The Necessity of Gibbsian Statistical Mechanics

David Wallace

November 10, 2018

Abstract

In discussions of the foundations of statistical mechanics, it is widely held that (a) the Gibbsian and Boltzmannian approaches are incompatible but empirically equivalent; (b) the Gibbsian approach may be calculationally preferable but only the Boltzmannian approach is conceptually satisfactory. I argue against both assumptions. Gibbsian statistical mechanics is applicable to a wide variety of problems and systems, such as the calculation of transport coefficients and the statistical mechanics and thermodynamics of mesoscopic systems, in which the Boltzmannian approach is inapplicable. And the supposed conceptual problems with the Gibbsian approach are either misconceived, or apply only to certain versions of the Gibbsian approach, or apply with equal force to both approaches. I conclude that Boltzmannian statistical mechanics is best seen as a special case of, and not an alternatve to, Gibbsian statistical mechanics.

1 Introduction

Unlike general relativity, or non-relativistic quantum mechanics, statistical mechanics lacks a generally-accepted axiomatic foundation. Not coincidentally, also unlike those other theories, the philosophical study ofstatistical mechanics has little if any consensus as to what the main foundational *problems* are, let alone how they are to be solved.

In lieu of such a consensus, much (non-historical) study of the foundations of statistical mechanics has focussed instead on the supposed contrast between two strategies for understanding statistical mechanics: the so-called "Gibbsian" and "Boltzmannian" programs.[1] The former characterises thermodynamic systems, and thermodynamic entropy, in terms of probability distributions over microstates of a system; the latter, in terms of individual systems. (The relation of the two approaches to anything actually said by the historical Gibbs and Boltzmann is a debatable point but will not be explored here; see Uffink (2007), Myrvold (2019), and references therein, for historically sophisticated discussions.)

Furthermore, in conceptual discussions of statistical mechanics (particularly, though not exclusively, in philosophy) something close to a consensus has

---

[1] The dichotomy is explicit in, e. g. , Frigg (2007), Callender (1999, 2001), and Albert (2000).

emerged: while Gibbsian statistical mechanics, and the Gibbsian definition of entropy (it is conceded) is a standard tool in *practical* applications of statistical mechanics, it is conceptually fatally flawed, and — unlike Boltzmannian statistical mechanics — lacks theresources to explain the key features of SM, notably the approach to equilibrium and the laws of thermodynamics. In brief summary (to be returned to later), Gibbsian statistical mechanics is criticised for trying to explain thermodynamic behaviour as a feature of our information about the world rather than as a feature of the world, for failing to identify entropy as a property of individual physical systems, and for evading rather than answering the challenges to statistical mechanics posed by time-reversibility of microdynamics and by the Poincaré recurrence theorem. (For examples of this criticism, see Albert (2000), Callender (1999, 2001), Goldstein (2001), Lebowitz (2007), and Maudlin (1995).) As such, despite the acknowledged technical advantages of the Gibbsian approach, the Boltzmannian approach offers the true *explanation* of the successes of statistical mechanics, and in particular of thermodynamics.[2]

My thesis here is threefold. Firstly, the Gibbsian and Boltzmannian approaches are not rival approaches to the statistical mechanics and thermodynamics of the same systems: rather, the Gibbsian framework is a more general framework in which the Boltzmannian approach may be understood as a special case. Secondly, the wider applicability of the Gibbsian approach is indispensible from a naturalistic perspective, inasmuch as a wide range of empirically successful applications of statistical mechanics cannot be understood within the Boltzmannian approach. And thirdly, the Boltzmannian criticisms of the Gibbsian approach largely miss their mark: in the main, they apply to some of the justifications, motivations and interpretations offered for Gibbsian statistical mechanics but not to Gibbsian statistical mechanics itself.

The paper is structured as follows. In section 2 I briefly summarise the main commitments of the two approaches. I then consider first statistical mechanics (sections 3-5) and then thermodynamics (sections 6-7), arguing in each case that the two approaches coincide when it comes to thermodynamic properties of sufficiently well-behaved systems (sections 3, 6) but that there are theoretically and experimentally relevant regimes where only the Gibbsian approach is applicable (sections 4-5, 7). In sections 8-10 I consider the main objections advanced against the Gibbsian approach. Section 11 is the conclusion.

I work throughout under the (of course false) assumption that the underlying microphysics is classical mechanics, so that *micro*states of a system are represented by points in a classical phase space. I do so to make contact with the contemporary literature rather than out of a belief that quantum theory is irrelevant here; at the end (section 12) I make some brief comments on how quantum theory affects my thesis.

---

[2]This marks a rather common pattern in philosophy of physics, where philosophers predominantly study and espouse a minority position in physics which is technically less productive, but conceptually clearer, than the alternative — other examples include algebraic vs. mainstream approaches to QFT (cf Wallace (2011), loop-space quantum gravity vs. string theory, and modificatory (e. g. , Bohm or GRW) rather than pure-interpretion approaches to the quantum measurement problem (cf Wallace (2018).

2   Gibbs vs Boltzmann: an overview

The main features of Boltzmannian statistical mechanics[3] are:

- Phase space is divided into *macrostates*, representing (something like) "macroscopically indistinguishable states". More precisely, each energy hypersurface in phase space is divided into macrostates.

- Systems are represented by phase-space points: that is, the state of a system is a (classical) microstate. The macrostate of a system is just that macrostate in which the system's microstate lies; due to the way that the macrostate partition is constructed, to know the macroscopic properties of a system we need know only its macrostate.

- The *Boltzmann entropy* of a system (up to a scale factor k_B) is the logarithm of the phase-space volume of the macrostate in which it lies (the "Boltzmann entropy" of the system).

- A system is in *Boltzmann equilibrium* if it lies in the largest of the macrostates (called the *equilibrium macrostate*) given the system's energy. The geometry of phase space for macroscopically large systems means that it is certain (for dilute gases) and heuristically plausible (in general) that the equilibrium macrostate for energy E is overwhelmingly larger in phase-space volume than all other macrostates of energy E. I'll call a system with this property *Boltzmann-apt'*.

- The approach to Boltzmann equilibrium is essentially a consequence of phase-space geometry combined with some reasonable assumptions about the dynamics: since almost all points (by phase-space volume) of a given energy lie in the equilibrium macrostate, either the dynamics or the initial state would have to be "ridiculously special" (Goldstein 2001) for the system not to approach equilibrium. Similarly, a system in the equilibrium macrostate is exceptionally unlikely to wander out of the macrostate, given that virtually all states it could evolve into are also in the equilibrium macrostate. This conception of equilibrium makes the approach to equilibrium a statistical or probabilistic matter, and (given the Poincaré recurrence theorem[4]) systems initially away from equilibrium will eventually evolve away from equilibium again, but for Boltzmann-apt systems the probability of equilibration is so close to unity, and the time taken for recurrence so large, that this can normally be disregarded. The *Boltzmann equilibration time* is the typical timescale after which a system reaches Boltzmann equilibrium.

---

[3]My account largely follows Albert (2000) and Frigg (2007); I take it to be mostly in agreeement with, e. g. , Carroll (2010), Goldstein (2001), Lebowitz (2007), North (2002), Penrose (1994).

[4]See Wallace (2015b) for a review, and for emphasis of the fact that the so-called 'problem of measure zero', according to which only 'almost all' states recur, is an artifact of classical mechanics without foundational significance.

- The *Boltzmann equilibrium values* of a system are just the macroscopic quantity values which specify the equilibrium macrostate. They may be measured by single measurements of the system after the equilibration timescale.

Note that at this stage, the terms 'equilibrium' and 'entropy are being introduced as terms of art, without any prior assumption about their relationship to similarly-named thermodynamic terms. (I consider the relation between statistical-mechanical and thermodynamic concepts of equilibrium and entropy in sections 6-7.)

At first pass, the main features of Gibbsian SM[5] are:

- Thermodynamic systems are represented by probability distributions over phase space (mathematically, by positive measures on phase space assigning measure 1 to the whole space).

- The *Gibbs entropy* of a system represented by probability measure ρ is

$$S_G(\rho) = -k_B \int \rho \ln \rho, \qquad (1)$$

the "Gibbs Entropy".

- A system is at *Gibbs equilibrium* if ρ is time-invariant under the system's dynamics, which (again at first pass) seems fair enough: "equilibrium", after all, is understood from a macroscopic perspective as a state that is unchanging in time). It then follows from ergodic theory that, if the system is ergodic, the equilibrium distribution must be uniform on each energy hypersurface (see Malament and Zabell (1980) for more discussion of this relation between Gibbsian SM and ergodic theory). The *Gibbs equilibration time* is the typical timescale after which a system reaches Gibbs equilibrium.

  For simplicity, we will assume that this equilibrium probability function is the *microcanonical distribution*, which is restricted to a single energy hypersurface and which is uniform on that hypersurface. (Extending to the more general case does not essentially change the story, but introduces distracting technical complications.)

- The *Gibbs equilibrium values* of the system are the expected values of the various dynamical quantities (microscopic or macroscopic) evaluated with respect to the equilibrium distribution. They may be measured in the usual statistical-mechanical way, by measuring many copies of the equilibrated system and taking an average.

There is an immediate problem with this first-pass version of the Gibbsian approach: it seems to have the corollary that real systems do not increase in

---

entropy or approach equilibrium. For if the probability measure is defined over individual microstates evolving under the system's Hamiltonian, the dynamics of the probability measure itself are determined: if the dynamics deterministically carry the points in region V at time 0 onto the points in region $V(t)$ at time t, then the conditional probability for a microstate of the system to be in region $V(t)$ at time t given that it was in region V at time 0 is unity. This suffices to uniquely determine the dynamics for ρ: they are given by Liouville's equation,

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho(t) = \{H, \rho(t)\}, \tag{2}$$

where H is the Hamiltonian and $\{\cdot, \cdot\}$ is the Poisson bracket. This is a first-order equation, so if the first derivative of ρ is zero at some given time it must have been zero at all previous times: that is, no system that starts away from equilibrium can reach equilibrium. Nor is there any prospect that the system even gets *closer* to equilibrium, at least as measured by the Gibbs entropy, since it is an easy consequence of Liouville's equation that the Gibbs entropy is invariant over time.

One influential way to evade this problem — due to Ed Jaynes — is to drop the idea that dynamics have anything much at all to do with equilibrium or the approach to equilibrium. To Jaynes and his followers,[6] to say that a system is at equilibrium is to say simply that we know nothing about its state except its energy; to say that a system is evolving towards equilibrium is to say that we are losing information about its state. This strategy has been robustly criticised in the philosophy literature,[7] most notably by advocates of the Boltzmannian program, as simply failing to recognise the obvious fact that the approach to equilbrium is an observed, empirically confirmed physical process. (See, e. g., Albert (2000).)

A more sophisticated variant of Jaynes' approach, due to Wayne Myrvold (2014), treats the probability distribution as epistemic, takes seriously its dynamical evolution, but asserts (i) that realistic agents will lose the ability to keep track of very fine-grained features of the probability distribution, and (ii) those fine-grained features are not dynamically relevant when it comes to calculating macroscopic quantities. As such, an agent can and must keep track not of the true probability distribution but of a coarse-grained variant of it.

A quite different approach to the problem is to keep the idea that the system evolves under some dynamics, but drop the idea that those dynamics are Hamiltonian. Most radically, we could assume that explicit time asymmetry must be introduced into classical physics to understand the approach to equilibrium (the work of Ilya Prigogine is generally interpreted as exploring this option; see Bishop (2004) and references therein for philosophical discussion). Alternatively, we might appeal to the fact that no physical system short of the

---

[6]See, e. g., Jaynes (1965, 1957a, 1957b).

[7]Though for acerbic pith in response to Jaynes, little beats physicist J. S. Rowlinson's quotation of Leslie Ellis: "Mere ignorance is no ground for any inference whatever. Ex nihilo nihil." (Rowlinson 1970)

whole Universe is truly dynamically isolated, and so will not truly have dynamics given by a Hamiltonian for the system alone. But the first approach is in danger of looking ad hoc, and the second does not seem to have the resources to explain time asymmetry.

The main strategy to reconcile the Gibbsian approach with the Hamiltonian dynamics involves a small modification to the former, rather than the latter:

- We introduce a *coarse-graining* map J on the space of probability distributions. At least in foundational discussions, the most common form given for J works like this:

    1. Coarse-grain phase space into cells.
    2. Replace each probability measure ρ with a measure that is uniform over each cell and which assigns the same probability to a cell as ρ did.

    This definition ensures that J is a *projection operator*: that is, $J^2 = J$. Maps like these (when linear) are sometimes called *Zwanzig projections*, and in fact can be defined by methods of coarse-graining much broader than the phase-space cell method; see, e.g., Zwanzig (2001) (or, for foundational discussion, Zeh (2007) or Wallace (2015a)) for more details. The action of J is often glossed as representing the fact that our measurements of a system have finite precision, although (as I discuss further in section 9 this is not really a viable way to think about it).

- The entropy of a probability distribution ρ (with respect to a given coarse-graining J) is now defined as the Gibbs entropy of the *coarse-graining* of ρ:

$$S_{G;J}(\rho) = S_G(J\rho) = -k_B \int (J\rho) \ln(J\rho). \tag{3}$$

- A system represented by probability function ρ is at equilibrium (with respect to J) if its coarse graining $J\rho$ is invariant in time.

There is no dynamical principle according to which the coarse-grained entropy is a constant of the motion; indeed, it is mathematically possible for the coarse-grained entropy to increase to a maximum value and then remain there indefinitely. Similarly, (and relatedly) nothing prevents a system initially away from equilibrium evolving into equilibrium, if "equilibrium" is defined in coarse-grained terms. (Note that this approach is at least closely related to Myrvold's proposal.)

Whenever it is significant, I will assume this last understanding of Gibbsian statistical mechanics — but for most of my purposes in this paper, all that matters is that there is an *effective probability distribution* for the system at any given time, which coincides with the actual probability distribution for any macroscopically relevant quantity and which evolves towards Gibbs equilibrium. Whether that distribution *is* the true probability distribution, or is simply a coarse-grained approximation to it, will be neither here nor there.

3 Equilibration in Boltzmann-apt systems

To begin our assessment of Boltzmannian and Gibbsian statistical mechanics, let us start by considering systems that are Boltzmann-apt: that is, each of which has a decomposition into macrostates such that, for any energy E, the equilibrium macrostate at that energy occupies the overwhelming majority of the volume of the energy hypersurface. In this context, the Gibbsian framework actually entails the Boltzmannian framework. For consider: if the Gibbs framework holds and if the system's initial energy is E, then after the Gibbs equilibration timescale the effective probability distribution is uniform on the energy-E hypersurface. But since the system is Boltzmann-apt, that means that with overwhelmingly high probability the system is in the Boltzmann equilibrium macrostate: that is, with overwhelmingly high probability the system is at *Boltzmann* equilibrium.

After the equilibration timescale, the macroscopic degrees of freedom at any time will be overwhelmingly likely to possess their Boltzmann-equilibrium values. So the Gibbs and Boltzmann equilibrium values coincide to an extremely high degree of accuracy. There will be a small probability that the result of a given measurement will diverge from these values, corresponding to the small probability that the microcanonical distribution assigns to the non-equilibrum macrostates. These *fluctuations* around the Boltzmann equilibrium values can be described either are fluctuations *within* Gibbs equilibrium, or as fluctuations *into and out of* Boltzmann equilibrium, but this is simply a semantic difference, and does correspond to any physical disagreement about the system.

So the user of Gibbsian statistical mechanics is fully entitled to use — and indeed, to accept the truth — of Boltzmannian statistical mechanics. In those systems to which the latter is applicable, the Gibbsian framework can be seen as grounding the Boltzmannian one.

The situation is not symmetric, for obvious conceptual reasons. The Boltzmannian framework *per se* contains no explicit notion of probability, and so does not permit us even to define the Gibbsian probability distribution. (Almost certainly the Boltzmannian framework requires some qualitative notion of probability — perhaps the notion of *typicality* advocated by Goldstein *et al* (2006) (see Frigg (2009) and references therein for further details) — but its advocates commonly claim that the full quantitative probability distribution is not required, and indeed is an incoherent concept in classical mechanics.) Nonetheless the Boltzmannian has a straightforward story to tell about the practical utility of Gibbsian methods: precisely because the equilibrium macrostate so dominates the volume of the energy hypersurface, and because macroscopic variables have almost-constant values on each macrostate, the average value of a macroscopic variable will be extremely close to its actual value at Boltzmann equilibrium. So the Boltzmannian can harmlessly use Gibbsian probabilistic methods as a calculational tool, without any commitment to their truth.

In summary: as long as we (a) confine our attention to Boltzmann-apt systems, and (b) wish to calculate only equilibrium values of macroscopic variables, the Gibbs and Boltzmann approaches are equivalent for all practical purposes.

And since the Gibbsian approach introduces a problematic notion of quantitative probability, and arguably has a murkier account of how equilibration works, one can see the case for preferring the Boltzmannian approach, and for setting aside as foundationally irrelevant the incontestable fact that working physicists use the Gibbsian approach.[8]

As we shall see shortly, however, (a) and (b) by no means exhaust the range of applications of classical statistical mechanics.

4 Beyond Boltzmann equilibrium: transport properties

What might we want to calculate of a (let's say Boltzmann-apt) system at equilibrium, other than the values of its macroproperties? One important class of properties are the *multi-time correlation functions*. For instance, consider a classical dilute gas at equilibrium, and let $X(t)$ be the position of some arbitrarily chosen particle at time t. At equilibrium, the expected value $\langle X(t) \rangle$ of that position is time-independent and equal to the center of mass of the gas; note that I can make this statement based on only the vaguest and most qualitative information about the gas.

But now consider the two-time correlation function

$$C(t) = \langle X(t)X(0) \rangle. \tag{4}$$

The form of this function *cannot* be worked out in any simple way from the equilibrium macrofeatures of the gas. Calculationally, it's clear that it involves the dynamics, since to evaluate it I need to work out how likely the system is to transition from one position to another over time t; conceptually, we have

$$\langle (X(t) - X(0))^2 \rangle = \langle X(t)^2 \rangle + \langle X(0)^2 \rangle - 2\langle X(t)X(0) \rangle = 2(\langle X(0)^2 \rangle - C(t)) \tag{5}$$

so that $C(t)$ may be recognised as a measure of how far a randomly-chosen particle diffuses through the gas after time t.

Knowing that the system is in the Boltzmannian equilibrium macrostate does not in any straightforward way provide us with enough information to calculate $C(t)$; indeed, since $C(t)$ is an explicitly probabilistic quantity, it is not even defined on the Boltzmannian approach. Of course this would not matter if $C(t)$ were simply a theoretical curiosity — but, as my discussion above may have suggested, two-time correlation functions like this are the main tool in calculating diffusion coefficients, rates of thermal conductivity, and the other quantitative properties of a system that characterise its behavior close to but not at equilibrium. Indeed, the huge subject of *transport theory* is largely concerned with calculating these two-time functions, and its methods are quite thoroughly

---

[8]See, e.g., Callender (1999, p.349): "[L]et me happily concede that for the practice of science, Gibbsian SM is usually to be preferred. Since the values of all the entropy functions I discuss agree at equilibrium, my arguments are necessarily philosophical in nature."

confirmed empirically. (See, e. g. , Zwanzig (2001), Altland and Simons (2010), and references therein.)

So: even for Boltzmann-apt systems, there are important cases where probabilistic methods seem necessary and do not reduce to Boltzmannian methods in any simple way. (Multi-time correlation functions by no means exhaust the list of such cases — indeed, I discuss another, the modern ('BBGKY') derivation of Boltzmann-type equations through truncation of the N-particle probability distribution to 2-particle marginals, in Wallace (2016) — but they already suffice to make the point.)

5   Beyond Boltzmann equilibrium: fluctuations

What about systems that are *not* Boltzmann-apt? One might *a priori* guess that statistical mechanics is inapplicable to such systems, since after all it relies for its efficacy (doesn't it?) on being able to average over a very large number of constituents. But this is not the case: as I now illustrate, statistical mechanics is frequently and successfully applied to systems where the Boltzmannian equilibrium state does not overwhelmingly dominate the energy hypersurface. Here I give two such examples, in neither of which we can simply assume that a system — even after the equilibration timescale — is "overwhelmingly likely" to be in one particular macrostate. (I give others in section 7 when I discuss the thermodynamics of small systems.)

The first example is *spontaneous symmetry breaking*, say (for definiteness) in ferrmomagnets. Here, "the" equilibrium macrostate below a certain temperature has a non-zero expectation value of magnetic spin, corresponding to the fact that it is energetically favourable for adjacent spins to line up. But it then follows from the rotational symmetry of the underlying dynamics that there must be another macrostate obtained by applying a rotation to each microstate in the first, of equal volume to 'the' equilibrium macrostate and equally justifiably called an equilibrium macrostate. In other words: systems with spontaneous symmetry breaking are not Boltzmann-apt, at least at energies corresponding to temperatures below the symmetry-breaking temperature.

It's possible to imagine a modification of the Boltzmannian framework to handle this case. We could generalise the definition of Boltzmann-aptness to allow for *many* equilibrium states, each related by a symmetry, and such that their *collective* volume dominates the energy hypersurface. (And notice that since Boltzmann entropy is logarithmic, replacing one equilibrium state with N reduces the entropy only by $k_B \ln N$, which will be a negligible shift if N is much smaller than the number of microscopic degrees of freedom). We could argue that unless the dynamics are 'ridiculously special', the system is overwhelmingly likely to end up in *one* of the equilibrium macrostates, and to remain there for a very long time.

But this does not suffice to save the phenomena. We require not just that the system will end up in one such state, but that *each is equally likely*. This follows directly from the assumption of Gibbs equilibrium (each has equal volume, so

each is equally likely) and is well-confirmed empirically. For instance, the pattern of symmetry breaking can be analysed in a ferromagnet, and it is clearly distributed at random. This assumption may seem obvious but it's not at all clear how Boltzmannian statistical mechanics can reproduce it, without being supplemented by an explicit probability distribution — we cannot, for instance, say 'typical states are equally likely to end up in each equilibrium macrostate', since 'being equally likely to end up in each equilibrium macrostate) is not a property that any given microstate can have in a deterministic theory.

The second example is *Brownian motion*. Here the system consists of one large particle in a bath of smaller ones, with the latter usually taken to be at Gibbs equilibrium. The large particle has no meaningful notion of 'macrostate' and evolves randomly, due to fluctuations in the number of particles colliding with it from any given direction; its probability distribution will converge on the Maxwell-Boltzmann distribution, but this in no way means that the actual state of the particle is time-invariant after its 'equilibration' timescale. Applying Gibbsian methods (such as by calculating the two-time correlation functions) we can derive the stochastic equation which the particle obeys (see Zwanzig (2001) for a more detailed discussion). The resultant equation, and variants on it which apply to similar setups, has been widely applied and thoroughly confirmed empirically.

Note that in both of these examples, probability is not simply playing a foundational role (as was the case in, say, the calculation of transport coefficients from two-point functions). Rather, the predictions of the theory are themselves expressed probabilistically, and don't have any direct re-expression in terms of categorical properties. Of course, the probabilities are measured through relative frequencies, and it is always open to the Boltzmannian to insist that apparently "probabilistic" predictions should be reinterpreted as, say, claims about what is typical when an experiment is repeatedly performed on a very large number of copies of the system. But this is just a claim about the general foundations of probability in statistical mechanics (specifically, that it should be understood on frequentist lines). It in no way eliminates probability from the actual statement and use of statistical mechanics.

To summarise sections 3-5: while there is a class of statistical-mechanical systems, and a class of properties of those systems, such that Gibbsian and Boltzmannian methods are equally applicable when calculating those properties, the scope of statistical mechanics is much wider than those classes and includes many phenomena that seem treatable only by Gibbsian means. As we will now see, essentially the same story recurs when we turn to thermodynamics.

6   Thermodynamics in macroscopic systems

As I (Wallace 2014) and others (see Skrzypczyk, Short, and Popescu (2014) and references therein) have argued elsewhere, thermodynamics is not a dynamical theory in the usual sense: not, that is, a theory of how undisturbed physical systems evolve over time. (Indeed, insofar is it is seen as such, thermodynamics

is essentially trivial: it is concerned with equilibrium systems, and the defining feature of such systems is that they do not evolve at all over time.) Rather, thermodynamics is a *control theory* (or, alternatively, a resource theory), concerned with which transformations can or cannot be performed on a system by an external agent, given certain constraints on that agent's actions. From this perspective, the First Law of thermodynamics disallows control actions where the work done on a system (i.e. the energy cost of the control operation to the agent), plus the heat flow into the system from other systems, does not equal the change in internal energy of the system. The Second Law, meanwhile, disallows those transformations which lead to a net decrease in thermodynamic entropy.

As a more precise (though by no means completely precise) statement of the content of thermodynamics, I offer the following: there are equilbrium systems, and they can be completely characterised for thermodynamic purposes by a small number of *thermodynamic parameters*: the energy U and some externally-set parameters — in typical examples, the volume V. (It is also possible to add some conserved quantities, such as particle number; for expository simplicity I omit this complication). The *equation of state* determines the thermodynamic entropy as a function of those parameters; schematically, we might write $S = S(U, V)$, understanding V to stand in for whatever are the actual external parameters and conserved quantities. Other thermodynamically relevant quantities can be calculated from the equation of state: therrmodynamic temperature, for instance, is the rate of change of U with S at constant V; pressure is minus the rate of change of U with V at constant S. The Second Law is the requirement that any allowable control process leave the sum of all entropies non-decreasing. (For a somewhat more detailed sketch of thermodynamics on these lines, see Wallace (2017, section 2).)

Recovering thermodynamics from an underlying mechanical theory, then, requires us to provide mechanical definitions of (inter alia) 'equilibrium', 'allowable control process', the 'thermodynamic parameters', and 'entropy' along with other reasonable mechanical posits, such that (a) these laws of thermodynamics can be derived (at least approximately) from those definitions and posits, (b) the definitions do reasonable justice to the informal, operational understanding of the thermodynamic terms (according to which, for instance, equilibrium states have macroproperties that are approximately constant in time) and (c) given that the equation of state of a system is an empirically measurable feature of that system, the definitions allow us to recover the actual, quantitative form of the equations of state of known systems.

In looking for such a recovery, we are not operating in a vacuum. After a century of statistical mechanics, it is well known how to calculate the thermodynamic entropy of a (large) system: taking the Hamiltonian of that system to be parametrised by the external parameters (so that, for instance, the Hamiltonian of a box of gas is a function of its overall volume), treat the thermodynamic energy as just directly representing the mechanical energy, and define the entropy as the log of the phase-space volume of all states with that energy (this makes the entropy indirectly a function of V as well as energy U, since the Hamiltonian depends on V). Mathematically speaking, this is the Gibbs entropy of the mi-

crocanonical distribution; empirically, this *works*, for a huge variety of systems, and so recovering it (at least to a high degree of approximation) is a *sine qua non* of any mechanical recovery of thermodynamics. (Call this the 'quantitative test' of a proposed recovery.)

We can also say something general about allowable control operations. At a minimum, such control operations ought to correspond to transformations allowable by the basic structure of classical mechanics. Any such transformations preserve phase-space volume when acting on closed systems, so we will assume that any such control operation is indeed volume-preserving. (It is possible to decrease phase-space volume by measuring the state and choosing the control operation accordingly, a strategy that leads towards Maxwell-Demon-style (apparent) counter-examples to the Second Law, but these lie beyond the scope of this paper; for discussion in the control-theory context, see Wallace (2014) and references therein.)

With all this said, let's consider what a derivation of thermodynamics for macroscopic systems might look like. In most cases (perhaps putting aside spontaneous symmetry breaking), we can reasonably assume such systems are Boltzmann-apt: that is, that for given U, V the phase space region is dominated by a single region in which the macroscopic variables are approximately constant. Let's start with the Boltzmannian approach, in which equilibrium is defined as occupation by the system's actual microstate of the largest-volume macrostate, and in which we eschew explicit and quantitative use of probabilities. A natural choice for control operations is then

Boltzmann-equilibration operations: in which the system's thermodynamic parameters are changed, some subsystems are brought into and out of thermal contact (that is, coupled or decoupled by some Hamiltonian) and then the system is allowed to evolve such that (in an irreducibly imprecise way) it is *almost certain* to reach equilibrium. To restrict to equiibration operations is to assume an agent who has control only over a system's bulk thermodynamic parameters.

By assumption, a Boltzmannian-equilibrium operation must have the desired effect for all (or at least the vast majority in phase-space measure) of points in the equilibrium region. So it follows that the volume of the post-operation equilibrium state must exceed that of the pre-operation equilibrium state. Hence Boltzmann entropy is non-decreasing under these operations, in accordance with the Second Law. Furthermore, since the equililbrium macrostate dominates the allowable region of the phase space, its Boltzmann entropy is numerically almost equal to the microcanonical entropy, so that the quantitative test is passed too. In short, this seems an entirely satisfactory (sketch of a) derivation of thermodynamics from Boltzmannian statistical mechancs, with the Boltzmannian notions of 'entropy' and 'equilibrium' mapping to the thermodynamic ones.

What about from the Gibbsian perspective? Here, to say that the system is at equilibrium is to say that its effective probability is uniform over the allowable region of phase space. Given that the system is Boltzmann-apt, recall that the system thus *almost certainly* has the Boltzmann-equilibrium values of

the macroproperties. (Again, there is a purely semantic difference here: the Boltzmannian says that the system is almost certainly at equilibrium and, if at equilibrium, certainly has such-and-such values of the macroproperties; the Gibbsian says that the system *is* at equilibrium and, as such, *almost* certainly has those values.) The natural choice here for control operations is:

Gibbs-equilibration operations: In which the system's thermodynamic parameters are changed, some subsystems are brought into and out of thermal contact (that is, coupled or decoupled by some Hamiltonian) and then the system is allowed to evolve such that if the effective probability distribution is originally uniform over the region determined by the old parameters, after the control operation it is uniform over the region determined by the new parameters.

Given the background coarse-graining assumptions of the Gibbs approach, the Gibbs entropy of the effective probability distribution cannot go down in the control operation. And the quantitative test is trivially passed. So again, we have a (sketch of a) satisfactory derivation of thermodynamics from Gibbsian statistical mechanics, with 'equilibrium' and 'entropy' here played by the Gibbsian rather than the Boltzmannian notions. As a further difference, because the Gibbsian approach describes the system probabilistically, the Gibbsian correlates for thermodynamic energy and work are *expected* values, not categorical values — although given the assumption of Boltzmann-aptness, the actual value will be extremely close to the expected value with extremely high probability.

The situation is parallel to the statistical-mechanical case. For the Gibbsian, there is no factive difference between the two approaches: the validity of the Gibbsian approach entails that of the Boltzmannian approach, and the two strategies differ only semantically. For the Boltzmannian, the Gibbsian use of probabilities is justifiable only on pragmatic terms.

Again in parallel to the statistical-mechanical case, Gibbsian thermodynamics would be required only if there are applications of thermodynamics where the Boltzmann-apt assumption fails, where the use of actual probabilities is unavoidable, and where statistical fluctations are non-negligible and measurable. And again, there are indeed many such cases.

7   Beyond macroscopic thermodynamics

A combination of theoretical and experimental advances have made the last twenty years a golden age for the statistical mechanics of small systems. On the theoretical side, the key advance has been *fluctuation theorems*, results derived in the Gibbsian framework which relate the probability distributions over different transformations between systems. For instance, in macroscopic thermodynamics, if a system is transformed between equilibrium states while remaining all the while in contact with a heat bath at temperature T, it is a standard result that

$$\Delta F \geq W \tag{6}$$

where ΔF is the decrease in the free energy between initial and final states, and W is the work extracted in moving from initial to final states. The inequality becomes an equality only in the quasi-static limit. The *Jarzynski equality* (Jarzynski 1997) sharpens this to:

$$e^{-\Delta F/kT} = \left\langle e^{-W/kT} \right\rangle \tag{7}$$

where the right hand side is an expectation value over different microphysical realisations of the transition between initial and final state. For large systems and slow changes, the fluctuations in the right-hand-side will be negligible and we will recover the quasi-static, non-probabilistic result, but the equality holds — according to Gibbsian statistical mechanics — even for small systems and for rapid transformations of those systems.

The (closely related) *Crooks fluctuation theorem* (Crooks 1998) again concerns transitions between equilibrium states A, B of a system in thermal contact with a reservoir. For that system, the theorem states that

$$\frac{\Pr(W|A \to B)}{\Pr(-W|B \to A)} = \exp\left(\frac{W - \Delta F}{kT}\right) \tag{8}$$

where $\Pr(W|A \to B)$ is the probability that a given transition from A to B will require work W, and $\Pr(-W|B \to A)$ is the probability that a given transition from B to A will extract work W. In the limit of large systems and slow processes, $W = \Delta F$ and the theorem just says that transitions in either direction are equally likely, but the result is again a mathematical consequence of Gibbsian statistical mechanics even for fast processes on microscopic systems.

These are results that cannot be derived in full generality in the Boltzmannian framework: they are explicitly probabilistic, and the free energies are themselves defined as expectation values. This would be of only limited significance, though, if they remained purely theoretical results, experimentally untestable on systems small enough to display meaningful fluctuations.

But of course they have been tested, extensively. The most well-developed examples have involved the stretching and unstretching of RNA and polymer chains (see, for instance, Collin *et al* (2005)). In these experiments, the work done on the chain in a given stretch-and-unstretch shows large thermal fluctuations (that is, we are way outside the Boltzmann-apt regime) but nonetheless the probability distributions over work done, and the expectation values, conform exactly to the predictions of the fluctuation theorems. Experiments have been done in a range of other small systems, and the field is moving too quickly to summarise in a foundational article like this one (for an already-dated review, see Bustamante, Liphardt, and Ritort (2005)) but suffice it to say that the experimental evidence for the fluctuation theorems looks compelling, even (especially?) in small systems where fluctuations are large.

Tests of the fluctuation theorems do not exhaust the range of recent experiments in microscopic thermodynamics. To give one more example, recall that Richard Feynman famously argued Feynman (1967, pp.116-9) that a ratchet

could not be used as a Maxwell demon to transfer heat between two reservoirs of equal temperature because of fluctuations. Bang *et al* (2018) have demonstrated this result empirically, using a microscopic 'ratchet' consisting of a colloidal particle in an optical trap. They verified that although on any given run of the experiment heat is sometimes converted into work, the expected work output (measured by averaging over many runs) is zero unless the two reservoirs are of unequal temperature.

In summary: modern physics is extensively applying, and testing, thermodynamics in the microscopic regime, where the Boltzmann-aptness assumption completely fails and predictions are explicitly probabilistic. In this regime, Boltzmannian statistical mechanics is inapplicable and the Boltzmannian conception of equilibrium is useless.

Now, it should be acknowledged that in using "thermodynamics" to apply to this regime, in which terms like 'work' and 'free energy' enter the Second Law only as expectation values, we are going well beyond their original use in nineteenth-century phenomenological thermodynamics. I could even concede that *that* subject reduces to Boltzmannian statistical mechanics as readily as to the Gibbsian version. But this is a semantic matter. Whether the subject that modern physicists call 'thermodynamics' is a precisification of that nineteenth-century subject or a genuine extension of it, it is a robust and empirically well-confirmed subject which relies for its formulation and its use on probabilistic — that is, Gibbsian — conceptions of statistical mechanics in general and equilibrium in particular.

I conclude that — even if we restrict attention to the classical case — the range of applications of modern statistical mechanics and thermodynamics vastly outstrips what can be analysed using just the methods of Boltzmannian statistical mechanics, without the explicit introduction and study of quantitative probabilities. Assuming (as I take it should be uncontroversial) that a foundation for statistical mechanics needs to be a foundation for *all* of statistical mechanics, and not just for the tiny fraction that had been developed by, say, 1900, then I don't see an alternative but to accept the Gibbsian framework as that foundation, with the Boltzmannian framework as a highly important special case of it applicable to Boltzmann-apt systems for the calculation of certain quantities.

I could end the paper here. However, critics of Gibbsian statistical mechanics have advanced a number of objections to the effect that it is conceptually incoherent, and those objections are not magically swept away simply by a naturalistic argument that we need Gibbsian methods (even if that need might make us more confident that the objections can somehow be met). In the next three sections, I consider what I take to be the main Boltzmannian objections to Gibbsian statistical mechanics, and argue that they either rest on misconceptions about the framework, or else apply to particular versions or developments of the Gibbsian framework but not to that framework in itself. For expository clarity, and to ensure that I am not engaging with straw men, I will concentrate on influential criticisms due to Albert (2000), Callender (1999, 2001) and Maudlin (1995).

8 Objections to Gibbs: modality and probability

One of the most common objections to Gibbsian statistical mechanics is that it makes thermodynamics in general, and entropy in particular, a study of modal, in particular probabilistic, features of a system, rather than of categorical features, when the latter is what is required to do justice to the phenomena. Indeed, this is often held up as a simple and straightforward *mistake*. As representative examples, consider Callender (2001, p.544):

> The problem is not the use of ensembles . . . The problem is instead thinking that one is *explaining* the thermal behaviour of *individual real systems* by appealing to the monotonic feature of some function, be it ensembles or not, that is not a function of the dynamical variables of individual real systems. It is impossible to calculate the intellectual cost this mistake has had on the foundations of statistical mechanics. The vast majority of projects in the field in the past century have sought to explain why my coffee (at room temperature) tends to equilibrium by proving that an ensemble has a property evincing monotonic behaviour. (Emphasis Callender's.)

Or Maudlin (1995, p.147):

> Since phenomenological thermodynamics originally was about . . . individual boxes [of gas], about their pressures and volumes and temperatures, 'saving' it by making it be about probability distributions over ensembles seems a Pyrrhic victory. It is remarkable, and not a little depressing, to see the amount of effort and ingenuity that has gone into finding something of which the phenomenological laws can be strictly true, while insuring that the something cannot possibly be the phenomena.

A particular concern of both authors is the Poincaré recurrence theorem. Given recurrence, we know that a system will eventually return arbitrarily close to its initial state, and so if entropy is indeed a function of the microstate of a system, then entropy must eventually return to its initial value, seeming to demonstrate that any account of entropy as monotonically increasing has lost touch with the microfoundations of thermodynamics.

There is a great deal to say in response to these objections. Here I identify six points that ought, jointly, to assuage such worries.

Firstly, what we want to explain in non-equilibrium statistical mechanics is itself something modal: not that systems *invariably* go to equilibrium but that they do so *almost certainly*. It might not be compulsory to quantify 'almost certainly' as 'with probability very close to 1' but at any rate it does not seem to involve a substantial change of focus. A probabilistic property of a system is poorly suited to explain why the system deterministically behaves in such-and-such a way, but it is well suited to explain why it very probably behaves in that way.

Secondly, in statistical mechanics (as distinct from thermodynamics) the entropy is ultimately no more than a book-keeping device to keep track of irreversibiliy in a system's dynamics. In Gibbsian statistical mechanics, irreversibility typically takes the form of an increasing dispersal of a probability distribution over the constant-energy hypersurface (which, in Boltzmann-apt systems, in turn entails the increasing likelihood of the system being in the equilibrium macrostate). Because (an appropriately coarse-grained) Gibbs entropy tracks this dispersal, it is a useful tool to study the approach to equilibrium, but it has no causal or explanatory role in its own right. (And the same is true for Boltzmannian statistical mechanics: the Boltzmann entropy, as far as non-equilibrium statistical mechanics is concerned, is a device for tracking a system's increasing likelihood of being found in increasingly-large macrostates.)

Thirdly, while in thermodynamics the entropy plays a much more quantitatively significant role, that role is itself modal. As I noted in section 6, the subject matter of thermodynamics is the transformations that an agent can bring about in a thermodynamic system. The word 'can' betrays the modality of this subject matter: if a system might be in one of many states, this constrains the transformations that the agent can bring about, at least without measuring that state. This modality is hidden in most applications of macroscopic, phenomenological thermodynamics — but in that context, the Gibbs entropy is just a property of the system's thermodynamic variables, and the modality is suppressed.

Fourthly, once it is recognised that in Gibbsian statistical mechanics 'equilibrium' is a statement about the probability distribution of a system, there is no contradiction between the (classical) recurrence theorem and the claim that entropy is non-decreasing. For the former tells us that any given system has some timescale at which it has returned to its initial state, and the latter (for Boltzmann-apt systems) tells us that at any time after the equilibration timescale the system is overwhelmingly likely to be in the equilibrium macrostate, and these statements are compatible. Nor is there any contradiction between the recurrence theorem and the claim that *thermodynamic* entropy is non-decreasing, for the latter concerns the interventions we may make on a system, and it is of no use to an agent to know that any given microstate will recur, absent knowledge that the system has in fact recurred at a given time.

Fifthly, while the quantum version of the recurrence theorem has a *uniform* timescale for recurrence, and so indicates that even the coarse-grained Gibbs entropy cannot be non-decreasing for all time, this simply indicates (assuming the orthodox coarse-graining version of the Gibbsian approach) that the assumptions underpinning the validity of coarse-graining cannot apply for arbitrarily long timescales (something that can in any case be read off the master-equation or BBGKY formalisms for irreversible Gibbsian dynamics; cf the discussion in Wallace (2016)). I suppose Callender and Maudlin should be pleased at this result: properly understood, Gibbsian statistical mechanics does not after all seek an exceptionless principle of non-decreasing entropy, but only an entropy that is nondecreasing over the physically significant timescales.

Finally, *Boltzmann* entropy is itself only superficially a categorical property

of a system. Yes, formally speaking the Boltzmann entropy depends only on a system's microstate, but it relies for its definition on a partition of the energy hypersurface into macrostates, and that partition is modal in nature — most obviously because the energy hypersurface itself depends on the dynamics. To make this vivid, suppose that a system has microstate x, and then perturb its Hamiltonian such that it remains constant in some small neighborhood of x but varies sharply over the rest of x's macrostate. Then the perturbation will adjust the macrostate partition, and thus change the Boltzmann entropy of x, even though no *categorical* property of the system has been altered.

9   Objections to Gibbs: coarse-graining

A second major locus of concern about the Gibbsian framework is the supposed inadequacy of any account based on coarse-graining (which, recall, is the mainstream — though not the only — approach to reconciling irreversibility with the time-invariance of fine-grained Gibbs entropy). It is variously described as ad hoc, as confusing our subjective limitations as experimenters with objective matters of fact, and as introducing some kind of spurious and empirically unsupported modification of the dynamics. Thus Maudlin (1995, pp.146-7):

> One can modify the underlying dynamics by adding some 'rerandomization' posit . . . but these surreptitious modifications simply have no justification.

Or Callender (1999, p.360):

> The usual response to the conservation of [the fine-grained Gibbs entropy] is to devise new notions of entropy and equilibrium, in particular, the coarse-grained entropy and a notion of equilibrium suitable for it. The motivation for these new notions is solely as a means of escaping the above "paradox" [time-invariance of fine-grained entropy], though it is usually defended with appeals to the imprecision with which we observe systems.

Here I identify three reasons that such concerns do not undermine the Gibbsian project (properly understood).

Firstly, in statistical mechanics the choice of coarse-graining is typically motivated neither by concerns about experimental limitations, nor through the ulterior motive of explaining irreversibility; rather, the motivation is the search for robust, autonomous higher-level dynamics. In the BBGKY approach to dilute-gas mechanics, for instance, the coarse-graining process is the discarding of three-body and higher marginals from the probability distribution; the motivation here is simply that we seem to be able to write down a well-defined and empirically successful dynamics for the so-truncated probability function. Similarly, in master-equation approaches to Brownian motion the reason for discarding information about the thermal bath in which the Brownian particle

moves is not (or should not be) that we do not have that information empirically; it is that there is a robust, empirically adequate, stochastic dynamics for the Brownian particle alone. (This is perhaps a good point to observe that statistical mechanics — contra the way it is often discussed in the foundational literature — is not itself a foundational project: its primary goal is to find empirically-adequate values and dynamical equations for collective degrees of freedom, not conceptually-adequate foundations for thermodynamics. See Wallace (2015a) for development of this point.)

Secondly, in thermodynamics there *is* reason to understand coarse-graining in terms of experimental limitations — but this is entirely appropriate given the control-theory understanding of thermodynamics. In the version of thermodynamics I sketched in sections 6–7, the coarse-graining can be understood as quantifying our operational limitations: if the best we can do is alter a system's macroscopic parameters and allow it to come to equilibrium, for instance, then the right coarse-graining is the one that replaces a distribution with the equilibrium distribution at the same volume and expected energy.

Thirdly, the macrostate partition at the heart of Boltzmannian statistical mechanics is just as vulnerable to these criticisms as is the Gibbsian coarse graining — indeed, it is a special case of that coarse-graining, corresponding in Gibbsian terms to replacing a distribution with that distribution which agrees on the probability of each macrostate and is constant across macrostates. Consider some standard descriptions of the coarse-graining:

> [W]e must partition [phase space] into compartments such that all of the microstates X in a compartment are macroscopically indistinguishable[.][9] (Callender 1999, p.355)

> Everyday macroscopic human language (that is) carves the phase space of the universe up into chunks. (Albert 2000, p.47)

If pushed, I suspect Boltzmannians would reply that it is not the epistemic indistinguishability of macrostates that is doing the work, but rather the possibility of writing down robust higher-level dynamics in terms of macrostates, and largely abstracting over microscopic details. But of course this is exactly what the Gibbsians have in mind when they speak of coarse-graining.

10   Objections to Gibbs: subjectivity

Probably the most severe criticism made of Gibbsian statistical mechanics is that it somehow conflates the question of how much we *know* about a system, with the question of how a system will *in fact* behave. Thus Albert (2000, p.58)

---

[9]Callender goes on to gloss this as ''that is, they share the same thermodynamic features''. But he does not define ''thermodynamic features'', and the standard definition — that thermodynamic features are restricted to the energy, volume and thermodynamic entropy, and functions thereof — does not suffice to define the macrostate partition.

> There's something completely insane (if you think about it) about the sort of explanation we have been imagining here . . . Can anybody seriously think that our merely being ignorant of the exact micro-conditions of thermodynamic systems plays some part in bringing it about, in making it the case, that (say) milk [mixes into[10]] coffee? How could that be? What can all those guys have been up to?

Or Callender (1999, p.360):

> Thermodynamic behaviour does not depend for its existence on the precision with which we measure systems. Even if we knew the positions and momenta of all the particles in the system, gases would still diffuse through their available volumes.

Many aspects of this concern overlap with worries about the modal or probabilistic nature of Gibbsian statistical mechanics, or the subjectivity of coarse-graining processes in particular, and so have been addressed in the previous two sections. Here I add just two further observations: one aimed at statistical mechanics, one at thermodynamics.

Firstly: insofar as we accept that statistical mechanics is in the business of making probabilistic predictions (and, as we saw in section 5, this is clearly implied by scientific practice) this reduces to a general concern about the *interpretation* of statistical-mechanical probability. For if we want to explain why a deterministic system will with high probability do X, probabilistic statements about its current state are pretty much all we can expect as explananda. If probabilities are here to be interpreted epistemically, then what is to be explained is why I have a high degree of belief in the system being (say) in the equilibrium macrostate τ seconds from now, and my present beliefs about the system are quite natural explananda.

It will be objected by (e.g.) Albert and Callender that when we say 'my coffee will almost certainly cool to room temperature if I leave it' we are saying something objective about the world, not something about my beliefs. I agree, as it happens; that just tells us that the probabilities of statistical mechanics cannot be interpreted epistemically. And then, of course, it is a mystery how they can be interpreted, given that the underlying dynamics is deterministic: perhaps as long-run relative frequencies, perhaps via the Lewis-Loewer best-systems approach (Lewis 1980, Loewer 2002), perhaps as the decoherent limit of quantum probabilites (Wallace 2016). But (on pain of rejecting a huge amount of solid empirical science, cf sections 4, 5, 7) *some* such interpretation must be available.

In the thermodynamic context, by contrast, it is far less clear to me why my knowledge of a system's state cannot play an explanatory role. To be sure, that information cannot explain why the system *spontaneously* approaches equilibrium; that is the domain of statistical mechanics. But there is no paradox in supposing that the transformations I can bring about of a system's state should

---

[10]Albert actually says 'dissolves in' here, but this is not strictly correct: milk, a mixture of water and water-insoluble lipids, is not itself soluble in water.

depend on my information about that very state. (If Albert had asked how our being ignorant of the exact microconditions of the coffee could make it the case that we are unable to un-mix the coffee and the milk, the rhetorical force is at least less clear.)

To be sure, in the phenomenological context this is largely irrelevant: information about a system is useless to me except insofar as I have access to operations which are sensitive to that information. (Even if provided with an exact readout of my coffee's microstate, I lack the manipulative precision to use that information to unmix it.) But that just brings us back to the previous section's discussion of coarse-graining.

11 Conclusion

The standard story in philosophy of statistical mechanics about the Gibbs and Boltzmann approaches relies on two assumptions: that the Gibbs and Boltzmann approaches are empirically equivalent and apply to the same physical systems, and that the Gibbsian framework has conceptual flaws that do not trouble the Boltzmannian.

I have argued that neither is correct. The standard objections to the Gibbsian approach either misunderstand the approach, or apply to certain ways of developing the approach but not to the approach itself, or apply with equal force to the Boltzmannian. And it is good that this is so, because physical as practiced in the 21st century requires the Gibbsian framework to handle a host of physically relevant situations in which the Boltzmannian approach is inapplicable. Correctly understood, modern statistical mechanics includes the Boltzmannian framework simply as a special case of the Gibbsian.

12 Epilogue: quantum statistical mechanics

My discussion in this paper is entirely classical. But the world is not even remotely classical, and so really we should ask: what does the Gibbs/Boltzmann distinction look like in quantum theory?

The obvious thought might be: replace phase space with (projective) Hilbert space; replace phase-space points with Hilbert-space rays; replace probability distributions over phase space with probability distributions over the space of rays. But macroscopic systems need not be, and in general will not be, in pure states. And mathematically, a probability distribution over the space of mixed states cannot be distinguished from a single element of that space.

So a hypothetical "Gibbsian" quantum statistical mechanics works with density operators *understood as probability distributions over mixed states*; a "Boltzmannian" statistical mechanics instead works with density operators *understood as individual mixed states*. But nothing at the level of the mathematics will distinguish the two approaches (see Wallace (2016) for further discussion). And I have been arguing that the *machinery* of the Gibbsian approach, not a hypo-

thetical *interpretation* of that machinery, is compatible with the Boltzmannian conception of statistical mechanics as objective. In the case of quantum theory, this seems to follow automatically: at the level of machinery, there *is* no difference between the two conceptions. In reality, modern quantum statistical mechanics simply studies the evolution of mixed states and coarse-grainings of those states, with no need within the formalism for any additionally statistical-mechanical conception of probability. Mathematically speaking the methods used are continuous with Gibbsian rather than Boltzmannian statistical mechanics, but it is mere semantics whether those methods should be called 'Gibbsian' or 'Boltzmannian'.

Acknowledgements

References

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

Altland, A. and B. D. Simons (2010). *Condensed Matter Field Theory* (2nd ed.). Cambridge: Cambridge University Press.

Bang, J., R. Pan, T. Hoang, J. Ahn, C. Jaryznski, H. T. Quan, and T. Li (2018). Experimental realization of Feynman's ratchet. *New Journal of Physics 20*, 103032.

Bishop, R. C. (2004). Nonequilibrium statistical mechanics Brussels-Austin style. *Studies in the History and Philosophy of Modern Physics 35*, 1–30.

Bustamante, C., J. Liphardt, and F. Ritort (2005). The nonequilibrium thermodynamics of small systems. https://arxiv.org/abs/cond-mat/0511629.

Callender, C. (1999). Reducing thermodynamics to statistical mechanics: the case of entropy. *Journal of Philosophy 96*, 348–373.

Callender, C. (2001). Taking thermodynamics too seriously. *Studies in the History and Philosophy of Modern Physics 32*, 539–553.

Carroll, S. (2010). *From Eternity to Here: The Quest for the Ultimate Theory of Time*. Dutton.

Collin, D., F. Ritort, C. Jarzynski, S. B. Smith, I. Tinoco, and C. Bustamante (2005). Verification of the Crooks fluctuation theorem and recovery of RNA folding free energies. *Nature 437*, 231–234.

Crooks, G. E. (1998). Nonequilibrium measurements of free energy differences for microscopically reversible Markovian systems markovian systems. *Journal of Statistical Physics 90*, 1481.

Feynman, R. P. (1967). *The Character of Physical Law*. Cambridge, Mass.: MIT Press.

Frigg, R. (2007). A field guide to recent work on the foundations of thermodynamics and statistical mechanics. In D. Rickles (Ed.), *The Ashgate Companion to the New Philosophy of Physics*, pp. 99–196. London: Ashgate.

Frigg, R. (2009). Typicality and the approach to equilibrium in Boltzmannian statistical mechanics. *Philosophy of Science 76*, 997–1008. Available online at http://philsci-archive.pitt.edu.

Goldstein, S. (2001). Boltzmann's approach to statistical mechanics. In J. Bricmont, D. Dürr, M. Galavotti, F. Petruccione, and N. Zanghi (Eds.), *In: Chance in Physics: Foundations and Perspectives*, Berlin, pp. 39. Springer. Available online at http://arxiv.org/abs/cond-mat/0105242.

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

Jarzynski, C. (1997). Nonequilibrium equality for free energy differences. *Physical Review Letters 78*(2690).

Jaynes, E. (1957a). Information theory and statistical mechanics. *Physical Review 106*, 620.

Jaynes, E. (1957b). Information theory and statistical mechanics ii. *Physical Review 108*, 171.

Jaynes, E. (1965). Gibbs vs boltzmann entropies. *American Journal of Physics 5*, 391–398.

Lebowitz, J. (2007). From time-symmetric microscopic dynamics to time-asymmetric macroscopic behavior: An overview. Available online at http://arxiv.org/abs/0709.0724.

Lewis, D. (1980). A subjectivist's guide to objective chance. In R. C. Jeffrey (Ed.), *Studies in Inductive Logic and Probability, volume II*. Berkeley: University of California Press. Reprinted, with postscripts, in David Lewis, *Philosophical Papers*, Volume II (Oxford University Press, Oxford, 1986); page numbers refer to this version.

Loewer, B. (2002). Determinism and chance. *Studies in the History and Philosophy of Modern Physics 32*, 609–620.

Malament, D. and S. L. Zabell (1980). Why Gibbs phase space averages work: the role of ergodic theory. *Philosophy of Science 47*, 339–349.

Maudlin, T. (1995). Review [of L. Sklar, *Time and Chance*, and L. Sklar, *Philosophy of Physics*]. *British Journal for the Philosophy of Science 46*, 145–149.

Myrvold, W. (2014). Probabilities in statistical mechanics. In C. Hitchcock and A. Hajek (Eds.), *Oxford Handbook of Probability and Philosophy*. Oxford University Press.

Myrvold, W. (2019). Beyond chance and credence. Forthcoming.

North, J. (2002). What is the problem about the time-asymmetry of thermodynamics? - a reply to Price. *British Journal for the Philosophy of Science 53*, 121–136.

Penrose, R. (1994). On the second law of thermodynamics. *Journal of Statistical Physics 77*, 217–221.

Ridderbos, K. (2002). The coarse-graining approach to statistical mechanics: how blissful is our ignorance? *Studies in the History and Philosophy of Modern Physics 33*, 65–77.

Rowlinson, J. S. (1970). Probability, information, and entropy. *Nature 225*, 1196.

Sklar, L. (1993). *Physics and Chance: Philosophical Issues in the Foundations of Statistical Mechanics*. Cambridge: Cambridge University Press.

Skrzypczyk, P., A. J. Short, and S. Popescu (2014). Work extraction and thermodynamics for individual quantum systems. *Nature Communications 5*, 4185.

Uffink, J. (2007). Compendium of the foundations of classical statistical physics. In J. Butterfield and J. Earman (Eds.), *Handbook for Philosophy of Physics*. Elsevier. Available online at philsci-archive.pitt.edu.

Wallace, D. (2011). Taking particle physics seriously: a critique of the algebraic approach to quantum field theory. *Studies in the History and Philosophy of Modern Physics 42*, 116–125.

Wallace, D. (2014). Thermodynamics as control theory. *Entropy 16*, 699–725.

Wallace, D. (2015a). The quantitative content of statistical mechanics. *Studies in the History and Philosophy of modern Physics 52*, 285–293. Originally published online under the title "What statistical mechanics actually does".

Wallace, D. (2015b). Recurrence theorems: a unified account. *Journal of Mathematical Physics 56*, 022105.

Wallace, D. (2016). Probability and irreversibility in modern statistical mechanics: Classical and quantum. To appear in D. Bedingham, O. Maroney and C. Timpson (eds.), *Quantum Foundations of Statistical Mechanics* (Oxford University Press, forthcoming).

Wallace, D. (2017). The case for black hole thermodynamics, part I: phenomenological thermodynamics. Forthcoming in *Studies in the History and Philosophy of Modern Physics*. https://arxiv.org/abs/1710.02724.

Wallace, D. (2018). On the plurality of quantum theories: Quantum theory as a framework, and its implications for the quantum measurement problem.

Forthcoming in S. French and J. Saatsi (eds.), *Scientific Metaphysics and the Quantum* (Oxford: OUP, forthcoming).

Zeh, H. D. (2007). *The Physical Basis of the Direction of Time* (5th ed.). Berlin: Springer.

Zwanzig, R. (2001). *Nonequilibrium Statistical Mechanics*. Oxford: Oxford University Press.