the assumption of a uniform Bayesian prior over past states.

8.6.6 A Tale of Two Clocks

A simple example will illustrate just how bizarre the results of attempting to deduce the past from the micro laws can be.

Suppose the closed system is an insulated box containing an ordinary mechanical clock with a mainspring coupled via gearing to the hands of the clock and to a conventional escapement governed by a balance wheel oscillating in resonance with a hairspring. Now suppose that at time "present" we inspect the clock and find that its mainspring is about half run down, and that the hands show 7 o'clock. We want to infer the state of the clock exactly one hour ago.

In such a clock, the movement is stationary most of the time, except for the balance wheel which oscillates, releasing the movement to advance by one tooth of the escapement wheel on each oscillation. For simplicity, assume that the present view is obtained at a time when the balance wheel is momentarily stationary at one extreme of its oscillation, and that the rest of the movement is also still. (This assumption is not crucial to the argument.) Now let us attempt to deduce its past state by applying our knowledge of the micro laws to backtracking from the present state. We can follow the process described in the previous sections.

First, choose randomly some state s consistent with the present viewN^V, then time reverse it, giving s'. What will s' look like? Since there was no macroscopic movement taking place in s, there will be none in s'. Thermal noise vibrations on the micro-scale will be reversed in direction, but these have no significant macro effects. In fact, s' is consistent with N^V. So s' looks just like a clock with half-wound mainspring, balance wheel about to start a swing, and hands showing 7 o'clock. Now we let this state evolve for an hour in conformance with the micro laws. If s' looks like the state of a clock, it will evolve like a state of a clock. After an hour's evolution, corresponding to an hour's devolution of s, state s' will have done what a clock's gotta do, and will have changed to a state r' showing exactly 8 o'clock with a mainspring rather more than half run down. Finally, we time reverse r' to obtain the deduced past state r. As before, time reversal does nothing significant to the state except perhaps altering the phase of the balance-wheel oscillation by up to half a cycle, a jump of no more than half a second.

Our deduction, supported with extremely high probability by the fundamental laws of Physics, is that the clock, which now says 7.00, an hour ago said 8.00 and was more run down than it is now! (A grandfather clock would give similar results, but would need a bigger box.)

Since this clock seems as queer as a two-bob watch, we discard it and build our own. We build a disc free to turn on a vertical shaft. The disc carries two radial tracks on which two masses can slide freely from near the centre of the disc to its rim. Springs running from fixed points on the disc to the masses pull the masses towards the centre with a force exactly proportional to the distance of the mass from the centre. If the disc rotates with a certain speed, the inwards forces of the springs on the masses will exactly equal the outwards centrifugal force on them, no matter what the radial positions of the masses. To "wind" the clock, we spin the disc until it reaches this speed with the masses nearly at the rim, then let it spin by itself. It will continue to spin at the critical speed for a long time without slowing down, the masses gradually moving inward as kinetic and spring energy is released to overcome friction. Low-friction gearing couples the disc to the hands of our clock.

Now suppose our new clock becomes the closed system under study. Let our "present" view N^V be that we find the hands showing 7 o'clock as before and the masses are half way between centre and rim. Again, we attempt to deduce the state of the clock an hour ago by backtracking some state s in N^V. The time reversed state s' now has the disc spinning with the critical speed, but backwards, and otherwise looking just like s. To devolve the system by one hour, we deduce the state r' reached by evolving s' for an hour. If the friction is low enough, r' will still show the disc spinning with the constant critical speed, but the masses will have crept inwards. An hour's backward spin will have left the hands showing 6.00. Time reversing r' to get the inferred past state r, we find that an hour ago our new clock showed 6.00, not 8.00, but was less "wound-up" than at present. Great news! We now have a clock which not only keeps proper time when no-one is looking, but also winds itself!

If both clocks are put in the box and both now show 7.00, we have shown by impeccable reasoning that almost certainly, an hour ago one clock showed 6.00 and the other 8.00. The nature of the different behaviours of the clocks is related to the following observations.

– Clock 1 cannot be made to run backwards. Clock 2 can: one just "winds it up" by spinning the rotor in reverse until it reaches critical speed with the weights far out.
– Clock 1 (or any clock using an escapement) will not work properly without some slight friction and/or inelasticity in the escapement. It needs dissipative phenomena. Clock 2 does not: the less friction the better.
– Clock 1, once fully wound, will run for a time independent of the amount of dissipation in its works (within limits). Clock 2, once fully "wound", will run for a time inversely proportional to the amount of dissipation.

An hourglass, or a digital watch, will resemble clock 1. A sundial resembles clock 2. I hope this tale is sufficient to show that attempting to recover the past state of a closed system by statistical deduction based on the fundamental microscopic laws of Physics is both unusual and unwise.

8.7 Records and Memories

It may seem that by concentrating entirely on what can be deduced about the past from a single, instantaneous, present view, I am setting up a straw man. Even for some isolated closed system, we often have memories or records of the past which are informative of the system's past, and if we include these in our deductions, we should usually be able to deduce a lower entropy past state for the system. However, this argument fails because the system we think we are studying is then not really isolated. It must have interacted with the record-making machine or person which now holds the record or memory. By bringing a record of the system's past into the deduction, what we are really doing is enlarging our present view to include, not just the state of the nominal system but also the present state of pieces of paper and/or nerve cells. If these are included in the present view, our deductions of the past should be based on deductions of the past of the extended system which includes the recording apparatus. Suppose that we know nothing of the closed system save its energy, and that it is a box undisturbed for several hours. We open the box and find that it contains some physical system S and, separated by an insulating partition, a self-developing camera connected to a timer, whose clock now says 6.00, and which was clearly set to trigger the camera at 5.00. A picture has emerged from the camera showing the system S in a lower entropy state than it now has. Can we now deduce from this present view of S and picture that S really was in a low entropy state an hour ago? I agree with Boltzmann that we cannot.

As with any isolated system governed by reversible non-deterministic laws, devolution of the present state is far more likely to result in a past state of higher entropy than the present. The tale of two clocks suggests that devolution of the present view is likely to show the clock running backwards, and lead to a view at 4.59 in which the clock shows 7.01, the picture was already in its developed state (which has higher entropy than the metastable state of undeveloped film), and S was in a higher entropy state than now. During the period 4.59 to the present (6.00) the clock never reached an indicated time of 5.00, and the camera was not triggered. Despite its ridiculous nature, this "view" of 4.59 (or 5.00) has far more posterior probability given the present view than has the commonsense conclusion.

Human memory is of course a more flexible recorder than a camera, but I see no reason why it too should not fall victim to an argument similar to the above. In a deductive reconstruction of the past system state using a memory of the past, we must also reconstruct the past state of the observer. Very probably, the reconstruction will follow one of the individually wildly unlikely but collectively hugely numerous paths to a past along which the

Fig. 8.3. Two-chamber + Pipe model; original "Then" state; pipe flow left to right.

system-observer interaction now recollected never took place, or if it did, happened differently from what is remembered.

I am not arguing here for a rejection of the commonsense conclusion in favour of the high-posterior view, merely arguing that deduction, even using what appear to be records showing a past state in accord with common sense, does not lead to the commonsense conclusion.

8.8 Induction of the Past (*A la recherche du temps perdu*)

We are forced to the conclusion that the kind of deductive reasoning we habitually and successfully use in predicting the future of a system is inappropriate for inferring its past. I will argue that, whereas our predictions of the future are based on deduction from the present view to the possible and probable future states, our reasoning about the past is usually inductive. The term is used in its limited sense of reasoning from the particular to the general, or at least reasoning from a data statement to a statement more general and stronger than the data statement.

If we think about the kinds of conclusions we make about the past, these do not seem to have the form of "posterior distributions" or means thereof. Rather, we will often conclude by presenting a view T^P of a past situation, in the sense used here of an incomplete and imprecise specification of a state, often with an appended account of how that state might have evolved into the present state or view on which we base our conclusion. In other words, we tend to conclude with a view of the past which satisfactorily explains the present.

To return to a simple example, if at time zero we observe a gas in a box with two chambers connected by a narrow channel, and find in the present view that the left chamber has higher pressure than the right, and that the net flow of gas in the channel is from left to right, what might a reasonable person conclude? (Figure 8.3.) Given an assurance that the box was well insulated and that there had been no external interference in the last five seconds, the reasonable conclusion which most would vote for is that five seconds ago, the pressure in the left chamber was even higher than it now is, the pressure in the right chamber was less, there was then more gas in the left and less in the right than is now the case, and that gas has been flowing from left to right from then till now.

Why does this conclusion seem reasonable? The past view postulated is not extraordinary, if true it would lead with good probability to the present view, and the postulated process leading from the past view to the present is in accord with our knowledge and experience of macroscopic behaviour. In particular, the asserted past view has lower entropy than the present one in accordance with our knowledge of thermodynamics.

This last point is common to most of our conclusions about the past of well-isolated systems: we postulate a past view or state of lower entropy than the present. Since in a closed system there are far fewer low entropy states than higher entropy ones, the assertion of a low entropy past state is a stronger statement than can be deduced from the present view. In common with all non-empty inductive conclusions, it is a more general and stronger statement than the data statement on which it is based. To be an admissible hypothesis, it must not imply anything contrary to the present view. That is, the asserted past state or view must give high probability to the present view. It must have high likelihood. However, being a stronger statement than the present view, it will in general imply more about the present state N than can be deduced from the present view N^V. If these implications about the present state can be checked by more detailed observation of the present, they may be found to be false. Thus, an inductive conclusion about T is always in principle falsifiable by additional data about N (which may include observations at times later than zero.)

A principled inductive conclusion about T from the data N^V will be based on the likelihoods of possible past states, i.e., the probabilities these states have of evolving into N^V. Under reversible, non-deterministic micro laws, the evolution of any state will, with very high probability, conform to the normal macroscopic and thermodynamic laws (the "macro laws"), and in particular will not result in a decrease in entropy. Thus, the only states in T^P with a high likelihood given N^V are those states which can evolve into N^V in a manner consistent with the macro laws. Denote the set of such states as C (for "core"). Although C accounts for only a tiny fraction of the posterior probability distributed over T^P, it is among the states of C that we expect to find our inductive conclusion. It remains to consider what kind of inductive principle will best perform the task

8.8.1 Induction of the Past by Maximum Likelihood

A simple and well-known inductive principle is the "Maximum Likelihood" (ML) principle. It suggests the selection of that hypothesis which has the greatest likelihood of yielding the observed data. In our context, Maximum Likelihood suggests that we infer the state at $t = -k$ to be that state in T^P with the greatest probability of evolving into N^V. The ML inference T_{ml} will of course be some member of C. Unfortunately, T_{ml} will often be a poor choice, not conforming well with a "commonsense" inference.

Given a choice among a family of hypotheses of different complexity, ML notoriously is inclined to "overfit" the data. That is, ML will usually choose a rather complex hypothesis even when the complex details of the hypothesis are not justified by the data. In our problem, while we expect a partially ordered state to evolve most probably in accordance with the macro laws, we also expect its evolution to show minor deviations from these laws. As was found in Section 8.6.3 when considering loss functions favouring adherence to the macro laws, the likelihood of a past state is also highest when its evolution into some state in N^V closely follows the macro laws. Hence, the maximum-likelihood choice will in be expected to show in exaggerated form fine-scale macro structure in N^V which in fact arose from statistical fluctuations in the evolution from the true past state. Given a fairly precise view N^V, it is likely that the ML inference T_{ml} will show an exaggerated feature not present in T, and arising from some feature of N^V which arose through thermal noise since the time of T. Even if the system has been in equilibrium since before the time of T, a sufficiently detailed present view will show some deviations from the maximum entropy view, and T_{ml} will try to explain these as the vestiges of past order. That is, the ML inference will overfit the data by striving to fit noise in the data which is not really informative about T.

This defect of Maximum Likelihood inference is well known, and various techniques have been developed to overcome it. A useful but rather *ad hoc* technique is to apply standard tests of statistical "significance" to the various features asserted by the ML conclusion, and to delete from the ML inference those features found to be insignificantly supported by the data. This approach requires the essentially arbitrary specification of a "significance level", and has severe technical difficulties in determining appropriate significance tests to apply during the sequential deletion of several features. A more principled approach may be based on "penalizing" the likelihoods of the competing hypotheses by a penalty which reflects in some way their various complexities. A well-founded way of combining the likelihood of a hypothesis with a measure of its complexity is provided by the Minimum Message Length (MML) principle and similar methods such as Minimum Description Length (MDL) and Bayes Information Criterion (BIC).

8.8.2 Induction of the Past by MML

I do not here mean to argue the special merits of MML as an inductive principle, but merely choose it as an example to show that there is at least one respectable inductive principle which can be expected to lead to inferences about the past which are close to our commonsense conclusions. I have chosen MML because long work with and on the principle has given me some claim to a fair understanding of how it would perform in this context, but I expect any other effective inductive principle, such as a significance tested ML or Bayesian approach, would give similar results.

Given some data and a family of possible models for or hypotheses about the source of the data, MML considers the construction of a message which encodes the data in (say) the binary alphabet. The message must be intelligible to a receiver who knows the observational protocol used to collect the data, the family of possible hypotheses, the system invariants such as total energy, and some (not necessarily informative) prior probability distribution over hypotheses.

The message is required to have a particular form. It must begin with a first part which specifies a single hypothesis. This is followed by a second part which encodes the data using a code such as a Huffman or Arithmetic code which would be efficient were the stated hypothesis true. Such a message is called an *explanation* of the data.

Standard coding theory shows that the length of the second part is given by the negative log of the probability of the data given the hypothesis. That is, the length of the second part is the negative log likelihood of the hypothesis. The length of the first part, which states a hypothesis, is also supposed to use an efficient coding scheme. The optimal scheme will depend in part on the prior probability distribution (MML is inherently Bayesian), but in any reasonable scheme the amount of information (and hence message length) needed to specify a complex hypothesis will be greater than for a simpler hypothesis. Further, the more precisely any free parameters of the hypothesis are specified, the longer the statement becomes.

The MML principle asserts that the best inductive hypothesis is that which minimizes the total length (first and second parts) of the message encoding the data. Thus, there is a built-in compromise between the complexity of the hypothesis and its likelihood given the data. A complex hypothesis requires a long first part to specify it, but may give a high likelihood and hence a short second part. A simple hypothesis with severely rounded-off parameters can be stated in a short first part, but may have a lower likelihood requiring a long second part. The best compromise (shortest message) is found to be reached when the hypothesis has just the complexity and precision justified by the data.

In a context such as the recovery of the past, where the conceptually possible set of hypotheses is fixed and discrete (here, the set of states in the state space), the notion of precision of the inferred hypothesis needs some clarification. The two-part explanation will begin by asserting a single, precisely defined hypothesis (i.e., past state) and the second part will encode the data (N^V) using a code expected to be efficient were the assertion true. How then can we think of the assertion as being of limited precision? The answer is that the expected explanation length is minimized if the code used for nominating the hypothesis (which must be determined before the data is known) can only nominate one of a subset of the conceptually possible hypotheses. The smaller this subset, the shorter the first part becomes, but, when the data arrives, the sender must choose one of this subset as his inference. If the subset is too sparse, no member of the subset may have a high likelihood given the data, and so the second part will be long. The optimum choice of subset balances these conflicting effects.

The MML principle is supported by several theoretical arguments. For our purposes, the most relevant is that it can be shown that in the shortest message, the statement of the hypothesis contains almost all the information in the data which is relevant to the choice of hypothesis, and almost no irrelevant information. In the present context, an MML hypothesis about T will contain almost all the information in N^V which is about T, and almost none of the information in N^V which derives from thermal noise processes occurring in the period since T. The magnitude of the "almost" qualification is small. Although no exact general bounds are known, it appears to be less than the log of the number of free parameters whose values are stated in the hypothesis.

The exact construction of a code for hypotheses which will minimize the expected length of the two-part message is computationally infeasible except in the simplest problems, but good approximations exist for the message length which allow MML inference to be applied in problems of considerable complexity.

In the present problem of making an inductive inference about the past state T, an approximate treatment of MML described in Section 4.10 leads to the following construction, assuming all past states are equally likely *a priori.*

Given the present view N^V, construct a set Q of past states which minimizes the approximate message length

$$L_Q = -\log(|Q|/V) - (1/|Q|) \sum_{r \in Q} \log(\Pr(N^V|r))$$

where $|Q|$ is the number of states in Q, V is the number of states in the state space, and $\Pr(N^V|r)$ is the probability that past state r will evolve into a state in N^V.

The first term in L_Q is the negative log of the total prior probability of states in Q, and gives the length of the first part of the message. The second term is the expected length of the second part if a randomly chosen member of Q is asserted as T. It is easy to show that for the optimum choice of Q, the smallest log-likelihood given N^V of any state in Q is only one less than the average log-likelihood.

In words, MML inference finds a "view" Q of the past which maximizes the prior probability of the view times the *geometric* mean of the likelihoods of the states in the view.

Were the *arithmetic* average used instead, the quantity maximized would be the total Bayesian posterior probability of all states in Q, which would increase monotonically with the size of Q and be maximized only when $|Q| \geq |T^P|$, the set of all possible ancestors of N^V. By contrast, Q as defined can contain no state with very low likelihood. Q will basically comprise the "core" set C of ancestral states with a good likelihood of evolving into N^V. However, its total posterior probability may be quite small.

Another way of regarding an MML hypothesis is that, of all hypotheses with a good likelihood, it has the largest "margin for error". That is, it represents a large group of similar hypotheses all of which have good likelihood, so it need not be specified very precisely. This property is shared by the true, lower entropy "Then" state which led to the present view. Simulations show that when a low entropy T evolves to a higher entropy partially ordered N^V, the evolution and the final view is insensitive to substantial changes in the microscopic detail of T, and to the detailed random events occurring in non-deterministic evolution.

In the real world, an explanation of a present view conforming to MML ideas would not normally assert a specific past state T. Rather, we would assert a view of the past roughly equivalent to the set Q, that is, a view asserting just enough detail about the past to ensure that any state compatible with that view would have a high likelihood of evolving into the present view. In general, we might also have to include in our explanation of the present some account of incidents occurring during the evolution whose occurrence was not implicit in the asserted past view. These additional details would form the second part of the explanation, and while not implied by the past view, should have reasonable probability given the past view.

8.8.3 The Uses of Deduction

I am not asserting that deduction of the past from a view of the present has no role. In some situations, our view of the present is essentially exact, the relevant physical laws are well-known, and no appeal need be made to statistical thermodynamic reasoning. The classical example is the deduction of past positions of major solar system bodies such as the planets, their moons, and the major asteroids. Their "laws of motion" are well-known, the bodies are so large that indeterminacy is irrelevant, and present observations of their positions and velocities are very accurate. There are no "dissipative" effects of much consequence. Hence, we can make reliable deductions of their past behaviour some thousands of years into the past.

In some situations where dissipation is not negligible, bulk motions apparent in the present view may still provide a basis for deducing at least some properties of the past. For instance, if the present view shows a spherical shock wave in the atmosphere, even a deterministic back-tracking of some state conforming to the present view will fairly well retrace the recent history of the wave, showing it to have been expanding from its centre with about the correct speed (although it will probably suggest that the intensity of the wave was less in the past than at present, see Figure 8.2).

If we are prepared to accept the well-founded inductive inference from many data sources that the Second Law and its irreversible fellows held in the past, we may accept them as premises in making deductions about the past. The instabilities which can arise in using these "laws" backwards are not necessarily fatal over short periods. If we accept our inductive inference that the memories and records which are part of our present view of the world give fairly reliable information about the past, deduction from this information may reliably imply other details of the past. However it still seems to be the case that deductive reasoning without reliance on inductive inferences about the past has at best a very limited scope.

Finally, deduction plays a vital role in supporting any inductively produced assertion about the past. We should accept the assertion only if we can deduce from it that the present view is significantly more probable than were the assertion not to hold.

8.8.4 The Inexplicable

An MML "explanation" message is considered acceptable only if it is shorter than the raw statement of the data. An explanation of a present view is acceptable only if it is more concise than the description of the present view. If no such explanation can be found, the MML principle suggests that the data is simply random with respect to the kinds of hypothesis we can assert about T, and no such hypothesis is supported by the data. Presented with such a present view, MML will offer no explanation, leaving only the conclusion that the present state is random with respect to the state at time $-k$. That is, we must conclude that N^V no longer contains any information about T.

Inexplicable present views are not ruled out by the fundamental laws of Physics. A simple inexplicable view of a box containing an ideal gas which has been undisturbed for an hour would be a view which showed a strong standing wave with a natural period of, say, ten milliseconds. There is no conceivable state an hour ago which could be expected to evolve into this view with reasonable probability, although of course there is a multitude of past states which could possibly do so.

The "inexplicable" is not confined to the virtually impossible. In fact, most of our day-to-day views of the world are inexplicable in this technical sense. When we see an oddly shaped cumulus, or the disposition of the stars in the night sky, or a collection of people in a shopping mall, we can usually offer no explanation of the view which does not require a set of assumptions as great as or greater than the features which they "explain". These everyday views are not views of a complete closed system, and we accept that what we see is largely the result of "novelty" entering our view from a wider sphere beyond the current reach of our senses. Even when we inspect an effectively isolated system, say, a rock in a glass box, we do not expect to be able to explain the fine detail of what we see, and just accept that each particular fleck of quartz is where it is for reasons which are beyond rational discussion.

Even if we find in a closed system some mild but significant departure from equilibrium, say, a small temperature gradient along a metal bar, the inference that the system had a greater disequilibrium in the near past, whilst perhaps the best possible guess and perhaps even true, is not really an "acceptable explanation" of the present. It requires us to assume a proposition at least as lengthy as the one it is trying to explain, and leaves us no wiser as to *why* the gradient is present.

We may at times be forced to make inferences about the past which are not acceptable explanations of the present, but only the best explanations which can be found. In such cases, it seems plausible to argue that we should still follow the inductive principles which would have led us to an acceptable explanation were one possible. Acceptable explanations are most likely to be found when the present view comprises a number of features which are logically independent of one another (i.e., each could reasonably exist in the absence of the others). Then an inferred past describable by fewer features which implies a high probability of the co-occurrence of the observed features will almost certainly be acceptable. Each observed feature acts as corroborative evidence of the inference.

8.8.5 Induction of the Past with Deterministic Laws

A case has been made that inductive reasoning from present data using MML or some similar inductive principle can be expected to reach inferences about the past agreeing with the inferences we normally accept, whereas deductive reasoning, including probabilistic deduction, will not. The argument has so far assumed non-deterministic fundamental laws. In this section, I consider whether MML induction will behave similarly if the fundamental laws are known and deterministic. I assume some inexact partially ordered view N^V of the present, and suppose it to have evolved after k time steps from some unknown state T which common sense suggests will have lower entropy and greater order than the present.

Given an inexact view N^V of the present state, we can in principle proceed to enumerate the ancestors at $t = -k$ of all the states in N^V, e.g., by backtracking, and thus arrive at a set T^P of the possible states at time $-k$. We then know that the true state T is some member of T^P, but not which one. Since the laws are deterministic and reversible, the cardinalities of T^P and N^V are equal. This is an essentially deductive mode of reasoning, applying to the past the kind of reasoning used to predict the future. In the task now considered, it seems, while perfectly valid, to be of little help.

The vast majority of states in T^P bear little resemblance to T. In particular, the vast majority will, according to our argument and simulation results, have entropies greater than the entropy of the present state N, which itself has entropy greater than T. Standard statistical reasoning also does not advance us. All states in T^P have the same likelihood (namely one) of yielding the observed "data" N^V, so the maximum likelihood principle will not help us find in T^P any clue about T. A Bayesian approach (taking all states in the state space to have the same prior probability before the "data" are considered) leads to a uniform distribution of posterior probability over the members of T^P, since all have the same likelihood. Thus, no state in T^P can be preferred on the basis of posterior probability, and the mean entropy over the posterior distribution is a terrible estimate of the entropy of T.

MML also does not seem to work in this case. Since all states in T^P have the same likelihood, choosing Q to minimize L_Q will result in $Q = T^P$ (see Section 8.8.2), and a random selection of a past state from Q would almost certainly select a state of higher entropy than N.

An alternative approach trying to avoid the equality of likelihoods would be to aim to infer a past view rather than a past state. If a past view V is asserted, one could identify its prior probability with the total prior probability of past states compatible with it, which is proportional to the number of states in the view, and identify its likelihood with the fraction of states in the view which would evolve into a state in N^V. This does not work either, as maximization of the posterior probability of V would lead to a view including all of T^P.

There is one inductive scheme which would probably yield a "commonsense" inference. Let the state space be divided into "cells" each containing a large number of states, such that all states in the one cell differ only microscopically and would present virtually identical macroscopic views. Then we could apply MML or some similar inductive principle to the inference of a past cell, in effect replacing the set of states by the set of cells as the conceptual hypothesis space. The prior of each cell would be proportional to the number of states it contained (which might be equal for all cells), and its likelihood given N^V would be the fraction of its states which will evolve into a state in N^V, i.e., the fraction of its states which are in T^P. The approximate MML approach used above for non-deterministic laws would then be expected, given N^V, to generate a set Q of cells all having a high likelihood and all containing states of lower entropy than N. In effect, the "cell" scheme replaces states with zero or one likelihood by cells with a full range of likelihoods, and introduces a form of indeterminacy by the back door. The problem with this idea is that the results would depend somewhat on the choice of cell size and, more importantly, on the criteria of similarity used in determining which states can be put in the same cell. These criteria will depend on the nature of the observations and measurements used in defining a "view". There will be states which are indistinguishable by some measurements but revealed as different by others. Given the subtlety and precision of the measurements afforded by modern Science, it is hard to be confident that any similarity criteria could be defined which could not be violated by some form of view-forming measurement. This problem makes me unwilling to accept the scheme as general and well-founded.

I have been unable to think of any general, principled method of induction which can be expected to make a "commonsense" conclusion in the presence of deterministic fundamental laws. While my failure is no proof that no such method exists, I am at present inclined to think that principled inductive inference of the past will yield commonly accepted results if the fundamental laws are reversible, but only if they are also non-deterministic.

One could argue that in practice, the evolution of partially ordered states under deterministic laws seems, in simulations, to be virtually indistinguishable at macroscopic scales from evolution under non-deterministic laws. Hence, we might as a pragmatic concession pretend that we do not know the deterministic laws, and make the kind of inference, MML or whatever, which would be justified by non-determinism. While this might usually work, it scarcely seems principled.

8.9 Causal and Teleological Explanations

The nouns "cause" and "effect" presumably enter our language because they are useful in some areas of communication. They are human constructs which like "mass", "electron" and "energy", do not refer to things directly accessible to our senses, but which, if hypothesized to apply to certain phenomena, allow what we can observe of these phenomena to be encoded concisely and used in forming coherent mental models of phenomena having some predictive power. As with these other terms, their meaning is ultimately defined in the sets of sentences in which they appear and which can be "decoded" by others to convey valid accounts of observations or verifiable assertions about future observations. Their use is, of course, not limited to just these sentences, but it is these which finally must be appealed to if one wishes to claim some use of the terms to be incorrect or meaningless.

Some such terms, like "mass" and "electron", have proved to be enormously useful in sentences with a mathematical form. By establishing just which mathematical sentences do decode to match observations, it has been possible to define a "calculus" for their meaningful use, i.e., to frame well-defined rules governing which uses are "correct". This does not imply that their meanings are immutable or even uniformly understood by all those who use them. Our current uses of the term "mass" might surprise Newton, while embracing all the uses he might have made. Further, some terms, such as "attraction" (as in gravity or electrostatics) are now considered to lack an exact calculus, action at a distance having been divorced from our canon of natural law, but remain useful in slightly less formal discourse which is content to approximate reality.

In my opinion, "cause" and "effect" may suffer the same demotion as "attraction", and we may have little better hope of defining an exact calculus for their use than we have of defining calculi for "sin" or "beauty". That said, I will try to add my tuppenceworth.

In an account of system histories in which dissipation plays no significant role, there seems little need to invoke notions of cause and effect. The entire history is summarized by a sufficiently detailed description of its state at any chosen time in the history, and the earlier and later states are implicit in this description. Given the positions and velocities of the planets yesterday, asking for the cause of their positions today is a question which, no matter how answered, leaves us no wiser. This applies even if there is some indeterminism in the system.

Since a reductionist history of a system to the most fundamental level relies only on the reversible fundamental laws, dissipation is not a necessary part of the history. The elastic collision of a pair of atoms is not a dissipative event, although its outcome may be uncertain, and the result of a multitude of such events may be described as dissipative at the macro level. So, at the fundamental level, there is little role for cause-effect language. In fact, notions of cause and effect simply do not appear in the expression of fundamental physical laws (and very rarely even in macroscopic laws).

I am led to think that these notions have their use in discourse concerning the macroscopic behaviour of dissipative systems, and are most likely to be invoked when we have only a partial understanding of the processes involved. Consider an acceptable explanation of the current situation which hypothesizes a past situation and encodes the present in terms of its probability of evolving from the past. The system of interest need not be closed: part of the encoded evolution path may further hypothesize some external intervention. For the explanation to be acceptable, i.e., shorter than a plain statement of the present facts, the evolution from past to present must be highly probable, and hence, if dissipative, must follow macro thermodynamics. This fact allows the encoded history of the asserted evolution to be encoded using the macro laws to define the trajectory probabilities from state to state, and allows the evolving states to be described in purely macroscopic terms. We need not even know the underlying fundamental micro laws, since the macro behaviour depends little on the micro details. I will call such an acceptable explanation a Causal explanation, and suggest that any proposition asserted in the explanation may be regarded as a cause of the present situation.

If the explanation explains the extinction of the dinosaurs, the postulated past climate and modes of life of the dinosaurs are "causes" (of why dinosaurs but not crocodilians copped it), the arrival of a comet or asteroid is a cause, the resulting global winter was a cause, the postulated presence of sulphate rocks under the impact was a cause (of global acid rain) and so on.

Even if the explanation is a generalized explanation which asserts a new "law", that law may have the status of a cause. When a postulated universal gravitational attraction was first used to explain planetary motions, it would have been natural to think of the gravitational pull of the Sun as causing the planets to accelerate towards the Sun. If such a postulated law becomes generally believed, it perhaps can no longer be regarded as a cause: it is simply a statement of the way things are and always will be.

Returning to the main theme, it is not logically impossible for the present state of a closed system to admit an acceptable (i.e., short) explanation in terms, not of evolution from some more-ordered past state, but of devolution from some more-ordered future state. Such an explanation, which I will call Teleological, asserts that some state F (or view F^V) will hold at some time $t = f$ in the future, and then shows that with high probability, all trajectories leading to F will show at the present time states much like the present view N^V.

An acceptable Teleological explanation of the present has the same logical structure as an acceptable Causal explanation, and receives as much support from theories of inference such as MML as does a Causal explanation. However, few Teleological explanations of (parts of) the present seem to be adopted. It may be that the present state of the Universe is such that, while many of its parts admit of acceptable Causal explanations, few if any admit of acceptable Teleological ones. If this conjecture is true, it describes an empirical property of the present: an inherent asymmetry which we might call a "causal arrow of time".

There is a more direct reason to expect Teleological explanations to be shunned. Recall that, in an acceptable Causal explanation, the asserted "Then" state or view has a high probability of evolving to a state in the present view N^V, but, as we have seen, states in N^V have very low probability of devolving into T. This fact, while perhaps surprising, is not embarrassing. The present does not devolve, so we cannot observe the divergence of the devolving trajectory away from the trajectory leading to T, and so cannot use this divergence to disprove the asserted "Then". The corresponding position in an acceptable Teleological explanation is that the asserted future state F has a high probability of devolving into N^V, but states in N^V have very little probability of evolving into F. Thus, we need only watch N evolve for a little while to observe (with very high probability) that the evolution of N has diverged from the path to F. In other words, a Teleological explanation of the present, even if acceptable, must be expected to suffer empirical disproof.

8.10 Reasons for Asymmetry

My main case is that we use deduction to reason about the future, but usually use induction to reason about the past. If it is accepted, it poses the question of why we use different reasoning for past and future. A crude answer might be that we then avoid ever having to reason deductively in the backwards time direction, thereby never exposing our reasoning to Boltzmann's theorem. About the future, we deduce forwards in time from the present. About the past, we postulate a past by whatever means, then check the postulate by again deducing forwards in time from past to present to see if the postulated past implies something close to the present. Similarly, we avoid Teleological induction of the future because we would have to check the postulated future by deducing backwards in time from the postulated future to see if it accounts for the present.

More seriously, I suggest we reason differently in the two time directions because we have different interests in the past and future.

Our interest in the future lies in the possibility of adapting our behaviour to benefit from or at least cope with the future. If we can guess what is likely to occur, we can perhaps do something about it. If we have no possibility of effective action, we might as well be fatalistic about the future: whatever will be, will be, so why bother to guess? If indeed our interest in prediction comes from our possibility of choice of action, the deductive modes, and in particular probabilistic deduction, are appropriate. They lead to at least a rough probability distribution over future states, conditional on our chosen acts. If we can evaluate the possible future states on our own scale of values, we can choose the acts giving the greatest expected benefit.

Our interest in the past is quite different. The past is gone: there is nothing we can do to alter it and whether the possible past states were good or bad on our scale of values, the good or bad has been done. There is no immediate value in knowing the past as far as choosing our future acts is concerned. Rather, our interest in the past is in helping us to understand the present.

An inductive inference about the past must, when deductively evolved to the present time, show a view in good accord with the view we actually have of the present. If we find an acceptable explanation, we have not only made a good guess about the past but now have a more concise account of the present. But it can do more. It may imply details of the present state which we did not notice in our observations. Thus, induction of the past may improve our view of the present. Of course, since no inductive conclusion is certain, we would be well advised to check these implications by more observation of the present before accepting them, but at least the induction has suggested what to look for. If we find the implications to be verified, we gain confidence in the inductive conclusion; if we do not, we have a refined present view to use as data for a better inference of the past.

Our understanding of the present may be improved more radically. If we consider an inference of the past chosen because it gives a short MML-style explanation of the present, there is nothing to prevent our hypothesis (in MML the first part of the explanation) from asserting a hitherto unknown natural law, the assumption of which, together with a hypothesized past state, provides a shorter explanation of the present than can be found using previously known laws. For instance, in our ideal gas examples, we might assert a new, more detailed form of the non-deterministic reversible collision laws which imply some subtle change in evolution probabilities giving a more probable evolution to the present. This may be the route by which most scientific advances are made.

As to why we should want to understand the present, the motives are presumably based on the possibility of choosing future actions. The better we understand the present (and the laws governing the processes by which the present state was reached) the better we are able to predict the future.

In summary, our deductions discover the direction of the point of the arrow of time, and validly imply a high probability of increasing entropy in the future. Similar deductions would imply with equal validity a high probability that entropy was higher in the past: a second point at the other end of the arrow. But we have no real interest in *deducing* the past: this can tell us nothing more than is implied in the present. Instead, we seek in the past an explanation of the present, and must resort to induction to do so. The feathers on the arrow of time are our inductive inferences, unprovable but falsifiable.

Deductive and inductive inferences about the past state of a closed system can be very different. When they differ, which should we believe? I suppose my answer must be that it does not matter what we believe, since the past is truly lost, and all we can ever know of the past of a closed system is what we know of its present. It is what we believe about the present which matters. This is not to say that inductions based on the present view of a closed system are pointless. They may lead to the elucidation of features of the present otherwise unnoticed or unexplained, an understanding of the present as a partially coherent pattern rather than a collection of undigested data, and in some cases to the discovery of general laws. These benefits do not seem to arise from deductions about the past state.

Despite this, if presented somehow with a situation where I were required to estimate the entropy of a past state, and would suffer a loss quadratic in the error of my estimate, I am not sure I would reject the deduction based estimate that the entropy was higher than it is now. Note, however, that such a situation cannot arise if the system was genuinely isolated during the relevant time. For me to suffer a loss because of error, the actual past value must still be exerting an influence over future events otherwise than through the present state of the system. That is, the system cannot have been totally isolated, and its past state has interacted with my environment.

8.11 Summary: The Past Regained?

A case has been made for a distinction in the modes of reasoning normally employed regarding the past and future of closed systems. We use deductive modes for predicting the future, but seem to switch to inductive reasoning to reach conclusions about the past. Deductive reasoning may be employed in forming a view of the past, but I believe it is fair to say that this is usually done only in situations where dissipation is insignificant, or when the deduction uses as premises the assumption of irreversible laws whose validity in the past can only be supported by inductive reasoning.

If the fundamental laws are assumed to be reversible and non-deterministic, no loophole is evident in Boltzmann's theorem that both the past and the future of a partially ordered state are probably less ordered than its present. However, with just the same assumptions, a well-founded induction of the past is likely to conform with our commonsense expectation that the past was more ordered. Further, if a past-based explanation of the present is allowed to assert not only a view of the past but also a new or refined “law”, inductive explanations may involve and justify the induction of new natural (or social, biological, economic, etc.) laws.

Under the same assumptions, Teleological explanations of the present, based on the induction of a future state, while logically possible, will almost certainly be disproved.

I have not been able to show that these conclusions follow if the fundamental laws are deterministic. On the other hand, I have not shown that the conclusions are incompatible with reversible deterministic laws.

Our different interests in past and future, based on the possibility of our effective action, are sufficient to motivate the switch of reasoning mode, and hence a subjective asymmetry in the Arrow of Time: deductive point at one end, inductive feathers at the other. Objective evidence may be found in an objective asymmetry in the present, if we find (as I believe to be the case) that the present admits acceptable explanations based on the inference of past states more often than it admits acceptable explanations based on the inference of the future.

8.12 Gas Simulations

Readers who are happy to accept the proposition that partially ordered states are most probably preceded and followed by less ordered states need pay little attention to the simulations, since they serve essentially only to establish the proposition in a very limited context.

The system imitates a two-dimensional ideal gas confined in an immovable rectangular box. In some experiments, internal divisions divide the box into two or more connected chambers. In order to implement exactly reversible “laws”, all arithmetic in the simulation is done with integers. Floating-point arithmetic is in general irreversible because of rounding-off errors. Thus, the atoms of the gas have X and Y positions which are integers, so the space within the box is a rectangular grid of points. Similarly, atom velocity components U (in the X direction) and V (in the Y direction) are restricted to integer values. As assumed throughout, time is represented by an integer t.

In the simulation of a time step from t to $(t + 1)$, a two-phase process is used. First, at each point of space occupied by two or more atoms, collisions are simulated using a Collision Table which defines the collision physics and which is set up at the beginning of the simulation program. The several atoms at a point are collided in pairs, in order of their identifying indices. The first and second atoms collide, then the second and third, then the third and fourth, and so on. The collision between two atoms effects a rotation of their relative velocity vector, their centre-of-mass motion being unchanged. If (u_1, v_1) is the initial relative-velocity vector, the new vector

(u_2, v_2) is determined from the collision table. The collision table ensures that $(u_1^2 + v_1^2) = (u_2^2 + v_2^2)$ to conserve energy. The integer-velocity-component restriction also requires the collision table to preserve the parity (oddness or evenness) of u and v, which ensures that the atom velocity components remain integers. The table is also set up so that

$$\text{if } (u1, v1) \to (u2, v2) \text{ then } (-u2, -v2) \to (-u1, -v1)$$

For convenience, the table also obeys the further condition

$$\text{if } (u1, v1) \to (u2, v2) \text{ then } (u2, v2) \to (u1, v1)$$

Note that these conditions on atom collisions do not imply that if for some states r and q, $r = M(q)$ then $q = M(r)$. To limit the size of the collision table, a collision which would result in one of the atoms having an absolute velocity component exceeding a fixed limit is not permitted. The limit is high enough that it is rarely reached.

The second phase of the time step moves the atoms according to their velocities. In general, the (x, y) coordinates of an atom are replaced by $((x + u), (y + v))$, except that if the new position would lie outside the box, one or more elastic collisions with the walls are simulated to bring the new position within bounds.

To simulate internal compartments or other fixed features within the box, it is possible to place "obstacles" at grid points within the box. Collisions among atoms moved to an obstacle point are inhibited. Instead, during the collision phase, atoms at an obstacle have their velocities altered in a way depending on the type of obstacle. The most useful obstacle simply reverses the velocity of any atom landing on it, so the next move phase will return the atom whence it came. An internal wall within the box can be constructed with a thick slab of such obstacles. The slab acts as a "soft" wall: an atom with a velocity component perpendicular to the slab of, say, u, may penetrate into the slab by up to a distance $(u - 1)$, since in move phases, atoms move from place to place without regard to intervening points (except for the box walls). Thus, a soft wall of thickness W may be passed through by an atom of velocity component perpendicular to the wall exceeding W. In most uses, the slab is made thick enough to reflect virtually all atoms. Soft walls are "rough": they reverse both velocity components of impinging atoms.

"Hard" walls like the bounding walls of the box have been simulated in one set of experiments to set up two rectangular chambers connected by a long, thin pipe. Hard walls are "smooth": collision with a hard wall does not affect an atom's velocity component parallel to the wall.

Given the above description, it is clear that the simulated gas is governed by deterministic, reversible laws. Both forward and reverse state sequences through time can be followed. The only difference between a forward time step and a backward step is that in the latter, the move phase is done before the collide phase, and subtracts velocities from positions. The asymmetry between the "collide-move" forward step and the "move-collide" backward step does not represent a time asymmetry in the gas "physics". It arises from the arbitrary decision to define the integer "time instants" as occurring just before, rather than just after, a collision phase. The physical progression simulated is

```
      Earlier ..... C M C M C M C M C M ..... Later
  Defined times ... T---T---T---T---T---T
```

It would of course have been possible, and exactly equivalent, to step back in time by reversing all velocities, proceeding as for stepping forwards, and finally re-reversing velocities. The introduction of non-determinism is described later.

8.12.1 Realism of the Simulation

Many tests have been done to see if this simple simulation resembles a thermodynamic system. The tests typically began with the gas in a partially ordered state with individual atom positions and velocities chosen pseudo-randomly in accordance with a specified macro view. The evolution of the system was then followed for some hundreds or thousands of time steps while tracking a physical quantity of interest.

In some experiments, the total gas energy was allowed to change by introducing a "heater" obstacle which in a collision phase, adds $+1$ to the V component of any atom at the heater's position if the time is even, or -1 if the time is odd. This action is deterministic and exactly retraceable. The effect of a heater is weakly to couple the gas to a heat source of infinite temperature, since the only velocity component distribution unchanged by the heater is the Uniform distribution.

Pressure was measured by accumulating the momentum transfer between one wall of the box and the atoms bouncing off it. A useful property of a gas state is its velocity component distribution. This was formed in a histogram, adding both U and V components to the same histogram since (at least near equilibrium) the two distributions should be similar and nearly independent. A "state entropy" was calculated from this distribution, which ignores the entropy of the spatial distribution of atoms (which, near equilibrium, is expected to be uniform) but serves as an indicator of the true entropy.

A visual picture showing the positions of all the atoms could be displayed at chosen intervals, and plot files written giving the history of chosen macro quantities such as entropy, temperature, etc.

Experiments showed the following properties of the gas:

– Near equilibrium (i.e., after sufficient time steps to erase any initial order), the velocity component distribution matched the expected (discrete) Maxwell-Boltzmann distribution, and the state entropy approached its maximum possible value.

- When the collision laws were modified to allow collisions only when exactly two atoms occupied the same position, the effect is to make the laws independent of the identities of atoms, and to prevent any two atoms ever having the same position-velocity state at the same time. The atoms should then behave as fermions rather than classical particles. When this was done, the velocity distribution matched the Fermi-Dirac distribution. As the difference is only apparent at very low temperatures, and the modified simulation was slower to approach equilibrium, all further experiments used the Maxwell-Boltzmann form.
- Temperature was measured by a maximum-likelihood fit to the velocity distribution. This measurement is meaningful only near equilibrium.
- The gas obeyed the classical gas law

$$Pressure \text{ proportional to } Density \times Temperature$$

- The Specific Heat at Constant Volume had the expected value.
- The speed of sound, as measured by oscillations in a pipe, had the expected value

$$Speed = \sqrt{\frac{\gamma \times Pressure}{Density}}$$

  where γ, the ratio of specific heat at constant pressure to specific heat at constant volume, has the value 2 for an ideal 2-dimensional gas.
- In an experiment where the gas is slowly heated by an obstacle at infinite temperature, the increase in entropy followed the classical law

$$\Delta Entropy = \Delta Heat/Temperature$$

- When the box was divided in two by a soft wall of thickness W and the initial state had all the gas on the left side of the wall with a mean speed much less than W, "hot" atoms slowly diffused to the right side, with a consequent initial cooling of the remaining gas on the left, but eventually the densities and temperatures of the gas on both sides became equal, with a temperature equal to the initial temperature. No quantitative analysis of the expected ideal gas diffusion behaviour has been done, but the simulation was at least qualitatively as expected.
  Similar qualitative realism was found when the box was divided into two chambers connected by a thin pipe, all with hard walls.
- Whatever the nature of the confinement or initial order, every simulation showed eventual convergence to an equilibrium view.

8.12.2 Backtracking to the Past

Whatever the details of the simulated situation, running for a long time showed convergence towards equilibrium. When the simulation was then reversed, the final state always returned exactly to the initial state, showing that the simulation indeed followed reversible deterministic laws. Further, if several runs were made with different initial states randomly chosen in accordance with the same initial macro view, the final states reached showed no significant macro differences. That is, the evolution of the macro properties of the gas is insensitive to initial differences on a micro scale.

If, however, the final state was perturbed by moving the position of a single randomly chosen atom by a single grid point, the reversed simulation invariably returned to an initial state with less order than the final state. In some experiments, the macro history running backwards would follow the true history fairly closely for 50 or so time steps before diverging towards a less ordered past, but divergence always occurred.

A simulation was done with a 10,000 atom system in a 200-by-50 box, where the initial "Then" state was in equilibrium with entropy close to the maximum possible, a uniform random distribution of atom positions, and a well-defined temperature of 2.0 (taking Boltzmann's constant as one.) Two "heater" obstacles were placed in the box, weakly coupling the gas to a high temperature. The simulation was run forward for 1000 steps to yield the "Now" state. As expected, the gas slowly heated up. At all times the gas had a well-defined temperature, which gradually increased from 2.0 in "Then" to 2.113 in "Now". Reverse simulation was then done from "Now" back to "Then". Starting with the exact "Now" state, deterministic devolution exactly recovered the "Then" state. Starting with the "Now" state with one atom shifted by one grid point, deterministic devolution followed the true temperature trajectory for about 50 steps, but thereafter showed an increase in temperature into the past, reaching 2.216 at the "Then" time. Non-deterministic devolution diverged from the true path almost immediately, reaching 2.229 at "Then". In both the latter cases, the backtracked "Then" had higher entropy than "Now", and of course higher than the true "Then".

8.12.3 Diatomic Molecules

An elaboration of the simulation imitates a gas such as hydrogen, in which pairs of atoms may combine with the release of energy to form diatomic molecules. In the simulation, the collision table has special entries which allow a pair of free atoms colliding with certain relative velocities to join, forming a molecule. The molecule then moves as a point unit according to the mean velocity of the atoms, and occupying a single grid position. The atoms retain individual velocities, their relative velocity cycling through a cycle of values under the control of another table, simulating a high-frequency internal vibration of the molecule involving a continual exchange between the kinetic energy of the atoms' relative motion and the potential energy of their bond. The total energy of a newly formed molecule exceeds the original total energy of the free atoms by a fixed "binding" energy. If nothing intervenes, the molecule will split into free atoms when its vibrational cycle returns to the state at which the molecule formed. However, collisions of its atoms with other free or bound atoms may alter its vibrational state, with the gain or loss of vibrational energy. If the molecule's vibrational energy falls below the binding energy, it cannot split.

A cool gas of free atoms may exist for a long time in a metastable state, since few of its collisions may reach the lowest relative velocity needed for combination. This threshold energy and the binding energy can be specified independently. If a small body of fast atoms enters the gas, triggering combinations, the local temperature may rise triggering more, and a simulated fire front or, in some cases, detonation shock wave appears.

The inclusion of this elementary "chemistry" into the simulated gas made no difference to the kinds of behaviour on which this study of the inference of the past are based. When the simulated laws are deterministic, the simulation remains exactly reversible, and a past state can be recovered exactly from an exactly known present state. Almost all partly ordered states evolve and devolve to less ordered states. When the present partly ordered state has evolved from a more ordered past state, exact retracing recovers the more ordered past state, but if retracing is attempted from an inexact view of the present, differing from the true present state only by a tiny perturbation, retracing almost always yields a past view less ordered than the present. With non-deterministic laws, virtually all attempts to recover the past state yield a less ordered view than the present.

8.12.4 The Past of a Computer Process

The reader may be forgiven for doubting the realism, and hence the relevance, of these simulation studies. A two-dimensional ideal gas in a quantized Newtonian space-time is a poor substitute for a real closed system. However, the simulation *is itself a real system with real conservation laws and real state transitions* even if it is only a pattern of data digits in a computer. It should be a very simple sort of system to reason about, since its state at any time step is very well defined and easily examined, its laws of behaviour are exactly known if one likes to inspect the program, and it is completely isolated.

The skeptical reader might ponder the following problem: given knowledge of the state of this real system after some thousand time steps, and given that the present state of the data arrays shows some order, what can he or she *deduce* about their state 500 steps ago? Given exact knowledge of the present data, there is no problem if the computer is not using an inaccessible random number source to simulate indeterminacy, but if it is, or if there are one or two tiny errors in the report of the present data, the results discussed here show that *deduced* past states (or probability distributions over past states) will almost certainly be misleading at least as to the degree of order in the past. If this real, but very simple, isolated and fully understood system presents deduction of the past with real trouble, can we expect to do better in the larger world?

8.13 Addendum: Why Entropy Will Increase (Additional Simulation Details)[1]

If we are presented with a closed system which appears to be in equilibrium, there is nothing which can be usefully predicted of its future, save that it will probably remain in equilibrium. We can hope to make interesting predictions about the future of a closed system only if it is now partially ordered, i.e., has entropy significantly below the maximum. I will now argue that we may confidently predict that its entropy will increase. Several cases will be treated.

(a) Deterministic Laws, Exact View.
Assume the Laws of Physics are deterministic, and that we have exact knowledge of the present state $s(0)$ at time $t = 0$. Suppose $s(0)$ is a state of low entropy H_0. The deterministic laws allow us to deduce the state $s(f)$ at some future time $f > 0$. The repeated application of the time-step mapping $s(t+1) = M(s(t))$ will no doubt have changed the state and its entropy. Since the mapping is based on the microscopic laws, which make no reference to entropy and are reversible, there is nothing inherent in the mapping which can lead us to expect that the entropy of a state will be related in any obvious way to the entropy of its successor, although, since a single time step will usually produce only a small change in the macroscopic view of the system, we may expect the resulting change in entropy to be small. After the f steps leading to time f, therefore, we can expect the initial entropy to have changed by some amount, but otherwise to be almost a random selection from the set of possible entropy values. Recall that in a closed system, the number of states of entropy H is of order $\exp(H)$. There are therefore far more high entropy states than low entropy ones, so the chances are that $s(f)$ will very probably have an entropy greater than the initial state $s(0)$. In a deterministic closed system of constant energy, the sequence of states from $t = 0$ into the future is a well-defined trajectory through the adiabatic state space at that energy. The trajectory may eventually visit all states in the adiabatic space, or only some subset. Given that the state space is discrete, an adiabatic space can only contain a finite, albeit huge, number of states. The trajectory from state $s(0)$ thus must eventually return to state $s(0)$, and since the mapping is one-to-one, must form a simple cycle in which no state is visited more than once before the return to $s(0)$. Since it is cyclic and produced by time reversible laws, the trajectory must show just as many steps which decrease entropy as increase it. However, because the trajectory can be expected to visit far

[1] The editors included this section for completeness. It contains material from Section 8.4 together with material on simulations which appeared in previous versions of the chapter. Chris Wallace removed these details to avoid distracting readers from the main argument. However, the editors felt that his careful simulations were interesting in their own right.

more high entropy states than low entropy ones, if we start from some low entropy state and advance for some arbitrary number of steps, we must expect to arrive at a state of higher entropy. If this argument is accepted, the Second Law will in general apply even given deterministic, time reversible laws. No paradox is involved.

(b) Simulation of Deterministic Laws, Exact View.

The two-dimensional gas simulation has been used as a model system to check the above argument. Although it simulates a very simple system, and has no provision for important real-world processes such as gravity, chemical bonding, electromagnetic interactions, etc., it is not obvious that the argument is sensitive to such phenomena, so the simulation should give some useful indication of the validity of the argument.

Some hundreds of experiments have been conducted in which a partially ordered (sometimes very highly ordered) state was set up in the gas and allowed to evolve. As it is infeasible directly to compute the state entropy as defined above, the simulator computes a simple but useful surrogate. For each state entered, a histogram of atom velocity components is collected. For each atom, both the U and V velocity components are entered in the histogram, as these are expected to be statistically independent degrees of freedom.

The entropy of the velocity component distribution is then computed as

$$H_v = \sum_v -n[v] \log(n[v]/N)$$

where $n[v]$ is the histogram count for component velocity v and N is the total of all counts, which is of course twice the number of atoms.

H_v is a measure of one component of the state entropy, but does not include any spatial-distribution entropy. However, it suffices as a measure of disorder in the velocity distribution of the state. Note that the number of states having a particular value of H_v is approximately proportional to $\exp(H_v)$.

At each time step, the simulator records on a file the time t and the per-freedom entropy $G_v(t) = H_v/N$. This allows the G_v history to be displayed as a plot versus time. Note that G_v may display distinct oscillations as time goes by, superimposed on a general increase, as G_v shows only one component of the system entropy, and in some situations, e.g., an initial standing wave in the gas, there is a periodic exchange of entropy between velocity and spatial components. Such oscillations tend to decrease as the system approaches equilibrium.

Further, the simulator can display a picture of the spatial positions of the atoms at chosen times, allowing a qualitative assessment of the behaviour. Other properties of the gas can also be recorded, such as the gas pressure, total momentum, and an estimate of temperature. The temperature T is determined by choosing the value which gives the best fit between the observed velocity component histogram and the equilibrium Maxwell-Boltzmann distribution $\Pr(v)$, which is proportional to $\exp(-v^2/(2T))$ (our "atoms" have mass 1).

Many forms of partial order have been simulated, some in boxes containing barriers. Situations simulated include the passage of gas through a channel from a chamber of high initial pressure to one of low initial pressure, the oscillation of a standing sound wave, the impact of a ball of cold, dense gas on a body of cold, low density gas, the diffusion of fast atoms over an "energy barrier", the exposure of the gas to an external "hot spot", and many others. To summarize the results, in every case the entropy of the system increased, finally (perhaps after some tens of thousands of time steps) becoming close to the equilibrium value, but fluctuating slightly from step to step. Qualitatively, the pictures of the gas showed the gradual decay of the initially visible order.

Quantitative results were also encouraging. The final states showed velocity distributions conforming to Maxwell-Boltzmann statistics. The variation of pressure with density and mean atomic energy was as expected. In the sound wave simulation, the measurement of the standing wave frequency yielded an estimate of sound velocity. This allowed estimation of "gamma", the ratio of specific heat at constant pressure to the specific heat at constant volume.[2] The estimated value was 2.02, compared with the theoretical value of 2 for a two-dimensional ideal gas.

In one batch of experiments, the collision law was altered to forbid collisions at a point unless exactly two atoms occupied it. This restriction removes the effect of atom indices from the collision phase, with the result that the atoms are no longer identifiable, and makes it impossible for two atoms to have the same position and velocity. Thus, the atoms become fermions rather than classical. The equilibrium velocity distribution was then found to conform to Fermi-Dirac statistics rather than Maxwell-Boltzmann.

Overall, no simulation of the future of a partially ordered initial state showed any macroscopic behaviour at variance with thermodynamic predictions. In particular, the invariable increase in state entropy confirmed that the Second Law could apply to a deterministic system governed by time reversible laws. Turning to Boltzmann's theorem that velocity distributions tend towards high entropy, note that his assumption of no correlation between the initial velocities of colliding particles seems not to be essential. In many of the simulations, it can be shown that most of the collisions occur between atoms which have collided before. If collisions produce correlations in the resulting atom velocities, these correlations must certainly have been present. However, the simulations agree with the conclusions of the theorem: the velocity distributions did tend towards higher entropy, and did approach the Maxwell-Boltzmann equi-

[2] See Section 8.12.1.

librium distribution despite whatever correlations in atom velocities may have been induced in earlier collisions. (It is perhaps worth noting that, in a collision between two atoms whose initial velocities have been independently sampled from the Maxwell-Boltzmann distribution, the final velocities are not statistically correlated.)

(c) Deterministic Laws, Inexact View.
If, rather than knowing the precise initial partially ordered state $s(0)$, we have only an imprecise view of the initial state, the argument for expecting entropy to increase can only be strengthened. We may well feel more justified in regarding the entropy at future time f as being in some sense a random selection from a range of values, since we no longer have any certainty about the future state $s(f)$. Given this uncertainty, the larger number of high entropy than low entropy states becomes perhaps a sounder basis for expecting $H_f > H_0$. Because each time step has but a small effect on the macroscopic view, we also expect the macroscopic behaviour of the system to be insensitive to the fine detail of the initial state, and to be well predicted by a macroscopic, or at least inexact, initial view.

(d) Simulation of Deterministic Laws, Inexact View.
The above expectations were tested by simulations of two sorts. In the first sort, a "true" partially ordered initial state $s(0)$ was defined and its evolution followed to a future time f. Then several inexact copies of $s(0)$ were made, each resulting from a change of plus or minus one in a randomly chosen position coordinate of a randomly chosen subset of the atoms, the fraction of perturbed atoms being specified. The evolution of each copy was then followed to time f and compared with the results from the "true" $s(0)$. The copies showed macroscopic behaviour and entropy histories very close to those of the true state, but microscopic comparison of the final states showed great differences in the positions and velocities of the atoms.
In the second sort, the system was set up to have two chambers linked by a narrow channel. The initial view specified only the densities and mean atom energies in the two chambers. Several initial states were constructed consistent with this view, with the atoms in each chamber having random, high entropy spatial and velocity distributions meeting the different specifications of the two chambers (see Figure 8.3). The evolutions of the several initial states were then compared, and again showed no significant differences at the macroscopic level.
In these two-chamber simulations, the state entropy was calculated by forming separate velocity-component histograms for the two chambers and adding the entropies calculated from each histogram. This approach takes into account the partial order which is shown by different densities and/or temperatures in the two chambers, and gives a lower value than given by the overall velocity distribution if such partial order exists.

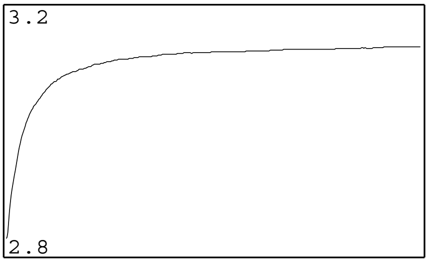


Fig. 8.4. Two-chamber + pipe model (from Figure 8.3); velocity entropy increase for 10,000 time steps from original "Then" state; non-deterministic laws.

(e) Nondeterministic Laws.
If we suppose the fundamental microscopic laws to be nondeterministic (but still reversible) the state-to-state mapping $M()$ is replaced by a probabilistic law

$$\Pr(s(t+1) = r) = M(r|s(t))$$

Given $s(0)$, we can at best compute a probability distribution over the possible states at time $f > t$, and are the more justified in expecting $s(f)$ to have higher entropy than $s(0)$.

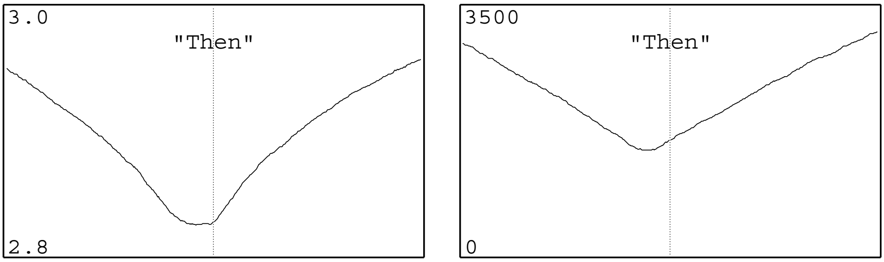
(f) Simulation of Nondeterministic Laws.
The simulation was modified to construct two different collision tables for effecting collisions. Both observed the constraints of time reversal, relative velocity component parity conservation, energy conservation, etc., described above. In collision phases, each pair-wise collision is effected by one or other table. The choice is determined by a conventional pseudo-random number generator of cycle length 2^{32}, and chooses each table with equal probability. The generator is a mixed congruential generator of 32-bit integers:

$$r[n+1] = ((69069 * r[n]) + 31415927) \text{ Modulo } \ 2^{32}$$

The inverse relation is

$$r[n] = (-1511872763 * r[n+1] + 670402221) \text{ Modulo } \ 2^{32}$$

which allows the simulation to be run backwards by retracing the sequence of choices. This scheme introduces uncertainty into the laws. Most collisions can have two equiprobable results. As expected, all simulations of the evolution of partially ordered states continued to show increasing entropy and macroscopic behaviour conforming to thermodynamic predictions. Figure 8.4 shows the long term increase in entropy in the two-chamber model.



(a) (b)

Fig. 8.5. Two-chamber + pipe model (from Figure 8.3). Deterministic laws. (a) State velocity entropy G_v evolution for + and − 250 time steps from original "Then" state; G_v range $[2.8, 3.0]$. (b) Atom count difference (Right chamber − Left) for + and − 250 time steps from original "Then" state; 30,000 total atoms.

8.13.1 Simulation of the Past

First, note that the correctness of the simulator has been well checked. In at least 100 experiments starting from partially or highly ordered states, the simulation has been run (usually but not always) until a near-equilibrium view was reached with a near-maximum G_v. The simulation has then been put into reverse and run backwards to time zero. In every case, the initial condition was exactly restored. In many cases, the simulation was then driven further back in time, to equally remote negative times in the past, then forwards again to time zero. Again, the initial state was always exactly restored. (When running backwards with nondeterministic laws, the pseudo-random number generator was run in reverse to reproduce the "random" choices of collision table at each collision.)

For all the initial states simulated, without exception, running the simulation backwards to a negative time b resulted in a state $s(b)$ with higher entropy than the initial state $s(0)$, provided only that b was earlier than about -10. Most runs began to show an increase in entropy after the first two or three backward steps from the "present" $t = 0$. Figures 8.5(a,b) illustrate the effect in the two-chamber-and-channel model.

To gain more insight into this phenomenon, another series of experiments was done with a different protocol. First, a partially ordered state was set up, say, T (for "Then"). This was run forwards in time for long enough to produce an unmistakable increase in entropy, giving a new state, say, N (for "Now"). The time was chosen so that N was still clearly ordered, with an entropy well below the equilibrium maximum H_m, but definitely greater than $H(T)$. Let the number of time steps from T to N be k. Then the simulator was reset to have N as initial state at time $t = 0$. Of course, running it backwards to time $t = -k$ recovered T, provided that, when "non-determinism" was enabled, the pseudo-random generator was put into reverse with an initial seed value equal to the final value reached in the generation of N from T.

We now have the situation that the “initial”, $t = 0$, state of the system is $s(0) = N$, and we know its true ancestral state at the “past” time $t = -k$ to be $s(-k) = T$.

Consider the task of trying to recover the ancestral state at $t = -k$ given only knowledge of the “present” state N, the invariants, and the assurance that the system has not suffered external interference in the period $t = -k$ to $t = 0$.

The simulator was used to illustrate the effects of trying to do this by deducing the past state from the present state and the known microscopic laws. The simulation was run backwards from the final state N for k steps and the final state $\hat{T}$ at time $t = -k$ compared with T. The results can be briefly summarized. Running with deterministic laws and beginning with the exact state N, $\hat{T} = T$, i.e., the past is exactly recovered. This case exposes the loophole in Boltzmann’s theorem. If N is the deterministic evolution of a lower entropy T, of course the collisions which produced the increased entropy will, when reversed, reduce the entropy of N back to that of T. The collisions in the evolution of T to N leave the final velocities in N correlated, but the correlation is not so much between the velocities of the last pairs of atoms to collide as in the entire pattern of atom positions and velocities.

More realistically, we cannot hope to have exact knowledge of N, so N was replaced by a state N_p obtained from N by randomly perturbing the positions of some of the atoms by plus or minus one grid point. The population of N_p states so generated can be thought of as a very precise but not quite exact view N_v of N. Then, running with deterministic laws and beginning with several perturbed versions of N, $\hat{T}$ was never equal to T and rarely looked close. In particular, $\hat{T}$ always had greater entropy than N, and so of course greater than T, even when the number of perturbed atoms was very small. It seems that the loophole in Boltzmann’s theorem is very small, and the pattern in N required to slip back through the loophole to a lower entropy past state is very fragile.

Running with non-deterministic laws was done with an unreversed pseudo-random generator with various choices of seed, since we could have no direct knowledge of the random events which occurred leading to the present state N. In these runs, even starting with the exact state N, T was never recovered, and all instances of $\hat{T}$ had greater entropy than N. The non-determinism appears to erase any loophole finding correlation. Figures 8.6 and 8.7 show results on the two-chamber + pipe model. Figures 8.1(a-d) and Figures 8.2(a-d) show results from two simulations of collisions between two bodies of gas.

8.13.2 A Non-Adiabatic Experiment

All the argument and simulation to this point have assumed adiabatic (constant energy) systems. While this assumption follows from the assumption of a closed system, many of the conclusions seem to apply to systems which are subject to some external influences. A simulation was done with a 10,000 atom

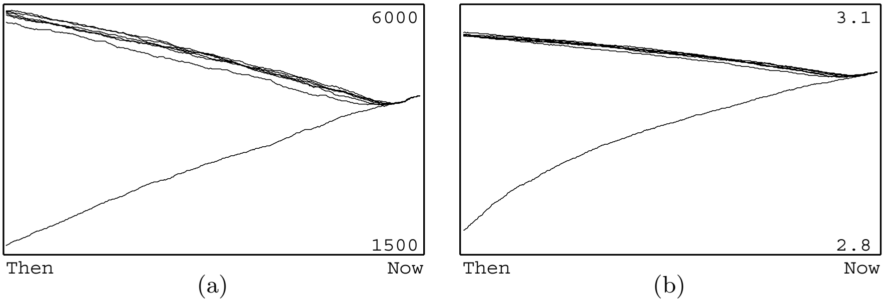


Fig. 8.6. Two-chamber + pipe model (from Figure 8.3). Deterministic laws. (a) Atom count difference devolved from "Now$_D$" to "Then"; traces show devolutions of "Now$_D$" with 0, 300, 100, 30, 10, 3 and 1 atoms perturbed by one grid point; only the unperturbed "Now$_D$" devolves to "Then". (b) State velocity entropy devolved from "Now$_D$" to "Then"; traces show devolutions of "Now$_D$" with 0, 300, 100, 30, 10, 3 and 1 atoms perturbed by one grid point; only the unperturbed "Now$_D$" devolves to "Then".

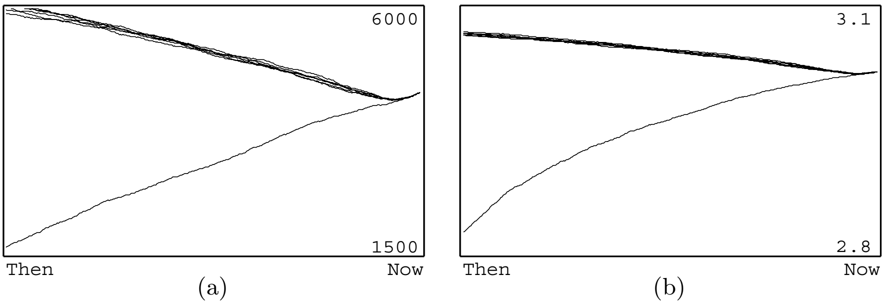


Fig. 8.7. Two-chamber + pipe model (from Figure 8.3). Non-deterministic laws. (a) Atom count difference devolved from "Now$_U$" to "Then"; traces show devolutions of "Now$_U$" with different random generator seeds; only the run with the seed equalling the final value reached in the evolution of "Now$_U$" devolves to "Then". (b) State velocity entropy devolved from "Now$_U$" to "Then"; traces show devolutions of "Now$_U$" with seven different random generator seeds; only the run with the seed equalling the final value reached in the evolution of "Now$_U$" devolves to "Then".

system in a 200-by-50 box, where the initial "Then" state was in equilibrium, with entropy close to the maximum possible, a uniform random distribution of atom positions, and a well-defined temperature of 2.0 (taking Boltzmann's constant as one.) Two special "heater" obstacles were placed in the box. In a collision phase, a "heater" adds +1 to the Y component of any atom at the heater's position if the time is even, or adds −1 if the time is odd. This action is deterministic and exactly retraceable. The effect of a heater is weakly to couple the gas to a heat source of infinite temperature, since the only velocity component distribution unchanged by the heater is the Uniform distribution.

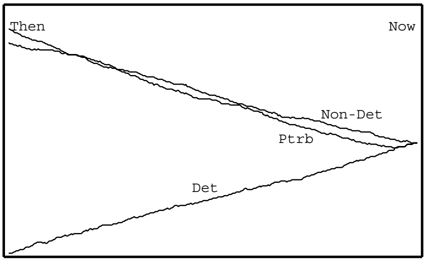


Fig. 8.8. Heater experiment temperatures; deterministic, perturbed deterministic and non-deterministic retraces; temperature range $[2.0, 2.25]$.

The simulation was run forward for 1000 steps to yield the "Now" state. As expected, the gas slowly heated up, so slowly that at all times the velocity distribution remained close to the equilibrium Maxwell-Boltzmann distribution, and G_v close to the maximum possible at the current total energy. Thus, at all times the gas had a well-defined temperature, which gradually increased from 2.0 in "Then" to 2.113 in "Now".

Reverse simulation was then done from "Now" back to "Then". Starting with the exact "Now" state, deterministic devolution exactly recovered the "Then" state. Starting with the "Now" state with one atom shifted by one grid point, deterministic devolution followed the true temperature trajectory for about 50 steps, but then showed an increase in temperature into the past, reaching 2.216 at the "Then" time. Non-deterministic devolution diverged from the true path almost immediately, reaching a temperature of 2.229 at "Then" (Figure 8.8). In both the latter cases, the entropy at "Then" was higher than at "Now", and of course higher than the true "Then" value. Although not shown in Figure 8.8, deterministic devolution of the true "Then" state to times earlier than "Then" also showed an increase in temperature into the past.

Clearly, this non-adiabatic situation presents the same problems for the deduction of past states as has been found in adiabatic systems, although now the increase in entropy occurring in forward evolution and backward devolution arises from an input of energy rather than the dissipation of order.