Foundations of Statistical Mechanics—
Two Approaches*

Stephen Leeds†

This paper is a discussion of David Albert's approach to the foundations of classical statistical menchanics. I point out a respect in which his account makes a stronger claim about the statistical mechanical probabilities than is usually made, and I suggest what might be motivation for this. I outline a less radical approach, which I attribute to Boltzmann, and I give some reasons for thinking that this approach is all we need, and also the most we are likely to get. The issue between the two accounts turns out to be one about the explanatory role probabilities play in statistical mechanics.

1. David Albert's *Time and Chance* (2000) is, by virtue of its clarity, concision, and sheer verve, a valuable contribution to the philosophical literature on statistical mechanics. It presents an approach to the foundations of statistical mechanics which is shared, at least tacitly, by many other people; but pushes it further, I think, than anyone has before, at least explicitly. I happen to be quite skeptical about whether the approach can be made to work, but I shall be presenting relatively little by way of argument against it. My goal is rather to point out an alternative approach I find more appealing, and to identify what is at issue between the two. The first section summarizes Albert's account, the second introduces my own, the third discusses what might be needed to show either account correct. The entire discussion, like the relevant part of *Time and Chance,* will be limited to classical statistical mechanics; I shall be pretending that the universe is a deterministic classical dynamical system.

Let us begin with the official formulation of Albert's reconstruction of classical statistical mechanics. This consists of a dynamical postulate, a hypothesis about the past, and a statistical postulate. The dynamical postulate is just F = ma. The Past Hypothesis is that "the world first came

into being in . . . whatever particular low-entropy highly condensed big-bang sort of macrocondition it is that the normal inferential procedures of cosmology will eventually present to us . . . ". The Statistical Hypothesis is that the "right probability distribution to use for making inferences about the past and the future is the one that's uniform, on the standard measure, over those regions of phase space which are compatible with whatever other information—either in the form of *laws* or in the form of *contingent empirical facts*—we happen to have" (Albert 2000, 96).

What does Albert mean when he speaks of the "right" distribution? He doesn't tell us explicitly, but surely the expectation is that use of the Statistical Hypothesis will turn out to be in some sense successful: indeed, Albert's reason for introducing the Big Bang hypothesis is precisely that without it, use of the microcanonical distribution will not be successful—it would presumably retrodict of a glass of ice-water that's been sitting a while in a warm room that it contained less ice an hour ago than it does now. Notice also that, as the last example illustrates, Albert expects that we will be able to use the 'right' distribution to make inferences—successful ones, presumably—about the past. It is worth pointing out how far Albert is willing to push such inferences: the microcanonical-distribution, suitably conditionalized, can be used not merely to pre- and retro- dict entropic rather than anti-entropic behavior; it also applies to commonplace situations in which the Second Law is not much of an issue. Thus, he thinks it validates inferences about whether there will be tomorrow or was yesterday a spatula in someone's bathtub:

> Suppose that I come upon an apartment about which I happen to have no direct empirical knowledge whatsoever other than the details of its architectural design and the fact that it contains a spatula. . . .
> If the distribution I use is one that's uniform over those regions of the phase space which are compatible *both* with everything I have yet been able to observe of its present situation *and* with its having initially started out with a *Big Bang, then* (and *only* then) there is going to be good reason to believe that (for example) spatulas typically *get to be where they are* in apartments only by means of the intentional behaviors of *human agents,* and that what human agents typically *intend* vis-à-vis spatulas is that they should be in *kitchen drawers.* (Albert 2000, 95)[1]

---

1. At one point Albert backs off from the very strong claim in the quotation—in fact, on the very same page as the passage quoted above (it's the passage about Mr. X, who bases his predictions on the total macrostate of the world). For more evidence that Albert's considered commitment is to the stronger claim, see the discussion on pp. 65–66, in conjunction with the claim, again on p. 96, that a virtue of the final version of the Statistical Hypothesis that it 'appears to get the story about spatulas *just right.*'

The idea that the microcanonical distribution, conditionalized on our knowledge, will be successful in guiding inferences about spatulas and the like is, I shall be soon be arguing, a claim which if true is not obviously so; certainly it goes well beyond what we find in the average statistical mechanics text. Nonetheless Albert has good reasons for advancing this idea. There's no single passage where these reasons are set out explicitly and in detail, but I think that a careful reading of the entire book will make it plausible that there is a line of argument which is at least compatible with Albert's discussion, and which leads naturally to the suggestion before us. The argument goes as follows:

Begin with the observation that in statistical mechanics we routinely cite probabilities in giving explanations: we explain the melting of this piece of ice on the grounds that it had a high probability of doing so. Notice that these probabilities are not the kind we find in quantum mechanics, where a complete specification of the state of a system is consistent with the existence of non-trivial probabilities governing its future behavior. Here, the complete specification of the state of a system (an isolated system, anyway) determines all its future behavior. The probabilities here are, then, not dynamical probabilities, probabilities of becoming; one might think of them as probabilities of *being:* an statistical mechanics probability is typically the probability that a system, given as being in a certain macrostate, is at the same time in a particular microstate which realizes that macrostate (or is in a particular family of microrealizations of that macrostate). So speaking of 'the probability of our ice melting' is a little misleading: what is extremely probable is that a piece of ice is in one of the microstates that guarantee deterministically that it will melt.

Now if one calculates probabilities by conditionalizing the microcanonical (henceforth mc) distribution over the macrostate of the system, it will turn out extremely probable as well that our glass of ice-water is in one of the microstates that can have come only from a glass of water with much less ice in it an hour ago. (Since this observation is a probabilistic updating of a famous objection of Loschmidt's, I will call it by Loschmidt's term: the Umkehreinwand, or Reversibility Objection.) Now this doesn't mean that we can't also explain why in fact there was much more ice in the glass an hour ago: we can, by describing the state of the glass two hours ago. We are not in the business of explaining past events on the basis of their futures, and there is no reason we need to be. Nonetheless, the Umkehreinwand raises a problem for our explanation. That is because when we cited the high microcanonical probability that the ice was in a destined-to-melt state as the explanation of its melting, we were committing ourselves to the claim that events with an high probability by that measure were reasonable to expect, or the claim that such events happened most of the time, or both of these. Nowadays most of us have

to some degree loosened the connection between explanations and expectations: we are willing to explain the occurrence of a low-probability event by quoting the probabilities. But there is this much connection between the probabilities we use in explanations and what we expect to happen: we demand that the probability measure we use in explanations comes close enough to the actual frequencies that it would be reasonable to use it as a guide in inference. The Umkehreinwand shows that the microcanonical distribution fails to meet this demand: however close it comes to the actual frequencies with which a piece of ice is in a destined-to-melt state, it is stunningly far from the actual frequency with which ice cubes are in has-recently-been-water states.

Why not respond to the difficulty by weakening the connection between explanation and expectation even further than we already have? Why not say that, given that the microcanonical distribution does come close to the frequencies that are relevant to the events we want to explain, there is no reason to demand that it match the actual frequencies of events that we don't use it to explain? Albert doesn't really address this question directly; my guess is that he would see this as giving up on the idea that we are *explaining* the melting of ice by citing microcanonical probabilities. The use of probabilities in explanation requires that these probabilities have some sort of physical reality, or at least physical significance; a probability distribution that doesn't even remotely approach the actual frequencies seems a poor candidate for such reality, or significance, however we understand these rather vague phrases.[2] Notice however that this argument

2. A recent discussion that attempts to shed some light on what it might be for the statistical mechanics probabilities to be physically real—and which might reflect some of what Albert intends here—is Loewer 1996. Loewer there argues that David Lewis' well-known treatment of chances can be carried over to the present non-dynamical situation. For Lewis, what makes e.g., the QM chances physically real, and usable in explanations, is that they are definable in a natural way, and conform closely to the actual frequencies. Loewer suggests that there is no real difference in the case of the statistical mechanics probabilities.

I do not find this entirely convincing. Surely a crucial part of our sense that the QM probabilities are 'physically real' is tied up with our sense that we need them in order to give the dynamics—more precisely, to give the probabilistic 'dynamics' whose role may reasonably be said to correspond to, or inherit, the role of dynamics in the deterministic case; of course the mc distribution plays no such role in the deterministic case. (It might be replied that in the Bohm interpretation the QM probabilities aren't needed to give the dynamics, but it is not so clear that that advocates of that interpretation think of them as 'physically real' either.) My own view is that there is no room for a notion of 'physically real' probabilities except where they can be shown to have a connection to the dynamics. So—to move a little off our topic of classical statistical mechanics—one way to show the mc distribution physically real would be to derive it from the stochastic QM probabilities, or the GRW probabilities that may underlie these. This would be a project rather far from what Albert has in mind: it would neither

makes no distinctions between prediction and retrodiction: the 'right' probability distribution should not only predict that ice will melt; it should also retrodict that a piece of ice in a warm room probably arose from the melting of a larger piece of ice. And in expecting that the 'right' distribution will deal with spatulas and the like, Albert is merely following out what is implicit in the account already. Any probability distribution will have something to say not only about the probabilities of macrostates conditional on earlier or later macrostates, but of macrostates conditional on less precisely described macrostates—the probability of there being a spatula in the bathtub given that there's one in the apartment. And if these are wrong by a mere 2 or 3 orders of magnitude, in contrast to the truly amazing numbers in the icewater example, this should be a serious concern to us as well. As with the Umkehreinwand, to say that the mc distribution is the 'right' one to use, although it is very wrong about spatulas and bathtubs, raises the suspicion that by 'right' we man only 'successful if used in certain ways,' and then the less it sounds as if we are *explaining* why the ice melted by quoting its high probability of doing so, as opposed to merely using an algorithm for predicting that it will melt.

Albert gives a plausibility argument that if there is a probability distribution that does well with retrodiction, spatulas, and the like, it might well be the microcanonical distribution conditionalized on the initial Big Bang state of the universe (henceforth I shall call this the mcc distribution)[3]. But do we have any reason to think that the mcc distribution, or indeed any natural distribution, *will* do well with them? In the passage I

---

validate retrodictive inference nor would it need to require a special initial state of the universe. I do not know how likely such a project is to succeed: in Quantum Statistics one uses ensembles just as in classical statistical mechanics—no one thinks to derive these from the measurement probabilities—but perhaps one can get more from the GRW probabilities.

3. Albert speaks of the 'uniform' distribution, rather than 'microcanonical'. 'Uniform' perhaps puts off for a while confronting the implications of the current view that the universe has infinite mass and presumably infinite kinetic energy. I think that sooner or later, both his approach and mine will require a finite universe, else it may well turn out that the kinds of Big Bang states we want to say are probable, and those we want to say are improbable, will both have infinite measure. We might do this by identifying the 'universe' with e.g., the past light-cone of the finite part that interests us.

I use the term 'mcc distribution' ambiguously. In its primary sense it is the mc distribution conditionalized over a certain macrostate—the Big Bang macrostate. This distribution, by updating, leads to a probability distribution at each time t for the possible microstates of the universe at time t, and when the context is clear, I will often refer to each of these as the mcc distribution as well.

Note that when I speak of macrostates, or macroscopic descriptions, I mean a description on some level of detail that allow us to say that entropy increases as ice melts in water: that is, such that a glass of ice-water is not in the same macrostate as the same glass with all the ice melted, even if the energy and mass of the two is the same.

quoted above, Albert raises the problem that the mc distribution over the phase space available to my apartment will assign a higher probability to the spatula's being in the bathtub than in the kitchen drawer. The mcc distribution is supposed to help with this: the idea is that, once we've conditionalized on the Big Bang, all the stories in which the spatula materializes in the bathtub out of thin air get eliminated, and ordinary stories about human intentions come to the fore: since people typically intend to keep spatulas in the kitchen drawer, this gets the higher probability. But this does not really solve the problem. The mc distribution over the gross thermodynamic features of Crystal Lake on a day in early October will predict that the gallon of water I just fished out will be in equilibrium— cool, but free of ice: the mc probability of there being any ice at all in the pail is extremely small. In fact, I might have noticed that there is already a fair amount of ice in the lake; I would be foolish to bet against ice in the pail at anything like the mc odds. Will conditionalizing on my knowledge give a more accurate answer? Suppose I know the contents of every part of the lake except what's in my pail; this doesn't make any difference at all to the calculation: the enormous majority of states consistent with all this knowledge place no ice in my pail.

Our problem with Crystal Lake is an instance of a familiar problem with uniform distributions: you cannot make a reasonable guess as to the contents of an urn filled with (labeled) red and black balls by beginning with a uniform distribution over all possible assignments of red and black, and then conditionalizing over samples taken from the urn. A uniform distribution over the various possible *proportions* of red and black balls will do much better, but the mc distribution over the positions and momenta of the molecules in Crystal Lake is of course uniform over the assignments of position and momentum to each molecule, not over the different position-momentum profiles.[4] Of course we haven't yet invoked the fact that we are conditionalizing over the Big Bang. What we now see is that if doing so leads to reasonable probabilities here, these cannot be accessed simply by taking the mc distribution, conditionalized on current information, and then using the Big Bang to eliminate all histories with counterentropic behavior. Rather, the implications of the Big Bang must be more extensive and subtler than that: somehow conditionalizing over it must turn the mc distribution into something closer to the uniform distribution over profiles without undermining the crucial ways in which statistical mechanics requires a distribution that is uniform not over pro-

4. I learned the nice word 'profile' from Albert's book: the profile of a distribution is the list of how many particles (and of what type) are in each 6-dimensional momentum-position 'box'. The single-particle density functions we deal with later are smoothed-out profiles.

files but over microstates (e.g., in assigning the largest measure to the equilibrium distribution). Perhaps this really is the case, but we have nothing whatever to go on by way of calculation or observation that would lead us to believe it. Our only reason to believe that conditionalizing over the Big Bang will lead to reasonable probabilities is that this demand seems an essential part of what might seem itself an inevitable way of looking at statistical mechanics. What we need now is to see whether there is another way of looking at the explanatory structure of statistical mechanics that doesn't demand so much.

2. Here is Larry Sklar discussing ergodic theory—the point however is a general one:

> Ergodic theory considers the question: Why does the natural probability distribution [the microcanonical measure] work? The answer it gives is the proven equality of phase-averages to infinite time averages. But there is a much simpler answer. And it is correct. And it is the *full* answer. And it is totally independent of any ergodic results. It goes like this: How a gas behaves over time depends on (1) its microscopic constitution; (2) the laws governing the interaction of its microconstituents; (3) the constraints placed upon it; (4) *the initial conditions characterizing the microstate of the gas at a given time.* (Sklar 1973, 210)

Sklar is surely right; there is a sense of 'explanation' in which we already have, at least in principle, an explanation of all the facts of thermodynamics. When we feel—as we surely do—that we don't yet have an explanation for the facts of thermodynamics, we must be looking for some above and beyond our ability in principle to derive them from initial conditions and dynamics. But what is it we are looking for? I want to suggest that corresponding to each part of Sklar's explanation, the dynamics and the initial conditions, there is something more we are looking for than the facts that Sklar cites.

In the case of the dynamics, it seems to me there is not much controversy about what we want—although not much agreement about how to find it. As Sklar reminds us, there is a sense in which, given the initial conditions of say our glass of ice in water, we understand the processes by which it melts. But it is also true that merely being able (in principle) to trace the microscopic history of one particular case of a system approaching equilibrium leaves us without a perspicuous grasp of what is going on. We want a story that applies to all cases of approach to equilibrium, and this means that we want to be able to abstract from the mass of detail, and describe what it is about the initial state of the ice-water in this case that it shares with other instances in which ice-water went to equilibrium—or indeed, went to equilibrium at this particular rate. There

are different strategies or research programs that propose to tell the story in different ways, and they differ greatly in the extent to which they promise to give us a sense of knowing what is going on when our ice melts. Among the least satisfactory are the accounts that appeal to the ergodicity of an isolated system: here, one claims that almost every (mc) initial state of the ice water is destined to spend most of its future in an equilibrium state. In such an account, the only characterization offered of the initial states that do go to equilibrium, as opposed to those that do not, is that the former is one of the many (mc) states that are destined to do so; the account leaves us in the dark about what physical feature distinguishes one class of states from the other. I think this is a defect; in any case, this approach has other serious problems, most notably its failure to say anything about the speed with which systems go to equilibrium. For this reason my own bet would be on the kind of story that Boltzmann offered early on, and which is still the one we tell if someone asks us to tell them quickly why ice melts: that an appropriate randomness in the initial distribution of the particles favors the collisions that drive the system in the direction of equilibrium. My guess is that only an account that talks about collisions in some detail can account for the known relations between frequency of collisions and rapidity of approach to equilibrium; and I think it is clear that such an account would do better than the ergodic account in telling us what physical feature distinguishes the initial states that evolve to equilibrium. It may be of course that this approach did not pan out: it might be that we could find no physically natural feature shared by the states that evolve to equilibrium at the appropriate rate. This would be disappointing, but I don't think it would be an explanatory disaster. It is a disaster to have no explanation whatever for a certain phenomenon, but in this case, as Sklar says, we do have one: what we are looking for is a certain level of perspicuity, and there is no guarantee that Nature will provide us with *that.*

Turning now to initial conditions, there is much less agreement about what sort of explanation we are looking for. Here the explanandum is the fact that the semi-isolated systems so often are in states which have that feature, whatever it is, that leads to an approach to equilibrium at the appropriate rate for that kind of system (as we just saw, this feature may amount to nothing more than being one of those states which will evolve to equilibrium at the appropriate rate). I will suppose with Albert that the first step in explaining this is to say something about the initial state of the universe, and I will also suppose that describing this initial state as a Big Bang low entropy state is not enough by itself to account for our explanandum: there is no reason to believe that *every* Big Bang low entropy state will lead to a world that obeys the Second Law. We need to find some other feature F of the early universe which (perhaps together

with the fact that that universe was a Big Bang low entropy universe) guarantees that virtually all semi-isolated systems are destined to go to equilibrium. As in a similar context above, we can hope that F is a physically natural feature, one we can describe perspicuously, but we cannot be sure it will be.

Now what should we ask for by way of explaining why the early universe had F? If 'early' means the very moment of the Big Bang, then of course it is out of the question to explain the presence of F in terms of earlier conditions that led to it. If 'early' means some time shortly after the Big Bang, then we might be able to do better—there might be an illuminating account of how some feature distinct from F, present in the still earlier universe, gave rise to the presence of F, analogous to the account in inflationary theories of how one form of low entropy at the time of the Big Bang gives rise to a later low-entropy state in which the low entropy takes a different form. But if F is some not particularly natural feature of the distributions of particle positions and momenta, we can easily imagine that no such story would be forthcoming: and in any case, we will only be pushing the explanatory story back. Sooner or later, it seems, we must come to the point where we say that the universe started off in a particular kind of state, and surely there explanation comes to an end. What can it matter, then, what we say about initial conditions? Why not say that they were such as to lead to second-law behavior and leave it at that?

The answer has to do with the plausibility of our account. If we ever knew with utter certainty that any of our theories was the whole truth about its subject matter, then whenever that theory presented us with what it labels a brute fact—an unexplained explainer—then there would be nothing left for us to look for, except perhaps to formulate the fact as perspicuously as we could. But we never know with certainty that any of our theories is true, still less the whole truth. Every one of our theories is in competition, if not with known alternatives, then with alternatives not yet formulated, and a big part of the competition consists in comparing what the theories treat unexplained.

Among the things any theory will leave unexplained are those ultimate unexplained explainers, its fundamental laws. And here there is general agreement that it matters greatly to the plausibility of a theory just what its laws are. There is also agreement that among the considerations we bring to bear here are some which might be called 'subjective' in the sense that one cannot infer them straightforwardly from a statement of our goals in doing science. Even if it is in some sense built into the nature of the scientific enterprise that we prefer simple laws to less simple ones, it is not built in that we measure simplicity in the particular ways we do: our particular standards are our own—one can easily imagine them as having

been different, and there is apparently no way to justify them except to say that one needs *some* standards, and these are ours. The same point holds even more strongly for all the other criteria we use which seem to have little to do with simplicity: for example, our preference for local, or for geometrical, laws in physics.

There are other things a theory may leave unexplained. Theories often postulate initial conditions, and these are unexplained explainers too; here too it matters to the plausibility of a theory, as compared with its competitors, what kinds of initial conditions it postulates. Given a choice between a theory that counts the initial conditions of the actual world as extremely unusual among the possible initial conditions allowed by the laws of the theory, and one in which these are more or less typical, we find the latter to that extent more plausible; the first theory will leave us with the feeling that there is a question—why is the actual world like *this?*— which we would like to see answered; the second theory allows us to dismiss the same question by saying 'Well, why *shouldn't* it be?' As in the case of laws, our evaluations rest largely on 'subjective' criteria—that is, criteria which we find natural to use, but which are not the only criteria we might have used. These surface when we give specific content to the word 'typical.' An initial condition is typical if *many* initial conditions are *like* it, and it is up to us to choose the respects of likeness we have in mind and what we will count as 'many'—the latter choice being particularly far from a routine matter when the family of possible initial conditions forms a continuum.

In the case of statistical mechanics, formulating the most plausible version of the theory we can has something of the nature of a five-finger exercise: classical physics is no longer in competition with any of the theories we take seriously. To see problems about the plausibility of initial conditions in a context where they matter to our choice among theories, one might look instead at the Bohm theory, as compared with its competitors.[5] But even if classical statistical mechanics is no longer in competition, we can still ask (hoping perhaps that the answer bears on current problems) how best to formulate it—in particular, whether the theory has a formulation in which our world is typical, according to some natural standard of typicality. There is a limit on how far we can go here. We have very good evidence that our world was once in a state of lower en-

5. In fact, one important strategy for explaining why it is that the quantum statistics hold, in terms of the Bohm theory—that of Dürr, Goldstein, and Zanghi (1992)—is a very close relative to the approach taken here. Dürr et al. show that, for any initial wave-function ψ for the entire universe, most actual distributions of particles will evolve in such a way that the quantum statistics are obeyed. 'Most' here is according to the standard $|\psi|^2$ measure, which they single out as the most appropriate one in this context solely on the grounds of its naturalness.

tropy, perhaps that it originated in a Big Bang; if classical statistical mechanics is to shed any light on the theories we take seriously, it seems likely it will need to include these facts about the world; but in any reasonable sense of 'most,' most of the worlds allowed by classical physics (or even contemporary physics) as physically possible do not arise from low entropy Big Bangs. We can still ask, however, that our world work out to be typical among worlds that arose from a low entropy Big Bang—that many such worlds be like our world. Given the features of the world we are trying to account for, it seems reasonable to ask that the respect in which many initial conditions turn out to be *like* those in our world leads to histories in which classical statistical mechanics holds—in other words, we want there to be many initial conditions with the feature I've been calling F. As for *many,* my suggestion is that we should read this as: 'many, according to some measure that we find reasonably natural'. And here the mc distribution seems as natural as one could hope for: it is in one clear sense a uniform distribution (and there is good reason to hope that whatever features are very common by the mc distribution will also be common according to the other distributions that we likewise see as 'uniform'—indifferent, unprejudiced), and of course it is a distribution that is natural in the sense of being invariant under canonical transformations.

The microcanonical measure thus has several quite distinct roles to play in the explanatory structure of statistical mechanics. We will want to show that the class of states with the crucial feature F is large according to the microcanonical measure conditionalized on the initial macrostate. But the role the microcanonical measure is playing here is quite different from the role it plays in characterizing equilibrium, or in explaining the approach to equilibrium. From the present point of view is a mistake to conflate the three: for example, to note, as so many discussions do, that statistical mechanics makes heavy use of the microcanonical distribution, call that a 'probability', and then go on to ask what kind of probability this might be—subjectivist, frequency, etc. The use of the microcanonical distribution in explaining the initial state of the world, to the extent it fits into any of these headings, is subjectivist: we are showing that our world is not an atypical one when we measure typicality according to a principle of indifference—of what *we* find indifferent. The use of the mc distribution in characterizing equilibrium is perfectly 'objective', but it is somewhat misleading to call it a probability: it is a mathematically definable measure on phase space, and there simply are no competing theories about what it is, any more than there are frequentist or propensity accounts of Lebesgue measure. Finally, when we say that initial states of thermodynamic systems typically go to equilibrium, or follow ergodic paths, the mc distribution doesn't enter at all: we are talking about a finite frequency. To

call this frequency a probability seems to be overstating things, since we haven't so much as defined an event space, any more than we do when we say that most people in the US live near cities.

In this account, the reason this piece of ice melted is that it was in a state of the kind that deterministically leads to melting—a kind which we hope to characterize in a more transparent way. Given that it was a piece of ice, why was it in such a state? There are answers at different levels (one might tell the microscopic history of the ice-cube); at the level of explanation we have been considering, the answer is: because the universe was in a certain kind of initial state—again one which we would hope to characterize in a more transparent way. Although probabilities have a role to play in this story, namely in allowing a certain claim about the initial state of the universe to be a plausible one, they have no direct role to play in explaining why ice melts. For this reason, no difficulty arises from the fact that the mc distribution, conditionalized on what we know of the ice, retrodicts that it arose from a more melted state: we are not in the embarrassing position of using the mc distribution to give explanations in one time direction, but refusing to be guided by its assignments in the other direction.[6]

The fact that mc-most states of our ice water encode the history of behavior that violates the second law is supposed to create two kinds of trouble for statistical mechanics: it is supposed to undermine the kinds of explanations that statistical mechanics can offer, and it is (though less frequently) held to undermine our confidence in the past having been as we believe it to be. With respect to explanation, there is one respect in which the present account abandons some terrain in the face of the

6. Notice another way in which the explanatory structure of statistical mechanics involves a 'subjective' element: in saying the state of the early universe belonged to a particular family which is large on the mcc distribution—the set of states with feature F—we are inevitably ruling out that it, or the state of the universe at any later time, belongs to some other families which are also mc-large, or which are large on the mc distribution when we conditionalize it on some other facts about the universe. If the initial state has F, then the final state of the world, if there is one, does not belong to the set of states whose time-reversed evolution obeys the second law; and this set of states is presumably large when we conditionalize the mc distribution by the macroscopic description of the world's final state. Likewise for most (mc) states of the world consistent with its current macroscopic description. So in saying that the initial state of the universe had F, we are ruling out many competing and in a sense symmetric claims we might have made instead. A choice was inevitable here: our universe can't of course belong to *every* mc-large class. Do we choose to explain statistical mechanics in terms of initial conditions because we prefer explanations in terms of initial conditions, or because nothing else has any chance of working? Both, I think: our preference for explanations that account for what happens at a later time in terms of what happened earlier is largely a consequence of the fact that in this world those are the only successful macroscopic explanations.

Umkehreinwand: we do not say, as so many writers on statistical mechanics have wanted to, that the high mc probability of melting explains the melting. On the other hand, the high mc probability of melting will still play a crucial role in the present account, as it should in any account that is true to the spirit of *statistical* mechanics, for (I will say more about this in the next section) it will figure in the reason that mcc-most initial states of the universe usually produce ice which goes on to melt. I think that, at least as far as explanation goes, the Umkehreinwand holds no further terrors for us. Sometimes one hears the suggestion that the difficulty raised by reversibility for our notion of explanation lies in the fact that we need to explain why the ice before us came from an earlier larger cube of ice, rather than from a glass of warm water, and that the reversibility objection blocks the most natural way to do so—namely by showing it is currently in a state which has a high probability of coming from water. I agree that this sort of explanation is blocked, but I think all that is needed to give a complete explanation here at the level we are seeking is to explain why it is that an isolated glass of warm water virtually never freezes (reason: the initial state of the universe had feature F), and then to point out the initial conditions that led to this block of ice—presumably a somewhat larger block of ice.

As for the epistemological problems that the Umkehreinwand is supposed to raise, it is true that if we want to assign a probability to the piece of ice before me having arisen by spontaneous freezing, the present account does not endorse doing so via conditionalization on the mcc distribution. That is not to say it allows no inference at all. The reasoning is of the same humble variety by which, given no special information about the next student who shows up for office hours, I believe rather strongly that her Social Security number will not be 327–89–6310: If this is a Bayesian inference at all, it is the ordinary 'non-objective' kind where one picks as uniform a prior as seems appropriate to the situation at hand. More plausibly, it is a frequency-based direct inference: believing that water spontaneously freezes at most once or twice in the whole history of the universe, and having no contrary statistics for any subclass containing this piece of ice, I let those statistics guide me in assigning a probability to the case before me. As for the weightier issues of skepticism about the entire past, the Umkehreinwand will have force here only if one accepts three doubtful propositions: a) that we have, or should retreat to, an evidence base—a set of known propositions—which we should use as a basis for assigning probabilities to every other proposition by using some 'objectively based' probability distribution, b) that this evidence base contains no propositions about the past, c) that the appropriate distribution to use is the mc distribution. Albert ends up rejecting b) and c). As I see it, there is no reason to accept a), so b) and c) become moot. One reason to reject a) is

this: we already have all sorts of opinions which don't count as knowledge—they are not held with anything near to certainty—but which nonetheless embody much of what we have learned about the world; indeed, much of what our species has learned, since undoubtedly some of these opinions have been selected for through the history of the species. To retreat to our 'evidence base' is to voluntarily surrender information with little hope of getting it back, and why should we want to do that?

3. Our two accounts—let us call them accounts A and B[7]—however different in their motivations, are in one respect pretty much in the same boat: they both rest on empirical and mathematical claims that no one has yet shown to be true. They are, however, in slightly different parts of the boat, in ways that are worth spelling out.

The biggest point of difference concerns how much is required of the initial state of the universe. Both accounts require it to be quite a low-entropy state, since both accounts try to support the idea that entropy has always been on the increase. Will this be enough? Or will we need to place stronger requirements on the macroscopic description of the initial state before it turns out that the mc distribution, conditionalized over that description, does what we want it to do? Because account B requires less than A, it has a better chance of doing without such extra requirements. And, in fact, the strategy that has traditionally suggested itself as a natural way to go about B is one which, if it is successful at all, seems likely to do without such requirements. The idea is to begin by justifying a kinetic equation—an analogue of the Boltzmann equation—for each kind of thermodynamic system. What one hopes to show is that for any macrostate of a semi-isolated system, the enormous mc-majority of microscopic realizations of that macrostate are destined to proceed towards equilibrium, at least for the lengths of time we commonly observe, at a rate which is predicted by some kinetic equation appropriate for that kind of system, and that macrostate. Something like this is, it seems to me, what most people expect to turn out to be true; progress on establishing it has been so far disappointingly slow, but by no means negligible. In particular, in the case of dilute monatomic gasses, there is Lanford's proof (summarized in Lanford 1983) that, in the Grad limit (as atomic diameter decreases and number of atomic increases, with constant nd^2) almost every

7. A for Albert; B because I believe the view is really Boltzmann's. The best-known statement of B is probably section 18 of Ehrenfest and Ehrenfest 1959. Contemporary writers in this tradition include Lebowitz, and—to some extent—Lanford and Grad. Grad's relations to this tradition are particularly complex and interesting, but I don't have space to discuss him here. Among philosophers, Paul Horwich and Huw Price take closely related positions—in each case with certain differences.

(mc) initial microstate compatible with any given single-particle number-density distribution will obey the Boltzmann equation, at least for a (very brief) while. There are unquestionably problems with the proof—in particular in connecting what it shows to be true in limiting the behavior of real systems—but I think the consensus is that the overall project is worth pursuing, and it does in fact continue to be pursued.[8]

Suppose that Lanford's proof can be generalized to hold for other thermodynamic systems, and that the Grad limit (or whatever limit is appropriate—different limiting theorems use different ones) turns out to be approached quickly enough so that actual systems count as near enough; we are still far from showing either that real systems obey a kinetic equation for a reasonable length of time, or that mc-most initial states of the world compatible with any given initial macrostate will lead to a world in which kinetic equations are generally obeyed. The two problems are closely related, and would both be solved if we could show that the past history of a system is in the appropriate sense irrelevant to its future behavior. The idea was already clearly stated by the Ehrenfests in 1912 (Ehrenfest and Ehrenfest 1959); here is a contemporary statement by Joel Lebowitz:

> " . . . for systems with realistic interactions the domain Γ_{ab} [of all states which at t_2 represent a system with macrostate M_b which has evolved from a system with macrostate M_a at t_1] will be so convoluted as to *appear* uniformly smeared out in Γ_{M_b}. [the set of all states which at t_2 represent a system with macrostate M_b, irrespective of previous history] It is therefore reasonable that the future behavior of the system, so far as macrostates go, will be unaffected by their past history. It would of course be nice to be able to prove this in all cases; . . . [although we can only in 'very simple situations' (the reference is to Lanford, and to some results on gas particles moving among an array of scatterers)] this should however be enough to convince a 'reasonable' person." (Lebowitz 1999, 349)

8. I don't think Sklar's degree of scepticism about the Lanford result in Sklar (1993) is quite justified. Sklar claims that taking the Grad limit here is not just a useful approximation: that the irreversibility the theorem claims to show (via claiming to show that most systems obey the Boltzmann equation) in some sense exists only in the limit. I suggest the theorem does indeed tell us something about systems of finite size, namely that, given the initial number-density function f_1 of a gas, normalized to apply to systems with arbitrary numbers of particles, and given a degree of approximation e, then mc-most gasses whose number-density functions are very close to f_1, and which have sufficiently many and sufficiently small particles, will evolve to have number-density functions within e of the Boltzmann time-development of f_1. To be sure, we don't know how small the particles have to be, but it is noteworthy that computer simulations of hard-sphere collisions by G. A. Bird and others have produced the required behavior with numbers of molecules in the mere thousands.

The same property—a close relative of the property called 'mixing', though limited to a particular class of states[9]—would straightforwardly allow us to claim that mc-most initial states lead to the kind of 2nd law behavior we observe: if we use the initial mc distribution to assign probabilities to the various macroscopic trajectories of the world through time, then for any point p along any trajectory, the conditional probability of continued Second Law behavior throughout the next short interval of time (given the macroscopic history of the trajectory up to p) can be calculated by calculating the probability that each component subsystem will exhibit Second Law behavior; but, by the mixing property, these probabilities can be calculated by conditionalizing the mc distribution over the macroscopic description of the component subsystem—and this predicts Second Law behavior. Of course, as Lebowitz mentions, actually proving the appropriate mixing property seems a very distant goal. I might mention that, as one searches the literature for work bearing on this goal, it is difficult to escape the impression that a certain amount of confusion has prevented people from focusing on it clearly. In particular, a damaging role has been played by the conviction that once one has shown that mc-most systems will obey a kinetic equation, one has thereby shown that such behavior is 'probable,' and that nothing else is needed. Likewise, the Gibbsian idea that statistical mechanics is entirely about the behavior of ensembles has led people to overlook even the possibility that one might want to prove a kinetic equation involving the single-particle number-density distributions (as opposed to the quite distinct marginal probability f_1 of the ensemble).

In Albert's approach, the situation seems to me somewhat less hopeful (you might think I should say: even more hopeless). Here the mcc distribution needs to predict not merely the approach to equilibrium but much else too. For this reason, it is harder to begin with something we know, or more or less firmly believe, about ordinary thermodynamic systems and try to build on that. We can with some plausibility think of the universe at the beginning as just one more thermodynamic system, and assume that what we know or believe about ordinary thermodynamic systems applies to it as well. So I think it would be reasonable to guess, or hope, that just

---

9. The limitation is essential: we do not want to claim that the states which encode an anti-2nd-law past history are as frequent in Γ_{ab} as they are in M_b. There is no *obvious* incompatibility with the underlying time-symmetric underlying dynamics; nonetheless, it hardly needs saying that the claim is one that requires proof. It is here that, even given the hoped-for Lanford results, there remain substantial difficulties. On the other hand, the suggestion of one referee that there is a difficulty about parallel time-development of branch systems seems to me unfounded: if most states of most systems follow the Boltzmann equation (as opposed to its time reversal) then they increase entropy in the forward time direction.]

as we predict the future of this glass of ice-water by conditionalizing the mc distribution over its macroscopic parameters, so we can use the mcc distribution to predict the future of the universe as a whole. The sort of retrodiction Albert is considering takes a different form: here, what we want to do for the whole universe is, by conditionalizing the microcanonical distribution on the initial macrostate I, to be able to retrodict, given a later macrostate L, what an intermediate macrostate M had been.[10] The problem with getting any support for this from the behavior of small systems is that the only analogous cases we observe for them are in a sense trivial. For ordinary systems, there is, given I, only one L we ever see: the one (call it L*) which has overwhelming probability according to the mc distribution conditionalized over I. And likewise only one M we ever see, again the M* predicted with overwhelming probability. The predicted probability of M* on L* (the ratio of the two numbers p(M* & L*) and p(L*), both near 1), is then automatically very close to 1, and of course we do see both M* and L*. The probability of M*, given some L′ other than L*, is non-trivial and interesting, but it makes no sense to ask how well this fits our experience, since we never experience L′. Now Albert's retrodictions—where the Big Bang plays the role of I and M and L are features of the whole state of the world like that of containing a spatula in a bathtub at a certain place—mostly involve probabilities which are surely less than 1. Even the main case that interests him—that of retrodicting of a bit of ice in water that it was previously more frozen—involves an L (there being a piece of ice here) and competing M's whose probabilities on the initial Big Bang microcanonical probability, or on this conditioned on what we know of the history of the world, are not close to 1. Our success in the trivial cases where the predicted probabilities are nearly 1 cannot lend much support to Albert's quite non-trivial claim about retrodiction, as against competing, more conservative (and equally simple) hypotheses—e.g. the hypothesis that the uniform distribution over macrostates correctly predicts the approach to equilibrium (and which says nothing one way or the other about retrodiction.)

If we set our sights a little lower, and ask only to find a distribution that lets us predict the macroscopic state of the world, given appropriate macroscopic information, there is something we can get out of Large Number theorems. Call a world mcc-well-calibrated if in it we can successfully use the mcc distribution, conditionalized over the macroscopic history up to any t, to predict the macroscopic state at $t + \Delta t$: specifically

---

10. More precisely, we are trying to retrodict certain *features* of the intermediate macrostate from certain features of the later macrostate: the kind of retrodiction we are interested in for the universe involves e.g. retrodicting the history of the Earth from its current macrostate, and these are really features of the whole universe's macrostate.

if the proportion of occasions on which those outcomes at $t + \Delta t$ which are predicted with probability r come to pass is, in the long run, 100r% (for all r). It then turns out we can show that mcc-most initial states of the world lead to mcc-well-calibrated worlds.

Perhaps one might improve this sort of result. Conditionalizing over all of past history is crucial to the argument, but one might hope that a version of a Markov principle applied: that conditionalizing over the present macrostate would give the same results. Likewise, one would wish to see a version of the argument that applied, not to the long run, but to some humanly significant time frame. Suppose these problems could be gotten around; what would we have? The result we're discussing is quite general—most D worlds are D-well-calibrated, for any distribution D-; for this reason I cannot see getting out of it anything stronger than the sort of conclusion I argued for earlier: that, the mcc distribution being a pretty natural one, we should not regard it as needing explanation if the actual macroscopic trajectory of the world were well-calibrated with respect to it. This however is no argument that the world *is* mcc-well-calibrated. For that, one needs some access to what the mcc probabilities in fact are, and how things have turned out with respect to them. This leads us back, it seems to me, to the only domain in which we have any real evidence: ordinary statistical mechanics. The uniform distribution over the current macroscopic state of a system gives the right probabilities for the approach to equilibrium—so we believe and hope one day to prove; if this agrees with the mcc distribution over the state of the system, or over everything we know about the world, well and good; but I see here no argument that that particular distribution is well-calibrated for inferences about the past, about spatulas, or anything else outside the usual domain of statistical mechanics.[11]

11. A referee has asked me to say something about what happens in accounts like mine (or Albert's) if we make things a little bit more realistic and suppose there is a stochastic element in the world à la GRW. To keep things as classical as possible, suppose that occasionally particles jump, according to some probabilistic law. Then there are really two questions to ask: does the presence of jumps make all postulates about the initial state unnecessary, since the jumps will knock our isolated systems into 'good' states anyway? and, given the presence of jumps, can it help to put any requirement on the initial state, since the jumps will erase all memory of it anyway?

At first glance, it might seem that one cannot answer 'No' to the first question and 'Yes' to the second: a Yes answer to the second requires that the feature F that drives future 'good' behavior holds in a large measure of states—large enough so a random jump will take us from F-states to F-states; and this may seem to entail that random jumps from not-F-states will take us to F-states too, so that the answer to the first question is also Yes. But the reasoning here is questionable. If a gas is not in a good state—say it is destined in the next few seconds to display anti-Boltzmann behavior—it is not plausible that the jump of a particle on Mars will knock it into one; even a

REFERENCES

Albert, David Z. (2000), *Time and Chance.* Cambridge, MA: Harvard University Press.

Dürr, Detlef, Shelly Goldstein, and Nino Zanghi (1992), "Quantum Equilibrium and the Origin of Absolute Uncertainty", *Journal of Statistical Physics* 67: 843–907.

Ehrenfest, Paul, and Tatiana Ehrenfest (1959), *The Conceptual Foundations of the Statistical Approach in Mechanics.* Translated by Michael J. Moravcsik. Cornell, NY: Cornell University Press.

Grad, Harold (1958), "Principles of the Kinetic Theory of Gasses", in S. Flugge (ed.), *Handbuch der Physik Vol. 12* Berlin: Springer, 205.

Horwich, Paul (1987), *Asymmetries in Time.* Cambridge, MA: MIT Press.

Lanford, O. E. III (1983), "On a Derivation of the Boltzmann Equation", in J. L. Lebowitz and E. Montroll (eds.), *Nonequilibrium Phenomena: The Boltzmann Equation.* North Holland, Amsterdam, 1–17.

Lebowitz, Joel (1999), "Statistical Mechanics: A Selective Review of Two Central Issues", *Reviews of Modern Physics* 71: S346-S358.

Loewer, Barry (1996), "Determinism and Chance" forthcoming in *History and Philosophy of Science.*

Sklar, Lawrence (1993), *Physics and Chance.* Cambridge: Cambridge University Press.

——— (1973), "Statistical Explanation and Ergodic Theory", *Philosophy of Science* 40: 194–212.

---

jump of another particle in the very same gas, though it plausibly will knock the gas into a path that will ultimately bring it to equilibrium, need not put the gas into an F state: changing the position of one particle may not force the gas to follow the Boltzmann equation for the next few milliseconds. It need not even be the case that *most* jumps will put the gas into an F-state: even if F-states are very frequent in every neighborhood of every point, jumps are not random throughout neighborhoods: they move only one particle, and so are confined to hyperplanes in which the states of all other particles are held fixed. So it seems that one might hope to say that the jumps are minor perturbations—that they don't undermine F, but they don't produce it either.