Probabilities in Statistical Mechanics: What are they?

Wayne C. Myrvold
Department of Philosophy
The University of Western Ontario
wmyrvold@uwo.ca

July 15, 2012

Abstract

This paper addresses the question of how we should regard the probability distributions introduced into statistical mechanics. It will be argued that it is problematic to take them either as purely ontic, or purely epistemic. I will propose a third alternative: they are *almost objective* probabilities, or *epistemic chances*. The definition of such probabilities involves an interweaving of epistemic and physical considerations, and thus they cannot be classified as either purely epistemic or purely ontic. This conception, it will be argued, resolves some of the puzzles associated with statistical mechanical probabilities: it explains how probabilistic posits introduced on the basis of incomplete knowledge can yield testable predictions, and it also bypasses the problem of disastrous retrodictions, that is, the fact the standard equilibrium measures yield high probability of the system being in equilibrium in the recent past, even when we know otherwise. As the problem does not arise on the conception of probabilities considered here, there is no need to invoke a Past Hypothesis as a special posit to avoid it.

Keywords: Statistical mechanics; thermodynamics; probability; chance; method of arbitrary functions.

the principles of statistical mechanics are to be regarded as allowing us to make reasonable predictions as to the future condition of a system, starting from incomplete knowledge of its initial state.

Richard C. Tolman (1938, p.1).

1 Introduction

Considerations of probability enter essentially in the formulation of statistical mechanics. There are, *prima facie*, two candidates for understanding such probabilities: either they are epistemic, having to do with our incomplete knowledge of the system, or they are ontic, physical facts about the system. Both seem problematic in the context of statistical mechanics. Probabilities are introduced because of incomplete knowledge of the state of a system—which suggests an epistemic reading—but are nonetheless used to generate testable predictions, which suggests that they reflect physical facts about the system, about which we gain knowledge by testing the predictions. Furthermore, for an isolated system that starts out in disequilibrium and then relaxes to equilibrium, no probability distribution over the complete state of the system, however it is construed, evolves into a standard equilibrium distribution. One symptom of this is that the fine-grained Gibbs entropy of an isolated system does not increase. Another is the disastrous retrodictions that result from taking an equilibrium distribution seriously as a probability distribution over the microstates of the system: the standard equilibrium distributions yield high probability that the system was in equilibrium in the recent past, even when we know that it was not. What, then, is the role of these distributions?

The key to resolving these puzzles lies in the family of techniques that is known (somewhat misleadingly) as "the method of arbitrary functions." The idea is that, for certain systems, a wide range of probability distributions will be taken, via the dynamics of the system, into distributions that yield approximately the same probabilities for some statements about the system.[1] Poincaré called the input probability distributions "conventions," which leaves their status unclear. In this paper, we take up a suggestion of Savage (1973), that these be regarded as subjective probabilities. The idea is that the evolution of the system can wash out considerable differences in credences about the initial state, if we restrict ourselves, not to all propositions about the system at a later time, but a limited class of them. Typically the propositions of interest will be those whose truth-values can be ascertained via feasible measurements. We don't need to argue that *arbitrary* probability functions about initial conditions will converge to the same probabilities, nor do we need to fix a unique probability function that all rational agents would be obliged to have. What is

[1]The method of arbitrary functions was pioneered by von Kries (1886) and Poincaré (1912), and elaborated by a number of mathematicians, notably Hopf (1934, 1936). For a systematic overview of mathematical results, see Engel (1992); for the history, see von Plato (1983).

needed is approximate convergence of probabilities of outcomes of feasible measurements for all credence functions that could represent the beliefs of reasonable agents, given the sort of information available to them.

If it so happens that the dynamics of the situation wash out differences between credence-functions of any two reasonable agents, with regards to the results of feasible experiments, we cannot come to this conclusion on the basis of either epistemic considerations or physical considerations alone. On the epistemic side, we require a restriction to a class of probability-functions that can represent the belief-states of reasonable agents with limited access to information about the system, and we require a limitation to certain sorts of measurements, since, for an isolated system, no invertible dynamics will wash out *all* differences between probability functions. On the physical side, we need the right sort of dynamics. If it is the case that the dynamics lead to probability distributions that yield approximately the same probabilities for the results of certain measurements, given as input any reasonable credence-function, then the values of these probabilities will be a matter of physics. We will call the probabilities that emerge in such a scenario *almost objective* probabilities, or *epistemic chances*.[2]

Though applicability of this idea will not be limited to equilibrium situations, it sheds light on the status of the standard equilibrium distributions. If we restrict attention to a limited set of properties of the system, namely, those that can be revealed by feasible measurements, then there is good reason to believe that, once a system has relaxed to macroscopic equilibrium, the standard equilibrium probability distributions yield approximately the same probabilities for the results of such measurements as would the result of evolving any probability distribution over initial conditions represented by a density function that doesn't vary too quickly over the phase space of the system. If so, then we can use the equilibrium distributions as calculational devices to determine the epistemic chances of results of measurements, without regarding these probability distributions as yielding either objective chances about the microstate of the system, or as representing reasonable credences about the microstate. Their role is as *surrogate* of the time-evolved credences of reasonable agents. More on this in §4.3.

For ease of exposition, we will be primarily concerned in this paper with classical statistical mechanics. Many of the conceptual issues will be essentially the same for quantum statistical mechanics; in particular, if a quantum mechanical mixed state is introduced because of imperfect knowledge of the actual state of the system, then the status of such mixtures will be much the same as that of classical mixed states. It should be noted, however, that, in light of the puzzles associated with classical statistical mechanical probabilities, some authors have suggested that quantum mechanics is required to make sense of statistical mechanics. In this vein, David Albert (2000, pp. 154–162) takes recourse to state-vector collapse, regarded as a real, and chancy, physical process. Along different lines, Linden et al. (2009) suggest that it is necessary to consider the reduced state of a system that is entangled

[2] This term is borrowed from Schaffer (2007).

with its environment. David Wallace, in unpublished work, has also argued that probabilities enter into classical statistical mechanics via quantum mechanics. These proposals merit serious consideration as an account of the origin of statistical mechanical probabilities, consideration which is beyond the scope of the current paper. The position taken in this paper is that classical statistical mechanics needs no help from quantum mechanics in making sense of its probabilities; classical statistical mechanics can stand on its own.

2 Concepts of Probability

2.1 Two senses of "probability"

The word "probability" is used in at least two distinct senses. One sense, which historically is the older,[3] has to do with degrees of belief. There is another sense on which probability is associated, not with agents and their beliefs, but with physical set-ups such as gambling devices. The probability of getting a pair of sixes on a roll of the dice is spoken of as being a characteristic of the dice and the circumstances of the throw—the "chance set-up"—whether or not this probability is known to anyone (for example, unbeknownst to all, the dice might be biased). The two have much in common. For one thing, there is a formal correspondence, as an assignment of objective chances and the degrees of belief of an ideally rational agent will both satisfy the usual axioms of probability. But they are distinct, nonetheless.

The literature on interpretations of probability sometimes suggests that we have to choose one or the other of these senses, or some other sense, as *the* meaning of probability. The attitude adopted in this paper is that both of these senses are useful concepts, and we would be foolish to discard either on of them. What is important is that we not conflate the two. In what follows, we will use the word "chance" for probabilities thought of as objective properties of physical systems, "credence" when we are speaking of the a degree of belief of some (perhaps idealized) agent.

Inattention to the distinction between chance and credence carries a risk of falling into an illusion that ignorance can, by some sort of alchemy, be transformed into knowledge. Suppose that we know of no reason why a coin will land heads rather than tails, and accordingly assign these alternatives equal probability. Since there is an equal probability of heads and tails, we then predict, with a high degree of confidence, that a large number of tosses will yield roughly equal frequencies of heads and tails. Ignorance has yielded knowledge.

Distinguishing between chance and credence, it is easy to see what has gone wrong. To have one's credence equally divided between heads and tails is not the same as being certain that the chances of heads and tails are equal. Equal credence for heads and tails is compatible with a wide variety of credences about the chance of heads, as any distribution of credences about the chance of heads that is symmetrical around $1/2$ yields equal credence for heads and tails. One might, for example, have credences about the chances represented

[3] See Hacking (1975) for a masterful overview of the history.

by a flat density function. That this is different from being certain that the chance is $1/2$ can be seen from the fact that different credences are obtained for the results of multiple tosses. An agent with flat credences about the chance of heads has credences $\{1/3, 1/6, 1/6, 1/3\}$ in the possible outcomes $\{HH, HT, TH, TT\}$ of a pair of tosses, whereas an agent who is certain that the coin-toss is fair has equal credence $1/4$ in each of these possible outcomes.[4]

Approaches to credence, or epistemic probability, differ on whether two reasonable agents in possession of the same information must agree on their credences. The programme of defending an affirmative answer to this is known as *objective Bayesianism*; on this approach, the axioms of probability are supplemented by some sort of principle that uniquely fixes credences. In connection with statistical mechanics, this approach has been strongly advocated by E.T. Jaynes, who advocates the use of the Principle of Maximum Entropy as a constraint on credences. Those who countenance credence differences between reasonable agents are known as *subjectivists*, or *personalists*. This is not an all-or-nothing distinction; one could hold that there are no principles of rationality strong enough to uniquely determine rational credence, while not accepting every coherent credence function as equally reasonable. For this sort of view, which is the one adopted here, Abner Shimony (1971) has coined the term *tempered personalism*.

2.2 Learning about chances

The chance of heads on a coin toss, if it is regarded as an objective feature of the set-up, is *ipso facto* the sort of thing that we can have beliefs about, beliefs that may be correct or incorrect, better or worse informed. Under favourable conditions, we can learn about the values of chances.

Particularly conducive to learning about chances are cases in which we have available (or can create) a series of events that we take to be similar in all aspects relevant to their chances, that are, moreover, independent of each other, in the sense that occurrence of one does not affect the chance of the others. The paradigm cases are the occurrence of heads on multiple tosses of the same coin, occurrence of a six on multiple throws of the same die, and the like. Consider a sequence of N coin tosses. If, on each toss, the chance of heads is λ, then the chance of obtaining any particular sequence of results having m heads and $N - m$ tails is

$$\lambda^m (1 - \lambda)^{N-m}.$$

Considered as a function of λ, this is peaked at the observed relative frequency m/N, and becomes more sharply peaked, as N is increased.

Let E be the proposition that expresses the sequence of results of these N tosses. Consider an agent who has some prior credences about the chance of heads, and updates them by Bayesian conditionalization. Let these credences be represented by a density function

[4] In general, as can be shown by a simple calculation, a flat distribution over chances of heads renders every possibility for the number of heads in N tosses equally probable.

$f(\lambda)$; that is, our agent's credence that the chance is in an interval Δ is given by

$$cr(\lambda \in \Delta) = \int_\Delta f(\lambda)\, d\lambda.$$

We can define a conditional credence function $cr(E|\lambda)$ that satisfies

$$cr(E \,\&\, \lambda \in \Delta) = \int_\Delta cr(E|\lambda) f(\lambda)\, d\lambda.$$

Updating by Bayesian conditionalization on the evidence shifts the credence-density function:

$$f(\lambda) \;\Rightarrow\; f(\lambda|E) = \frac{cr(E|\lambda)\, f(\lambda)}{cr(E)}.$$

It seems natural to suppose—and, indeed, this is often assumed without explicit mention—that our agent's credences set $cr(E|\lambda)$ equal to what the chance of E is according to the hypothesis that the chance of heads on each toss is λ, as is required by what Lewis (1980) has dubbed the *Principal Principle*.[5] This gives, for our example,[6]

$$cr(E|\lambda) = \lambda^m (1 - \lambda)^{N-m},$$

and

$$f(\lambda) \;\Rightarrow\; f(\lambda|E) = \frac{\lambda^m (1 - \lambda)^{N-m}}{cr(E)}\, f(\lambda).$$

This has the consequence that, provided our agent's prior credences don't assign credence zero to some interval containing the observed relative frequency, her credence in chance-values close to the observed relative frequency is boosted, and her credence in other values, diminished. Moreover, since the likelihood function $\lambda^m (1 - \lambda)^{N-m}$ is more sharply peaked, the larger the number of trials, relative frequency data becomes more valuable for narrowing credence about chances as the number of trials is increased.

Note that there are three distinct concepts at play here: chance, credence, and relative frequency in repeated trials. None of these three is to be identified with any of the others. They do, however, connect in a significant way: relative frequency data furnish evidence on which we update credences about chances.

[5] Since one of the chief sources of significance of the Principal Principle is the role it plays in learning about chances, readers are urged to resist the temptation to gloss the Principle as the injunction to set one's credence equal to the chance, when the chance is known. It is when the chance is *not* known that the Principal Principle is most valuable!

[6] Recall that E expresses the particular sequence obtained, not the weaker statement that there are m heads and $N - m$ tails.

2.3 Two non-senses of "probability"

Missing from our classification are the two conceptions of probability that arise most frequently in discussions of probability by physicists. These are the classical conception of probability, founded on a Principle of Indifference, and frequentism. A few words are in order about why neither of these are satisfactory.[7]

In the classical vein, there is a temptation to attempt to define the probability of an event A as the ratio of the number of possible cases favourable to A to the total number of possible cases, if the state space is discrete, or as the ratio of the volume of the state space in which A holds to the total available volume of state space, in the continuum case. This can't be the whole story, however. In the discrete case, a judgment is needed as to which way of partitioning the state space yields equiprobable cases. In the continuum case, a judgment is required about which measure on the space is the appropriate one. We may say: a uniform measure, represented by a flat density function, but then we must remind ourselves that a probability distribution that is uniform with respect to one parameterization of the space will not be uniform with respect to others. If the appropriate distribution is uniform, then a judgment is required about which variables it is uniform in.

One of the best comments on attempts to define probability in terms of mere counting of cases occurs in Laplace's *Philosophical Essay on Probabilities*, in which an incautious formulation is first enunciated, then corrected.

> *First Principle.*— The first of these principles is the definition itself of probability, which, as has been seen, is the ratio of the number of favorable cases to that of all the cases possible.
>
> *Second Principle.*— But that supposes the various cases equally possible. If they are not so, we will determine first their respective possibilities, whose exact appreciation is one of the most delicate points of the theory of chance (Laplace, 1951, p. 11).

It is certainly true that, given a judgment that a certain partition of the state space is to be regarded as equiprobable, such a judgment fixes the probability of all boolean combinations of elements of this partition; moreover, this is a very useful fact, as it reduces a great many problems in the theory of probability to combinatorics. But Laplace's First Principle does not suffice as a *definition* of probability, since, as pointed out by Laplace himself, it requires supplementation by a judgment of which cases are equiprobable.[8]

[7] The literature on both of these topics is vast, and some readers may find the following inadequate. But a thorough discussion of these matters is beyond the scope of this paper. For further discussion of the classical conception, see, *e.g.*, van Fraassen (1989). For more on frequentism, see Jeffrey (1992), Hájek (1997, 2009).

[8] "Equally possible" is Laplace's phrase. For discussion of the meaning of this, see Hacking (1971). Laplace has been accused of circularity, on the grounds that "equally possible" can only mean *equally probable.* Even if Laplace can be defended against the charge, the point remains: what we have does not suffice as a definition of probability unless supplemented by an account of which cases are to be regarded as equipossible.

There are also attempts to identify objective probabilities with relative frequencies of events, in either actual or hypothetical ensembles of events. There is, indeed, a connection between chances and relative frequencies. For example, if we consider the case of a ball being drawn from an urn containing a number of balls, with each ball having an equal chance of being drawn, then the chance that the drawn ball will be black is equal to the proportion, or relative frequency, of black balls in the urn. We cannot, however, simply equate the chance of a black ball being drawn with this proportion, as the conclusion that the chance of drawing a black ball is equal to the proportion of black balls relies on the condition that each ball have an equal chance of being drawn, and this requires a notion of chance distinct from that of relative frequency of balls in the urn. The situation is worse for other events. In the case of drawing from an urn, it is clear that the relevant reference class should be the balls in the urn. In other cases it is less clear what the appropriate reference class should be, and what chance we ascribe to an event may vary widely depending on the reference class.

Recourse might be made to limiting relative frequencies in a hypothetical infinite sequence of repeated events of the same type; we may be tempted to define the chance of heads in a coin toss as the value that the relative frequency would converge to, if the coin were to be tossed infinitely many times. But why should we think that there is such a value? How are we to evaluate the counterfactual? Appeal may be made to the Strong Law of Large Numbers, which assures us that, in an infinite sequence of identically distributed and independent events, the relative frequency of any outcome-type will converge to a limiting value. But we need to be careful. There are, of course, possible sequences on which the relative frequency does not converge, and possible sequences on which the relative frequency converges to the wrong value. What the Strong Law says is that the set of such sequences has probability zero. This requires us to be able to ascribe probabilities to propositions regarding whether or not there will be a limiting relative frequency, and to propositions regarding the value of the limiting relative frequency, if there is one, and this requires a notion of probability distinct from limiting relative frequency. Although relative frequencies of events in repeated trials have a bearing on chances, in that they are in many cases our best evidence about the values of these chances, they are conceptually distinct from chances.

Both the classical conception and the frequency conception are attempts to introduce an objective notion of probability that is compatible with deterministic laws of nature. We will be able to sidestep the issue, as it will be argued below that, whether or not there is a fully objective notion of probability available, we can introduce a notion that will suffice to play the role, for the purposes of statistical mechanics, that objective probability is meant to play.

3 Puzzles about statistical mechanical probabilities

The state of a classical system is represented by a point in its phase space, which is specified by specifying the values of all coordinates q_i in its configuration space and their conjugate momenta p_i. These change over time according to *Hamilton's equations*:

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q_i},$$

where H is the Hamiltonian of the system, that is, the total energy of the system, expressed as a function of coordinates and momenta $\{q_i, p_i\}$. We will use x as a variable ranging over phase space points: that is, over full specifications of all coordinates and momenta. As the state of the system evolves, so too will any probability distribution over the phase space. If $\rho(x,t)$ is a probability density function over the state of the system at time t, then *Liouville's equation* expresses the time-dependence of this probability density function:

$$\frac{\partial \rho}{\partial t} + \sum_i \left(\frac{\partial \rho}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial \rho}{\partial p_i} \frac{\partial H}{\partial q_i} \right) = 0.$$

Any distribution that is uniform with respect to canonical phase space variables will be unchanging in time, as will be any distribution given by a density function that is a function only of the total energy H.

Statistical mechanics invokes certain standard probability distributions for systems in equilibrium. For an isolated system with conserved total energy known to lie in a small interval $[E, E + \Delta E]$, we use the *microcanonical distribution*, on which regions of phase space with equal phase-space volume have equal probability, provided that the energies associated with points in the region lie entirely in the interval. For a system that is in contact with a heat bath of temperature T, one uses the canonical distribution, with density function

$$\rho(x) = Z^{-1} e^{-H(x)/kT},$$

where k is Boltzmann's constant, and Z is chosen so that the function is normalized. Given a phase-space probability density ρ, we define the *Gibbs entropy*

$$S[\rho] = -k \int \rho(x) \log \rho(x) \, dx,$$

where the integral is taken over the entire phase space of the system. It follows from Liouville's equation that the Gibbs entropy of an isolated system is constant.

The question we want to ask is: are these probabilities to be regarded as subjective degrees of belief, or objective chances?

Textbook expositions typically begin by observing that the systems to which we apply thermodynamics possess a large number of degrees of freedom, and that our knowledge of the system usually consists of measured values of a relatively small number of parameters. The state of the system is, therefore, incompletely known. For this reason, we introduce a probability distribution over the states of the system compatible with what is known about it. One textbook tradition, with its roots in Tolman (1938), bases probability assignments in statistical mechanics on what Tolman calls the "hypothesis of equal *a priori* probabilities in the phase space." About this postulate, Tolman writes,

> Although we shall endeavour to show the reasonable character of this hypothesis, it must nevertheless be regarded as a postulate which can ultimately be justified only by the correspondence between the conclusions which it permits and the regularities in the behaviour of actual systems which are empirically found (Tolman, 1938, p. 59).

Tolman's argument for the reasonableness of adopting a probability distribution that is uniform with respect to canonical variables is based on Liouville's theorem, which demonstrates that such a distribution is a stationary distribution; this shows that "the principles of mechanics do not themselves include any tendency for phase points to concentrate in particular regions of the phase space" (p. 61).

> Under the circumstances we then have no justification for proceeding in any manner other than that of assigning equal probabilities for a system to be in different equal regions of the phase space that correspond, to the same degree, with what knowledge we do have as to the actual state of the system. And, as already intimated, we shall, of course, find that the results which can then be calculated as to the properties and behaviour of systems do agree with empirical findings (p. 61).

This suggests an application of a Principle of Indifference, albeit not an incautious one that disregards the need for a judgment about the variables with respect to which the distribution is to be uniform. We seem to be skirting dangerously close to the alchemy of turning base ignorance into golden knowledge. It looks as if ignorance probabilities, introduced on the basis of a Principle of Indifference, yield empirical predictions which are found to be corroborated. Indeed, this is strongly suggested by some subsequent textbooks. To take one example, in the opening chapter of E. Atlee Jackson's *Equilibrium Statistical Mechanics*, a version of the Principle of Indifference is introduced.

> to *predict* the probability of a certain event, one uses the general rule:
>
> > **If there is no apparent reason for one event to occur more frequently than another, then their respective probabilities are assumed to be equal** (Jackson, 1968, p. 8).

On the basis of this and some plausibility considerations, Jackson introduces

> **The Basic Assumption of Statistical Mechanics**. All microstates of a system that have the same energy are assumed to be equally probable (p. 83).

This is immediately followed by the remark,

> This simple *assumption* is the basis of all of statistical mechanics. Whether or not it is valid is a matter that can be only be settled by comparing the predictions of statistical mechanics with actual experiments. To date there is no evidence that this basic assumption is incorrect. A little thought shows that this agreement is indeed remarkable, for our basic assumption is little more than a reflection of our total ignorance about what is going on in the system.

It would, indeed, be remarkable if an assumption of ignorance could be turned into reliable empirical predictions. But ignorance cannot be transformed into knowledge. There must be something else going on.

We may be tempted to simply reject all of this as confusion stemming from a conflation of chance and credence. If we do so, then, it seems as if we must adopt a single reading of "probability": either epistemic or ontic. Neither one seems able to do all that we ask of it, however. Probabilities are introduced in the first place because we have incomplete knowledge of a system. Yet on the basis of these probabilities we calculate expectation values of measurable quantities, as well as probability distributions for deviations from these expectation values, and we find our expectations to be confirmed by experiment. This suggests that the probabilities in question are something more than a reflection of ignorance. We have a puzzle: how can it be that a postulate about probabilities, introduced because of incomplete knowledge of the state of the system, can be tested by experiment?

Furthermore, it seems problematic to take the standard equilibrium probability distributions as representing *either* our credences about the system's microstate, or objective chances regarding the system's microstate. For a system consisting of a large number of degrees of freedom that has been prepared in a nonequilibrium state and allowed to relax to equilibrium in isolation from its environment, there will be propositions about the state of the gas which are true and, moreover, might be known to be true, to which the equilibrium measure assigns absurdly small probability. Let A be the set of states compatible with the system's non-equilibrium macrostate at t_0. Let the system relax to equilibrium, while isolated from its environment, and let $T(A)$ be the set of states that evolve from states in A. The system's state must be in $T(A)$. Since evolution of an isolated system is measure-preserving, it has the same measure as A, which, being a non-equilibrium macrostate, has vanishingly small measure compared to the system's equilibrium macrostate.

An illustration of this is the fact that, were we to take the equilibrium distribution seriously as yielding probabilities, epistemic or ontic, about the microstate of the system at t_1, we would be led to make disastrous retrodictions. We would ascribe vanishingly low probability to the system's state at t_1 being in $T(A)$, which has the consequence of vanishingly low probability of the system's state at t_0 having been in A.

On an objective chance reading, the equilibrium distribution is clearly in trouble. If the state was in A at time t_0, then it has no chance of being in a state outside of $T(A)$ at t_1, though the equilibrium distribution assigns most of its probability to the complement of $T(A)$. On an epistemic reading, the same considerations apply, *if* that we know the dynamical map T and have no difficulty in applying it. If we are able to back-evolve credence functions, then taking the standard equilibrium distribution as representing our credences about the state of the system at time t_1 would lead us to retrodict that the system was in an equilibrium macrostate at t_0, even if we know that this is not the case. (If, however, an agent is uncertain of the dynamics, or unable to do the calculation, the equilibrium distribution *might* represent the agent's credences; more on this in §4.6).

Thus, if we ask whether statistical mechanical probabilities are epistemic or ontic, we run up against two sorts of puzzles. First, they are standardly introduced on the basis of incomplete knowledge of the system, which suggests an epistemic reading, and yet they are said to generate testable predictions, which suggest an ontic reading. We seem to require *both*, in an inconsistent way. Second, on either reading, the status of the standard equilibrium distributions is puzzling, as they seem to reflect neither objective chances nor the knowledge we actually have of the system, which suggests that *neither* reading is acceptable.

4 Hybrid probabilities: Epistemic chances, or almost objective probabilities

The notion of probability that we will be invoking is one that is neither wholly epistemic, nor wholly ontic, but intertwines considerations of both sorts. As mentioned above, we take our cue from the literature on the family of techniques that goes by the name of the "method of arbitrary functions." We consider a class of credence functions about the state of a system at time t_0, regarded as the sorts of credences that a reasonable agent could have. We examine the evolution of these credences under the dynamics of the system. It can happen that all credence functions about states of affairs at t_0 lead, via the dynamics of the system, to virtually the same credence about some propositions about the state of the system at time t.

The ingredients that we start with are:

- a class $\mathcal{C}$ of credence-functions about states of affairs at time t_0 that is the class of credences that a reasonable agent could have, in light of information that is accessible to the agent,
- a dynamical map $\mathcal{T}$ that maps states at time t_0 to states at time t_1, which induces a map of probability distributions over states at time t_0 to distributions over states at time t_1,
- a set $\mathcal{A}$ of propositions about states of affairs at time t_1, and
- A tolerance threshold for differences in probabilities below which we regard two probabilities as essentially the same.

Given a dynamical map $\mathcal{T}$, we define, for any probability function P about states of affairs at t_0 the time-evolute $\mathcal{T}^*P$, which assigns to a proposition S about states of affairs at t_1 the probability that P assigns to the state at t_0 being one that, when evolved to t_1, makes S true. If A is a subset of the system's state space, and ω the system's state,

$$\mathcal{T}^* P(\omega \in A) = P(\omega \in \mathcal{T}^{-1}(A)).$$

If, for some $S \in \mathcal{A}$, there exists a value λ such that $|\mathcal{T}^* P(S) - \lambda| < \epsilon$ for all $P \in \mathcal{C}$, then we will call this value an *almost-objective probability*, or an *epistemic chance* of S. Note that these probabilities are those yielded by evolving credences about the state at t_0 via the actual dynamics of the system, whether these are known to the agent or not. We are not talking about intersubjective agreement; the value of an epistemic chance might be unknown to all, and might not represent the credence of any agent. Under propitious circumstances, however, agents can gather evidence about the values of such epistemic chances and by conditionalizing on such evidence come to agreement (see §4.2, below).

Quantities such as these are suited to play a role analogous to objective chance, even if the underlying dynamics are deterministic. Note that the definition includes both physical considerations, in that the dynamics must be the right sort, and epistemic considerations, having to do with limitations on accessible knowledge about the system. Whether or not a given proposition S will have an epistemic chance depends both on the dynamical map $\mathcal{T}$ and on the class of $\mathcal{C}$ of reasonable credences (though, if all goes well, will not depend too sensitively on the latter); the value of an epistemic chance of S, when it exists, is largely a matter of the dynamics.

Most of the mathematical literature on the method of arbitrary functions considers limiting behaviour as some parameter characterizing the physical set-up approaches a limit, and comes to conclusions about the behaviour of arbitrary probability functions representable by a continuous density function, or arbitrary probability functions representable by a density function of finite total variation.[9] Though theorems about limiting behaviour will be of interest to us, it is not limiting behaviour that we are really concerned with, but approximate convergence of all probability functions in the class $\mathcal{C}$, for a fixed physical set-up. Limiting theorems will be of interest insofar as conclusions about such approximate convergence can be extracted from the proofs of such theorems.

A common approach to statistical mechanics invokes coarse-graining. Instead of a probability distribution over the precise microstate of the system, one adopts some coarse-graining procedure to yield a probability distribution over coarse-grained states. A probability distribution over coarse-grained states yields probabilities only for observables that are expressible in terms of the coarse-grained description. From the perspective adopted here, this can be seen as one way of specifying a limited class of propositions $\mathcal{A}$, about which there will be almost-objective probabilities.

[9]For definition, see footnote 10.

4.1 Examples

To get a sense of this method, let us begin with two examples. Our first example will be a case in which it is possible to prove rigorous theorems about washing-out of differences in credence functions. Though much simpler than any cases of genuine interest, it will give us a flavour of how things might go for more complicated systems.

4.1.1 Poincaré's wheel.

Consider the following variant on the wheel example discussed by Poincaré in *Science and Hypothesis* and in *Science and Method*, and analyzed by Poincaré in his lectures on probability (Poincaré 1912), and in considerably more detail by Hopf (1934). A roulette-like wheel is divided into a large, even number n of equal sectors, alternately colored red and black. The wheel is spun, with an imperfectly known initial angular velocity. Poincaré and Hopf consider the case in which the wheel comes to rest via friction. For ease of analysis, let us consider the frictionless case: Given a credence function over initial positions and angular velocities of the wheel, what should one's credence be that, at a time T later, a pointer fixed to the wheel's support points to a red sector?

We first consider a case in which the initial position of the wheel, as indicated by the angle θ between the fixed pointer and some reference mark on the wheel, can be known exactly, or, at least, with degree of uncertainty that is small compared to the width of one sector. A probability distribution over the initial angular velocity then determines a probability for the pointer pointing to a red sector a time t later. It follows from a theorem due to Kemperman (Engel, 1992, Th. 3.9.a) that

$$\left| \Pr_t(Red) - \frac{1}{2} \right| \leq \frac{\pi}{2nt} V_\omega,$$

where V_ω is the total variation of ϕ, a measure of how 'wiggly' the density function for ω is.[10]

[10]The total variation, not to be confused with the variance, applies to probability distributions on $\mathbb{R}$ that can be represented by a density function. Alpine hikers will find the concept intuitive. Let $f(x)$ be the probability density function of a random variable X, and imagine that it represents the altitude of terrain over which one has to hike. The total variation V_X of X is the total vertical ascent, plus the total vertical descent, that one would have to perform while hiking the length of the real line. An almost flat density function has total variation close to zero. A unimodal distribution—that is, one given by a density function that has a unique maximum—has total variation equal to twice the maximum height of f; one has to hike to the top of the peak, and then back down. For a distribution with a finite number of modes, the total variation is the height of the first peak, plus the absolute difference between the height of that peak and the height at the next local minimum, plus the absolute difference between the height at that local minimum and the next local maximum, and so on, up to the height of the last peak.

If f is continuous and piecewise differentiable, then

$$V_X = \int_{\infty}^{\infty} |f'(x)| dx.$$

For any probability distribution over ω with finite total variation, $\Pr_t(Red)$ converges to 1/2 as $nt \to \infty$. We thus have convergence to equiprobability of *Red* and *Black* for any initial probability distribution, provided it has finite total variation, if we imagine a sequence of setups with longer and longer times, or more and more sectors, or both. It is this sort of behaviour that led Poincaré to speak of "arbitrary functions."

But we are not interested in limiting behaviour; we are interested in a set-up with T and n fixed. The conclusion that we can draw is, not that $\Pr_T(Red)$ is close to 1/2 for arbitrary credence functions about initial conditions, but rather, for all those in a certain class. All probability distributions over ω for which V_ω is small compared to nT will yield approximately the same probability for *Red* at time T, namely, 1/2.

If nT is large enough, this will not be a particularly stringent condition on credences about the initial conditions, and it may very well be one that we can expect all reasonable credences to satisfy. Let ϵ be our threshold for regarding two probabilities to be essentially the same. Then, if all credence-functions in $\mathcal{C}$ have $V_\omega < 2nT\epsilon/\pi$, we will have

$$\left|\Pr_T(Red) - \frac{1}{2}\right| < \epsilon,$$

for any time-evolute Pr_T of any credence function in $\mathcal{C}$.

We can also consider the case in which the initial position is not known precisely. Unsurprisingly, uncertainty about initial position cannot increase predictability, provided that the agent's credences about initial position and angular velocity are independent. Suppose that this is the case, that is, we take it that the spin imparted to the wheel by the croupier does not depend on the initial position of the wheel. For such a case, the result cited above still holds (Engel, 1992, Prop. 3.12).

Uncertainty about velocity will, as time goes on, increase uncertainty about the position of the pointer. Take T large enough that the evolution of our agent's credences from the time the wheel is spun, t_0, to time $t_1 = t_0 + T$ yields a credence-function for position that closely approximates a uniform distribution over the circle. This limiting distribution is stable, and will be maintained as time goes on. Now consider the fact the dynamics of the system are invariant under time-reversal. Does it follow that we can apply the analysis to the agent's credences about the state of the system at time t_1 in the reverse direction, to get the conclusion that her credence about the state of the wheel at t_0 are also closely approximated by a uniform distribution?

If f is piecewise differentiable but not continuous, we add to this the sum of the absolute values of the jumps that f makes at each of its points of discontinuity. In general, we define the total variation as follows. Take any numbers $x_0 < x_1 < ... < x_n$, and consider the quantity

$$\sum_{i=1}^{n} |f(x_{i-1}) - f(x_i)|.$$

V_X is defined to be the essential supremum of this quantity, over all choices of $x_0, ..., x_n$.

No. We can conclude that uncertainty about velocity increases uncertainty about position, for times after t_0, on the condition that our agents' credences about position at t_0 are independent of her credences about velocity. If this independence condition holds at some time t_0, it cannot hold at other times. For instance: if our agent is certain that the position of the wheel at t_0 is θ_0, but uncertain about what its angular velocity is at that time, then (provided that the wheel spins freely between t_0 and t_1), she is uncertain about both position and angular velocity at time t_1, but these are not independent uncertainties; conditional on the supposition that the angular velocity is ω, she is certain that the position at time t_1 is $\theta_0 + \omega T \pmod{2\pi}$. The conditions of the theorem cited above are not met.

Consider, now, the credence-function that results from the following recipe. We start with the credence-function cr_0 of an agent who knows the position of the wheel at t_0 exactly (say, $\theta_0 = 0$), but is uncertain of the velocity. We evolve this credence-function forward to a time $t_1 = t_0 + T$, to form a credence function cr_1. Suppose that T is large enough that cr_1's distribution for position closely approximates a distribution uniform on the circle. Form a new credence function cr_1^R, the temporal reversal of cr_1, which ascribes to a set A of states the credence that cr_1 ascribes to the set of states that results from reversing all velocities in A. Must this new credence function be a reasonable one if cr_0 is?

A moment's reflection shows that the credence-function cr_1^R would be an odd belief-state to have, about the state of the system at t_1 or any other time. It would be the belief-state of an agent who was highly uncertain about the position of the wheel, uncertain also of its velocity, but certain that, whatever the position is, its velocity is precisely the velocity required to make the reference mark on the wheel line up with the fixed pointer after a lapsed time T. There might be physical set-ups for which such a belief state would be reasonable, but the set-up we have described is not one of them.

The moral to be drawn is that the temporal reversal of a reasonable credence function need not be a reasonable credence function. In the example as we have described it, we have assumed that the agents have some way of knowing, with high precision, the initial position of the wheel, but no way of knowing future positions except by prediction based on applying the dynamics of the system to credences about its current state. We have a temporal asymmetry that is due, not to temporal asymmetry of the dynamics, but to temporal asymmetry in the class of credence functions deemed reasonable.

It can be shown that,[11] though θ_t and ω can be independently distributed at most one time (except in the special case of uniform distribution of θ, which is a stationary distribution), the covariance of these variables tends to 0 as time goes on, for any initial distribution with V_ω finite. This means that, if we replace the complicated joint distribution of θ_t and ω by one on which these two quantities have the same marginal distributions but are independent, we will not go far wrong in calculating quantities that are functions of θ_t alone, of ω alone, or of the product of the two. Thus, though we do not regard these quantities as independent, for certain purposes this independence will make little difference,

[11]Those who are interested may contact the author for details of the proof.

and we can use a distribution on which they are independent as a surrogate for their joint distribution.

4.1.2 The Boltzmann equation

Suppose we have a gas, not too dense, consisting of identical molecules that repel each other strongly at short distances but exert negligible forces on each other at longer distances. The evolution of each molecule can be approximated by free motion punctuated by occasional collisions that produce abrupt velocity changes. Suppose we start with a probability function over the state of the gas on which each molecule is identically, independently distributed, and forward-evolve it according to the dynamics of interaction between the molecules. Since whether or not a molecule undergoes a collisions depends on where the other molecules are, independence will not be preserved under evolution; the distribution evolves into a distribution with complicated dependencies between the states of molecules.

If, however, we consider a quantity, such as total kinetic energy, which is a sum of quantities pertaining to single molecules, we do not need the full distribution to calculate its expectation value. We can form single-particle distributions by integrating over the degrees of freedom of all molecules but one. Similarly, we can form two-particle distributions, three-particle distributions, *etc.*.

If interaction is negligible except for brief collisions, and if the gas is rarified enough that the probability of three-body collisions is negligible, the evolution of a one-particle distribution function will depend only on itself and the two-body distribution function, and not higher-order distribution functions, the evolution of a two-particle distribution function will depend only on itself and the three-particle distributions function, and so on. This leads to a hierarchy of coupled equations called the *BBGKY hierarchy*.

Suppose that the term involving two-particle distributions in the equation for the one-particle distribution is virtually the same as what would result from replacing the two-body distribution by the product of one-particle distributions. This is not as far-fetched as it might seem at first; it means that probability that a molecule will collide with a particle with velocity $\mathbf{v}_2$, conditional on the first particle having velocity $\mathbf{v}_1$, does not change appreciably as we vary $\mathbf{v}_1$. This amounts the assumption that, though collisions build up correlations between molecules, the effect of any given particle on the others becomes so distributed throughout the gas that the velocities of particles about to collide with any given molecule are effectively independent of the velocity of the molecule.[12] If this is correct, then we can stop the BBGKY hierarchy at the first level, yielding *Boltzmann's equation.* Boltzmann's equation has the result that any initial distribution of velocities approaches the Maxwell-Boltzmann distribution.

Though rigorous theorems are not as easy to come by as they are for the Poincaré wheel, it is certainly plausible that the effective lack of correlation between molecules, as far as the evolution of the one-particle distribution is concerned, will hold for a wide class of

[12] This is an instance of what Wallace (2011) has called *the Bold Dynamical Conjecture.*

initial probability distributions. If this is correct, then all such distributions will agree, not only that the distribution of velocities will approach the equilibrium distribution, but also on the rate at which this occurs.

4.2 The first puzzle resolved: testing hypotheses about almost-objective probabilities

For real systems having a macroscopic number of degrees of freedom, we may not know the exact Hamiltonian, and, even if we did, typically could not do the detailed calculation required to evolve a non-equilibrium probability function over a substantial period of time. An agent might, nevertheless, believe that, for some propositions about the system, there exist almost-objective probabilities, even if she doesn't know the values of these probabilities. In such a case, our agent can entertain hypotheses about what these values are, and have degrees of belief in such hypotheses. We can imagine her wondering what advice would be given by a Laplacean oracle that knew the dynamics of the system in question and could evolve her credences about states of affairs at t_0 into credences about states of affairs at t_1, and she can have credences about what that advice would be.[13] That is, epistemic chances, like objective chances, are the sorts of things that our agent can have credences about. Let A be some proposition about the system, and let our agent's credences about the value λ of the epistemic chance of A be given by a credence-density function $f(\lambda)$. She will also have credences of the form $cr(A \,\&\, \lambda \in \Delta)$, and these can be used to define a conditional credence function $cr(A|\lambda)$. Suppose, now, that her credence in these things is such that her credence in A, conditional on the supposition that the oracle would recommend credence λ in A, is equal to λ. That is,

$$cr(A|\lambda) = \lambda.$$

This condition is an analog of the Principal Principle, applied to epistemic chances; call it the PPE. It is reasonable to expect our agent's credences to satisfy this. If her credences about state of affairs at time t_1 are not the evolutes of her credences about states of affairs at t_0, this is either a failure of coherence or reflects uncertainty (due to uncertainty about the actual dynamics, or what the result of applying those dynamics would be) about what would result from evolving her credences about the state at t_0. We can gloss the PPE as saying that, if our agent had access to a Laplacean oracle, learning consisting of conditionalization on its pronouncements would result in acceptance of its recommendations.

If A is a proposition whose truth or falsity can be ascertained by measurement or observation, and we have multiple copies of the system (where a copy means that the two systems have the same dynamics and that our agent's credences about initial conditions are essentially the same for both), then, provided her credences satisfy the PPE, our agent can

[13] This is a calculating oracle. We don't imagine that the oracle knows what the actual state of the system is; rather, it tells one what credences about states of affairs at time t_1 are consistent with one's credences about states of affairs at time t_0 and the laws of physics.

do experiments and update her credences about what the recommendations of the oracle would be, in a manner precisely analogous to learning about chances, as outlined in §2.2. This will have the consequence that credence in hypotheses that accord higher epistemic chance to the observed results will be boosted relative to credences that accord them lower epistemic chance. With sufficient evidence, we can end up with arbitrarily high confidence that we know what the actual epistemic chances are.

4.3 The second puzzle resolved: the status of equilibrium distributions

Our puzzle was: If the standard equilibrium probability distributions represent neither objective chances nor our credences, why, then, do we bother with them?

Consider again a Laplacean oracle, with perfect knowledge of the dynamics of some isolated thermodynamic system, and unlimited capacity to perform calculations on the basis of these dynamics. Suppose that the system is at time t_0 in some non-equilibrium state, and that it relaxes to an equilibrium macroscopic state by time t_1. Consider Bob, an agent with finite cognitive capacities, who is able to ask questions of the oracle. If Bob tells the oracle the credence-function representing his beliefs about the state of the system at time t_0, the oracle will be able to evolve this probability distribution to time t_1. Knowing the result of such a calculation would tell Bob what his credences about the state of the gas at t_1 ought to be, given his credences about its state at t_0 and taking full advantage of the oracle's expertise in classical mechanics. It is not at all clear, however, that the oracle would be able to communicate this distribution, in all its detail, to Bob, or that it would be of much use to Bob if he had it, as the distribution would be monstrously complex; for one thing, the support of the distribution might be a complicated, finely fibrillated shape threading through the system's phase space. If, however, Bob's interests are in forming expectations about the results of feasible measurements (we might imagine him making bets on the outcomes of such measurements), then the oracles's communication task is much simpler. For most practical purposes, the oracle could simply provide Bob with the expectation value and dispersion of a finite number of measurable quantities.

It is conceivable—and, in fact, we have good reason to believe that this is the case for the sorts of systems to which we successfully apply thermodynamics—that the oracles's communication task will be even simpler than this. Even though the time-evolute of Bob's credences about the gas at t_0 is enormously complex, if the dynamics are of the right sort there will be simplicity in this complexity, because the support in phase space of Bob's credence about initial conditions will end up so finely spread through phase space that the probabilities of results of macroscopically feasible measurements will not differ appreciably from those yielded by an equilibrium distribution that the demon can communicate succinctly, and Bob will be able to calculate with that.

When this obtains, then the fact that Bob knew the system to be in a non-equilibrium state has negligible bearing on the probabilities of outcomes of feasible measurements yielded by the time-evolved probability function t_1. What this requires is the following plausible hypothesis.

> *Hypothesis of Washing-Out of the Past.* Once a system has relaxed to a macroscopic equilibrium state, no feasible measurement will be informative about its past state.

If this is true, then the dynamical process of relaxation to equilibrium will wash out very considerable differences between credence-functions over initial conditions, including credence functions that attribute very different non-equilibrium macroscopic states to the system at earlier times, and the equilibrium probability distribution can be used as a surrogate for the time-evolved credences not only of Bob, but of Alice, who believes the system to have been in a wholly different nonequilibrium macrostate at t_0, and of Charles, who takes the system to have been in equilibrium at that time. The time-evolutes of all these credence-functions will yield indistinguishable expectations about the results of measurement from those yielded by a stationary distribution, such as the standard equilibrium distribution.

This, then, will be the role of the standard equilibrium distributions. Though these typically discard much of what we know about the past of the system, for certain, limited purposes, namely, calculation of probabilities of the results of feasible experiments, this discarded information will be irrelevant, and the equilibrium distribution can be used as a surrogate for our time-revolved credences.[14]

Typically, there will be a temporal asymmetry in the use of equilibrium distributions. We will in a great many cases be able to use them, as surrogates for the credences of reasonable agents, to make predictions. We will rarely, if ever, be in a position to use them for retrodictions. Since we will typically be dealing with systems whose dynamics are symmetric under time-reversal, we may very well wonder where the source of the asymmetry lies. If, at time t_1, we can attach some epistemic chance value to the result of some measurement performed some time after t_1, does it not follow that the same epistemic chance must attach to the result of that measurement if performed the same amount of time *before* t_1?

Though the dynamics of the system may be time-symmetric, the dynamics are not the only ingredient that go into epistemic chances. Another ingredient is the class $\mathcal{C}$ of reasonable credence functions. As we have already pointed out, there is no reason to expect this to be invariant under time-reversal, and, typically, it won't. We have knowledge, in the form of records or memories, about the past, that is of a different nature from the sort of

[14] *Cf.* Oliver Penrose,

> we may say that in the limit of large t the true ensemble and the equilibrium ensemble become observationally equivalent with respect to single observations. It will be shown ... that this equivalence extends to compound observations as well, so that the equilibrium ensemble can be used in place of the true ensemble for all calculations of equilibrium properties. Owing to the particularly simple form of the equilibrium ensemble, this equivalence property is extremely valuable (Penrose, 2005, p. 135).

knowledge we can have about the future.[15] If any reasonable agent can have non-negligible credence that the system was in a nonequilibrium macrostate in the recent past, then an equilibrium distribution will not suffice as a surrogate for that agent's credences, if used to compute probabilities for past events.

David Albert (2000) blocks retrodictions that would result from using equilibrium probabilities for retrodiction by introducing the *Past Hypothesis*, which 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" (p. 96). This is required as a patch on a flawed Statistical Postulate that privileges uniform distributions and leads to disastrous retrodictions. The Statistical Postulate that Albert initially introduces, which recommends a probability distribution as uniform as possible, consistent with the macroscopic state of a system, would undermine ordinary inferential procedures, as it would lead us to regard everything we take to be a record of the past to be very probably an illusion, something that arose recently via a random fluctuation from equilibrium, or worse, a hallucination in the imagination of an isolated brain or whatever minimal system is required to support consciousness. It is for this reason that a special postulate is introduced that enjoins us to trust normal inferential procedures as applied to cosmology.

In this paper, we have not taken a starting point that privileges uniform distributions, and so have not needed to apply a patch to a flawed initial postulate. The use of equilibrium distributions for inference is limited to circumstances in which they lead to essentially the same inferences as would the time-evolved credences of any reasonable agent. The credences of reasonable agents who know about cosmology are adapted to whatever conclusions are arrived at from cosmological data via normal inferential procedures, but there is no reason to single out cosmology in this regard. Our advice to agents, if advice were needed, would be: attach high credence to whatever conclusions are arrived at via normal inferential procedures applied to anything you know.

4.4 Statistical Regularity and Mixing

Let Γ to be the phase space of some system, and let $\mathcal{S}$ be a σ-algebra of subsets of Γ, which we will take to be the measurable subsets of Γ. Let T_t be a flow on Γ, that is, a family of mappings $T_t : \Gamma \to \Gamma$ such that $T_{r+s} = T_s T_r$ for all $r, s \in \mathbb{R}^+$.[16] We assume that, for every t, T_t is a measurable function.

Following Hopf (1934), we will say that $A \in \mathcal{S}$ is *statistically regular* iff there is a probability function Q such that, for every probability function P on $\langle \Gamma, \mathcal{S} \rangle$ that can be represented by a density function, $P_t(A) \to Q(A)$ as $t \to \infty$. This means: differences in probabilities over initial conditions wash out, and we have convergence on a common limiting probability for the state being in A. We will also say that a measurable function

[15] Needless to say, no *explanation* of this fact has been offered in this paper. But it is a fact!

[16] Note that we are not requiring the flow to be invertible, though in most cases of interest it will be.

f is statistically regular iff there is a probability function Q such that, for any probability function P on $\langle \Gamma, \mathcal{S} \rangle$ that can be represented by a density function, the expectation value of f with respect to P_t converges to the expectation value of f with respect to Q.

A flow T_t is said to be *strongly mixing* (sometimes just *mixing*) if and only if, for any sets $A, B \in \mathcal{S}$,

$$\lim_{t \to \infty} \mu(T_t A \cap B) = \mu(A)\mu(B), \tag{1}$$

where μ is Lebesgue measure.

It can be shown that a measure-preserving flow is strongly mixing if and only if every $A \in \mathcal{S}$ is statistically regular, or, equivalently, if every measurable function f is statistically regular. There is thus a close connection between mixing and the concept of statistical regularity.

We say that there is an almost-objective chance that the phase point is in A, at a time $t_1 = t_0 + T$, if, for every probability function P in the set $\mathcal{C}$ of reasonable credence functions, $|P_T(A) - Q(A)| < \epsilon$, where ϵ is our threshold of tolerance for regarding $P(A)$ and $Q(A)$ to be "close enough." We do not require equality in the limit, nor do we require closeness for all input probability functions. However, it will typically be the case that from a proof of statistical regularity one can extract explicit bounds on the rate of convergence, as we have seen above, in our discussion of the Poincaré wheel. If these bounds involve some quantity, as, in our case, total variation, that is relevant to the reasonableness of a credence function, then it will often be the case that we will be able to extract from the proof a criterion for a proposition to have an almost-objective chance.

4.5 Non-equilibrium probabilities

Non-equilibrium statistical mechanics deals with systems that are initially out of equilibrium and evolve towards equilibrium, or else are maintained in a non-equilibrium state by external influences.

Though there is a greater diversity of approaches to nonequilibrium statistical mechanics than there are to equilibrium statistical mechanics, there are some commonalities. A device that is frequently used is the postulate of *local equilibrium.* Even if the macroscopic properties of a system vary from place to place (say, there is a temperature or pressure gradient), we might be able to treat the system, over small regions, as approximating an equilibrium state corresponding to the local values of macroscopic parameters. This will be appropriate when local equilibration rates—which, from our perspective, are the rates at which differences in credences over initial conditions wash out—are fast compared to the rate at which macroscopic parameters equilibrate. When this obtains, the justification for use of the postulate of local equilibrium probability densities will be much the same as it is in the equilibrium case.

Weaker than the postulate of local equilibrium is the practice, already mentioned in the discussion of the Boltzmann equation, above, of treating certain correlations as negligible for the forward evolution of the probability distribution. This is a procedure that Sklar (1993, p. 93) refers to as "rerandomization." If the legitimacy of this procedure required the variables whose correlation is being treated as negligible to be probabilistically independent, then, as Sklar rightly emphasizes, this would inconsistent with the dynamics of the system. But we do not require independence; what we require is that these correlations be small enough to be negligible for all reasonable probability distributions over initial conditions. We have already seen, in the example of Poincaré's wheel, a case in which one variable is a function of another, and yet this dependence does not show in any tendency of one to increase or decrease as the other increases. In this case it can be proven that the covariance becomes small, and explicit bounds can be put on the covariance, in terms of the total variation of the input probability functions.[17] In other cases, we have, not rigorous theorems, but plausibility arguments, based on a combination of physical considerations and numerical simulations.

4.6 Non-Liouvillean evolution and entropy increase

So far we have been talking about isolated systems. Isolation is at best approximate. If we begin with a probability distribution over system + environment and evolve it, then, though the evolution of the probability distribution over the state of system + environment is measure-preserving, the marginal distribution of the system (that is, the distribution that results from restricting one's attention to the degrees of freedom of the system and averaging out the degrees of freedom of the environment) may increase in entropy.

The system will typically be small (few degrees of freedom) relative to the environment with which it interacts. There are a number of arguments, of varying rigor and generality, that show that, if the system interacts only weakly with its environment, so that the total Hamiltonian can be approximated by the sum of the system's energy and the environment's, then, provided the system's degrees of freedom are few compared to the environment's, the marginal distribution for the system that is yielded by a microcanonical distribution over system + environment closely approximates a canonical distribution.[18] The result holds for both classical and quantum mechanics, and is often presented as a justification for using the canonical distribution for a system in thermal contact with a heat bath. Recently, this result has been generalized in interesting and significant ways. Although the usual quantum argument assumes a uniform distribution (a totally mixed state) for system + environment on its energy subspace, it has been shown that this assumption can be considerably weakened. When the environment is much larger (that is, has a much higher dimensional Hilbert space) than the system, then most pure states of the system + environment yield reduced states for the system that closely approximate a canonical distribution.[19] This has the consequence

[17] Again, interested readers may contact the author for details.

[18] See, *e.g.*, Gibbs (1902, pp. 181–83), Khinchin (1949, §20), Thompson (1988, §2.4).

[19] This follows from a theorem of considerable generality due to Popescu et al. (2006), which encompasses restrictions on the global state other than the usual restriction to a fixed energy subspace. Significant earlier work related to this issue includes Goldstein et al. (2006). See Lloyd (2006) for discussion, and further

that any probability distribution over the state of the system + environment will yield, as marginal for the system, an approximation to a canonical distribution, except for those probability distributions that give large weight to the exceptional pure states that yield marginals for the system that aren't close to canonical.

This result has a classical analog. It can be shown that most (in some sense of 'most') classical probability distributions over a composite of a system weakly coupled to a much larger environment yield marginals for the system that approximate the canonical distribution (Plastino and Daffertshofer, 2008). This result was proven for a discrete classical state space; it would be interesting to see an extension of the result to distributions over a continuous classical phase space.

The upshot of all this is that, even if our agent's credences about the state of a system and its environment differ wildly from the standard statistical mechanical probability distributions, her credences about the system itself might be closely approximated by the canonical distribution, which can then be used a as a proxy for her actual credences when calculating probabilities of the results of subsequent measurements.

Thus, for non-isolated systems, it *can* be the case that the standard statistical mechanical probability distribution does reflect the agent's credences about the microstate of the system, considered as a marginal distribution derived from a distribution about the system and environment. Does this lead to disastrous retrodictions? The canonical distribution, applied at time t_1, would yield high probability that the system was in equilibrium at time t_0, some short time before, if it were further stipulated that the system was isolated in the interim. If it was not isolated, then the dynamics of the system alone entail nothing about its state at earlier times; we need also to specify how its environment acted upon it. Our agent can, without contradiction, have credences about state of affairs at t_1 such that her credences about the system are closely approximated by the canonical distribution (an equilibrium distribution), while having credences about the state of system+environment that yield, when back-evolved, high probability to the system not having been in equilibrium at time t_0. Once again, we can get away with using an equilibrium distribution for the state of the system at t_1, without encountering disastrous retrodictions.

There is another way for the entropy of a credence-function to increase. As mentioned, our agent will typically not know the precise dynamics of the system, and, even these are known, she may be unable to perform the calculation. Thus, the agent will be uncertain about what the result is of evolving her credence-distribution ρ_0 about the state of the system at t_0, via the actual dynamics, from t_0 to t_1. She may entertain conjectures about what the result would be, and, provided that her credences satisfy the PPE, her credences about states of affairs at t_1 will be a weighted average of the candidates for what the time-evolute of ρ_0 is. If she is sufficiently uncertain about this, then her credence-distribution about states of affairs at t_1 might be uniform, or nearly so, over the region of phase space accessible to the system at this time. Again, there is no temptation to make disastrous retrodictions.

references.

Each of the candidates for the evolute of ρ_0 back-evolves, via the conjectured dynamics associated with it, into ρ_0, even though the mixture of these candidates is invariant under each candidate evolution.

In case this sounds confusing, here's another way of expressing the point. Suppose there is some set $\{\mathcal{T}_\alpha\}$ of dynamical maps, such that Alice believes that one of these yields the actual dynamics of the system. Suppose that her credence-distribution over this set is given by a density $\mu(\alpha)$. Then her credence about the state of affairs at t_1 is given by

$$\bar{\rho}_1 = \int \mathcal{T}_\alpha \rho_0 \, d\mu(\alpha).$$

The averaged distribution $\bar{\rho}_1$ could be an equilibrium distribution. This would mean that, for each $\mathcal{T}_\alpha$, $\bar{\rho}_1$ is invariant under $\mathcal{T}_\alpha$, and hence $\mathcal{T}_\alpha^{-1}$ also leaves it invariant. But, though $\bar{\rho}_1$ is an invariant distribution that represents Alice's credences about the state of the system at t_1, it doesn't follow that it reflects her credences about the state of the system at t_0, as, for each α, $\mathcal{T}_\alpha \rho_0$ back evolves, via $\mathcal{T}_\alpha^{-1}$, into ρ_0.

5 What is Thermodynamics?

Our account of statistical mechanical probabilities combines physical considerations with considerations about feasible measurements and reasonable credences. When the latter are introduced in the context of physical theory, there is resistance from some quarters. We can expect to hear objections that notions such as measurement and credence have no place in a fundamental physical theory.

A goal of statistical mechanics is to recover thermodynamics in a macroscopic limit. In discussing what sorts of consideration are and aren't appropriate to introduce into statistical mechanics, it is worthwhile to pause and consider the nature of thermodynamics.

Thermodynamics begins with a distinction between two modes of energy transfer. Energy may be transferred from one system to another as heat, or via work being done by one system on another. The First Law of Thermodynamics says that any change in total energy of a system can be partitioned into these modes of energy exchange. The Second Law says that heat extracted from a system cannot be converted without residue into work.

If heat were a substance, then the heat content of a system would be a function of the state of a system, and there would be no question of how to partition a change of energy into a part due to heat exchange and a part due to work done. A fundamental fact of thermodynamics is that energy that enters a body as heat can be extracted as work, and *vice versa*. On the kinetic theory of heat, heating and doing work are both processes in which parts of one body act on parts of the other to change their state of motion. The difference is that, when we do work on a body, say, by raising a weight, we move its parts in a systematic way, that we can keep track of; in heating, motion is imparted to the molecules of a body in a haphazard way, and the added energy is quickly distributed among the molecules that make up the body. There is no way to keep track of it in such a way as to wholly recover this energy as work.

This suggests that the distinction between heat and work is relative to the means available to us; it is a matter of what *we* can keep track of. This, in fact, was Maxwell's view. "Available energy is energy which we can direct into any desired channel. Dissipated energy is energy we cannot lay hold of and direct at pleasure, such as the energy of the confused agitation of molecules which we call heat" (Maxwell 1878a, p. 221; Niven 1965, p. 646).

In the opening section of his *Philosophical Essay on Probability*, Laplace famously invited the reader to consider a being that knew the precise state of the world at an instant and all the laws of nature, and was capable of performing the requisite calculations; "for it, nothing would be uncertain and the future, as the past, would be present to its eyes" (Laplace, 1951, p. 4). The point of this passage is to explain why, in spite of the presumed deterministic nature of the laws of physics—which Laplace seems to have taken as an *a priori* truth—there should be such a subject as probability theory. Laplace's answer is that, though such an intelligence would have perfect knowledge of all events, past, present, and future, our own state of knowledge will always remain infinitely removed from this ideal. Hence, we must use probability theory in order to cope, systematically, with less-than-perfect knowledge.

Maxwell, in his *Theory of Heat* (1871, pp. 308–309), invites us to imagine a being that could keep track of and manipulate individual molecules. The actions of such a being would not, according to Maxwell, be subject to the Second Law of Thermodynamics. Moreover, the very concepts needed to express the law would not occur to it; "we have only to suppose our senses sharpened to such a degree that we could trace the motions of molecules as easily as we now trace those of large bodies, and the distinction between work and heat would vanish, for the communication of heat would be seen to be a communication of energy of the same kind as that which we call work" (Maxwell, 1878b, p. 279). Just as Laplace's demon would have no use for probability theory, Maxwell's demon would have no use for the science of thermodynamics.[20]

Thermodynamics, in its very formulation, requires a distinction between those aspects of a system that are within the scope of our knowledge and control, and those that are not. It is because of our inability to keep track of and manipulate individual molecules that we regard some processes as dissipative. This distinction is reflected in our statistical mechanical treatment of macroscopic systems. In statistical mechanics, we distinguish between variables that we regard as known (or knowable), and use to define a thermodynamic state, and those over which we define a probability distribution. We also distinguish between types of interaction between a system and the rest of the world. Consider a gas in a container that is in thermal contact with a heat bath. The gas exerts a pressure on the walls of the container; this is due to the forces of repulsion between the molecules of the gas and the walls of the container. But the gas is also in thermal equilibrium with the walls of the container, which,

[20]For more on Maxwell's view of statistical mechanics and thermodynamics, see Myrvold (2011).

if non-insulating, may conduct heat from a heat bath. The walls have finite temperature; the molecules that make up the walls are fluctuating about their equilibrium position, and this means that the forces exerted on a molecule that approaches the wall will also be subject to fluctuations about their mean values. Implicitly,[21] we partition the interaction of the gas with the walls of the container into two terms, a mean value associated with the macroscopic position of the walls, and a term responsible for thermalization of the gas. It is via the former that we do work on the system, by manipulating the macroscopic position of the walls; energy transfer via the latter is regarded as heat transfer. This distinction between two sorts of interaction is essential to any statistical mechanical construal of a distinction between heat and work.

This doesn't mean, of course, that we can only apply thermodynamics to systems that are within our reach and which we are actually manipulating. Means-relative distinctions are distinctions about the *sorts* of manipulations we can perform, whether or not, in a particular instance, we are in a position to do so. If two systems of unequal temperature come into thermal contact and equilibrate, energy has been dissipated and there has been an increase of thermodynamic entropy, because a temperature difference that *could have* been used to do useful work is no longer there. This is true even if the system is separated from us by a vast gulf of space or time.

If the goal of statistical mechanics is to recover thermodynamics, and if the very concepts with which the laws of thermodynamics are formulated only make sense with reference to some agent's capacity to keep track of and manipulate molecules, then it should come as no surprise that such considerations enter into the formulation of statistical mechanics. Yet there are laws of thermodynamics, which talk about an agent can and cannot achieve, and these depend only on very general assumptions about what the agent is able to do—according to Maxwell, the second law of thermodynamics is valid only insofar as molecules are not being manipulated individually or in small numbers. The conception of probability outlined in this paper—incorporating both epistemic and physical considerations—thus seems appropriate for the goal of recovering thermodynamics, so conceived.

6 Explaining Relaxation to Equilibrium

The preceding will not still all qualms about introducing subjective considerations into a physical context. It might look as if we are giving up on the goal of explaining thermodynamic behaviour of physical systems, as our knowledge and beliefs about such systems are surely not to be included in an explanation of why they behave as they do![22]

It is certainly correct that considerations of limitations of our knowledge and manipulative prowess are out of place in explanations of the behaviour of systems (except those

[21] And sometimes explicitly; see, *e.g.*, Thompson (1988, §2.5).

[22] For a particularly vivid expression of this point, see Albert (2000, pp. 54–65); see also Loewer (2001, p. 611).

that we happen to be manipulating). A system behaves as it does because of its dynamics, together with initial conditions. Explanations of relaxation to equilibrium will have to involve an argument that the dynamics, together with initial conditions of the right type, yields that behaviour, plus an explanation of why the sorts of physical processes that give rise to the sorts of systems considered don't produce initial conditions of the wrong type (or rather, don't *reliably* produce initial conditions of the wrong type).

There is a connection, however, between the epistemic considerations advanced in this paper, and what would be required of an explanation of relaxation to equilibrium. The processes that are responsible for relaxation to equilibrium are also the processes that are responsible for knowledge about the system's past condition of non-equilibrium becoming useless to the agent. Thus, an explanation of relaxation to equilibrium is likely to provide also an explanation of washing out of credences about the past. Moreover, an explanation of why no process reliably produces initial conditions that lead to anti-thermodynamic behaviour would also explain the reasonableness of credences that attach vanishingly small credence to such conditions. Our judgments about what sorts of processes occur in nature and our judgments about what sorts of credences are reasonable for well-informed agents are closely linked; if there were processes that could reliably prepare systems in states that lead to anti-thermodynamic behaviour, then it would not be unreasonable for an agent to attach non-negligible credence to the system having been prepared in such a state, and we would adjust our judgments about what are and are not reasonable credences accordingly.

Furthermore, reflection on the considerations above help us to see what is wanted from such an explanation. If it can be shown that, given fairly weak constraints on an initial probability distribution and on the dynamics, the distribution evolves into one that is, with respect to probabilities of results of macroscopic measurements, effectively indistinguishable from an equilibrium distribution, this would show that weak assumptions about the initial state and the dynamics entail that the system evolves into a state that is macroscopically an equilibrium state. Of particular note in this regard is a series of recent results regarding approach to equilibrium in quantum mechanics. For a very broad class of Hamiltonians (namely, those with nondegenerate energy gaps), the reduced state of a small subsystem of a large quantum system will equilibrate (Linden et al., 2009), provided only that the state of the large system be composed of a large number of energy eigenstates. The equilibrium state is independent of the precise initial state of the bath. Also of interested is the result due to Goldstein et al. (2010), which demonstrates approach to macroscopic equilibrium for arbitrary initial states and "typical" Hamiltonians. One expects that there are ways to extend these results to classical mechanics; to work out exactly what assumptions about the classical system are required would take some care but would likely yield insight into conditions of equilibration in classical mechanics.

7 Conclusion

It has been argued that epistemic chances are the right sort of thing to do what is required of probabilities in statistical mechanics. To show that they actually do the work we need, we need to establish approximate convergence of probability results for realistic systems. This leaves us with a well-motivated research programme, namely, examination of the conditions under which the dynamics of a system will yield almost-objective probabilities—not in the infinite long run, but in finite time. Results concerning limiting behaviour will, of course be relevant, as they can often be rendered informative about finite time behaviour; in many cases results about convergence of probability distributions can yield bounds on closeness at finite times, as has been illustrated in our discussion of the Poincaré wheel.

The familiar dichotomy of epistemic and ontic probability does not fit well with statistical mechanics, which requires use of probabilistic concepts though the state evolution is deterministic (this remains true of quantum statistical mechanics as currently practiced, which deals with deterministic, unitary evolutions rather than dynamical state reduction), and turns assumptions about probabilities into verifiable empirical predictions. This is not surprising if we consider thermodynamics, the science that it is the goal of statistical mechanics to recover in approximation, to involve a mix of physical considerations and considerations regarding what it is in our power to keep track of and manipulate. Epistemic chances, whose very characterization requires both considerations of epistemic limitations and physical dynamics, seem to be just what are required for the purpose.

8 Acknowledgments

I thank John Norton for helpful comments on an earlier draft. The members of the University of Western Ontario Philosophy of Physics Reading Group and audiences at various presentation provided useful feedback. I am grateful also to David Wallace, Tim Maudlin, Barry Loewer, and David Albert for useful discussion, and in particular for pressing me with regards to points on which we don't see eye-to-eye. This work was supported by a grant from the Social Science and Humanities Research Council of Canada.

References

Albert, D. (2000). *Time and Chance*. Cambridge: Harvard University Press.

Engel, E. M. (1992). *A Road to Randomess in Physical Systems*. Berlin: Springer-Verlag.

Gibbs, J. W. (1902). *Elementary Principles in Statistical Mechanics: Developed with Especial Reference to the Rational Foundation of Thermodynamics*. New York: Charles Scribner's Sons.

Goldstein, S., J. L. Lebowitz, C. Mastrodonati, R. Tumulka, and N. Zanghì (2010). Approach to thermal equilibrium of macroscopic quantum systems. *Physical Review E 81*, 011109.

Goldstein, S., J. L. Lebowitz, R. Tumulka, and N. Zanghì (2006). Canonical typicality. *Physical Review Letters 96*, 050403.

Hacking, I. (1971). Equipossibility theories of probability. *British Journal for the Philosophy of Science 22*, 339–355.

Hacking, I. (1975). *The Emergence of Probability*. Cambridge: Cambridge University Press.

Hájek, A. (1997). 'Mises redux'—redux: Fifteen arguments against finite frequentism. *Erkenntnis 45*, 209–227.

Hájek, A. (2009). Fifteen arguments against hypothetical frequentism. *Erkenntnis 70*, 211–235.

Hopf, E. (1934). On causality, statistics, and probability. *Journal of Mathematics and Physics 13*, 51–102.

Hopf, E. (1936). Über die Bedeutung der willkürlichen Funktionen für die Wahrscheinlichkeitstheorie. *Jahresbericht des Deutschen Mathematiker-Vereinigung 46*, 179–195.

Jackson, E. A. (1968). *Equilibrium Statistical Mechanics*. New York: Dover Publication, Inc.

Jeffrey, R. (1992). Mises redux. In *Probability and the Art of Judgment*, pp. 192–202. Cambridge: Cambridge University Press.

Khinchin, A. I. (1949). *Mathematical Foundations of Statistical Mechanics*. New York: Dover Publications.

Laplace, P.-S. (1814). *Essai Philosophique sur les Probabilités*. Paris: Courcier. English translation in Laplace (1951).

Laplace, P.-S. ([1902] 1951). *A Philosophical Essay on Probabilities.* New York: Dover Publications. Translation of Laplace (1814).

Lewis, D. (1980). A subjectivist's guide to objective chance. In R. C. Jeffrey (Ed.), *Studies in Inductive Logic and Probability*, Volume II, pp. 263–93. University of Caifornia Press.

Linden, N., S. Popescu, A. J. Short, and A. Winter (2009). Quantum mechanical evolution towards thermal equilibrium. *Physical Review E 79*, 061103.

Lloyd, S. (2006). Excuse our ignorance. *Nature Physics 2*, 727–728.

Loewer, B. (2001). Determinism and chance. *Studies in History and Philosophy of Modern Physics 32*, 609–620.

Maxwell, J. C. (1871). *Theory of Heat.* London: Longmans, Green, and Co.

Maxwell, J. C. (1878a). Diffusion. In *Encyclopedia Britannica* (Ninth ed.), Volume 7, pp. 214–221. Reprinted in Niven (1965, pp.625–646).

Maxwell, J. C. (1878b). Tait's "Thermodynamics". *Nature 17*, 257–259, 278–280. Reprinted in Niven (1965, pp. 660-671).

Myrvold, W. C. (2011). Statistical mechanics and thermodynamics: A Maxwellian view. *Studies in History and Philosophy of Modern Physics 42*, 237–243.

Niven, W. D. (Ed.) (1965). *The Scientific Papers of James Clerk Maxwell*, Volume Two. New York: Dover Publications. Reprint of Cambridge University Press edition of 1890.

Penrose, O. (2005). *Foundations of Staistical Mechanics: A Deductive Treatment.* New York. Reprint of Pergamon Press edition, 1970.

Plastino, A. and A. Daffertshofer (2008). Classical typicality of the canonical distribution. *Europhysics Letters 84*, 30006.

Poincaré, H. (1912). *Calcul des probabilités* (2nd ed.). Paris: Gauthier-Villars.

Popescu, S., A. J. Short, and A. Winter (2006). Entanglement and the foundations of statistical mechanics. *Nature Physics 2*, 754–758.

Savage, L. J. (1973). Probability in science: A personalistic account. In P. Suppes (Ed.), *Logic Methodology, and Philosophy of Science IV*, pp. 417–428. Amsterdam: North-Holland.

Schaffer, J. (2007). Deterministic chance? *The British Journal for the Philosophy of Science 58*, 113–140.

Shimony, A. (1971). Scientific inference. In R. Colodny (Ed.), *The Nature and Function of Scientific Theories*, pp. 79–172. Pittsburgh: Pittsburgh University Press. Reprinted in Shimony (1993).

Shimony, A. (1993). *Seach for a Naturalistic World View, Volume I: Scientific Method and Epistemology.* Cambridge University Press.

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

Thompson, C. J. (1988). *Classical Equilibrium Statistical Mechanics.* Oxford: Clarendon Press.

Tolman, R. C. (1938). *The Principles of Statistical Mechanics.* Oxford: Clarendon Press. Reprint, Dover Publications, 1979.

van Fraassen, B. C. (1989). Indifference: the symmetries of probability. In *Laws and Symmetry*, pp. 293–317. Oxford: Oxford University Press.

von Kries, J. (1886). *Die Principien Der Wahrscheinlichkeitsrechnung: Eine Logische Untersuchung.* Frieburg: Mohr.

von Plato, J. (1983). The method of arbitrary functions. *The British Journal for the Philosophy of Science 34*, 37–47.

Wallace, D. (2011). The logic of the past hypothesis. Available at `http://philsci-archive.pitt.edu/8894/`.