Taming Abundance: on the Relation between Boltzmannian and Gibbsian Statistical Mechanics

Charlotte Werndl and Roman Frigg

Published in Valia Allori (ed.): *Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism and Laws of Nature*, Singapore: World Scientific 2020, pp. 617-46

Abstract

Theoreticians working in statistical mechanics seem to be spoilt for choice. The theory offers two different theoretical approaches, one associated with Boltzmann and the other with Gibbs. These approaches are neither theoretically equivalent nor in any obvious way inter-translatable. This raises the question about the relation between them. We argue that Boltzmannian statistical mechanics is a fundamental theory while the Gibbisan approach is an effective theory, meaning that the former provides a true description of the systems within its scope while the latter offers an algorithm to calculate the values of physical quantities defined by the fundamental theory. This algorithm is often easier to handle than the fundamental theory and provides result where the fundamental theory is intractable. Being an effective theory, the Gibbsian approach works only within a certain domain of application. We provide a characterisation of the limits of the approach and show that BSM provides correct results in cases in which the two theories disagree.

Contents

1 Introduction

Theoreticians working in statistical mechanics (SM) seem to be spoilt for choice. The theory offers two different theoretical approaches, one associated with Ludwig Boltzmann and the other with J. Willard Gibbs. There are significant differences between the two approaches, which offer distinct descriptions of the same physical system. We refer to them as Boltzmannian SM (BSM) and Gibbsian SM (GSM) respectively. While one can cherish theoretical pluralism as a virtue, it does raise the question about the relation between the two approaches because two approaches are neither theoretically equivalent nor in any obvious way inter-translatable. The question of the relation between BSM and GSM becomes even more acute when we realise what functions they perform in the practice of SM. GSM is the drudge of SM. It provides the tools and methods to carry out a wide range of equilibrium calculations, and it is the formalism that yields the results in applications. However, as Lavis (2005) notes, when the question arises 'what is actually going on' in a physical system, physicists are often quick to desert GSM and offer an account of 'why SM works' in terms of BSM. And discrepancies are not restricted to foundational issues. In non-equilibrium situations BSM is usually the theory of choice because despite many attempts to extend GSM to non-equilibrium, no workable Gibbsian non-equilibrium theory has emerged (see Frigg (2008), Sklar (1993), and Uffink (2007) for reviews). But the practice of using one approach for everyday equilibrium calculations while explaining the non-equilibrium behaviour of physical systems and giving a foundational account of SM using the other approach is of questionable legitimacy as long as the relation between the two approaches remains unclear. What we need is an account of how the two approaches relate, and the account must be such that it justifies the customary division of labour. Abundance must be tamed.

Unfortunately, attempts to give explicit accounts of the relation between BSM and GSM are few and far between. In part this is due to the fact that there is a strand of arguments that downplays the problem. What drives such views is the claim that the two approaches are empirically equivalent, at least as far as equilibrium calculations are concerned.[1] Theoretical differences can then be brushed aside because discrepancies concerning foundational issues are something that one can live with. This argument is problematic for two reasons. First, while it is true that GSM and BSM produce the same predictions in many cases, agreement is not universal. In fact, there are important cases where GSM and BSM make conflicting predictions, which implies that GSM and BSM are not empirically equivalent (Werndl and Frigg 2017b, 2019a). This forecloses the escape route of non-committal theoretical pluralism.

[1]See, for instance, Davey (2009, 566-567) and Wallace (2015, 289). Arguments for special cases are given in Lavis (2005).

Where the status of one theory vis-à-vis the other is explicitly discussed, either the view is that GSM and BSM have to be reconciled (Lavis 2005), or it is suggested that GSM is the preferred formulation of SM (Wallace 2015). We are taking a different route and claim that BSM is a fundamental theory while GSM is an effective theory. This means that BSM provides a true description of the systems within the scope of SM; GSM offers an algorithm to calculate values defined by the fundamental theory. The algorithm is often easier to handle than the fundamental theory and provides result where the fundamental theory is intractable. As every effective theory, GSM works only within a certain domain of application. We provide a characterisation of the limits of GSM and show that BSM provides the correct results in cases in which the two theories disagree.

We discuss both approaches in the setting of classical systems. For want of space we state definitions and theorems only for the deterministic case. The generalisation to stochastic classical systems is straightforward,[2] and in Section 3 we briefly discuss an example that has a stochastic time evolution. There is an interesting question whether our approach generalises to quantum statistical mechanics. While we are optimistic that it does, we acknowledge that, given the current state of the discussion, this claim is largely speculative. The reason for this is that no generally accepted quantum formulation of BSM is currently available,[3] and so we lack the theoretical basis to compare BSM with GSM.

The paper is structured as follows. In Section 2 we introduce BSM and GSM, and in Section 3 we note that they are not empirically equivalent. In Section 4 we draw a contrast between fundamental and effective theories and argue that GSM is an effective theory while BSM is a fundamental theory. Effective theories are not universally applicable, and the most useful effective theories are ones for which we know the domain of applicability. In Section 5 we offer sufficient conditions for GSM to provide correct results. In Section 6 we briefly summarise our results and point out that regarding GSM as an effective theory has important repercussions for a number of projects, in particular attempts to turn GSM into a non-equilibrium theory.

[2]Statements of the relevant definitions and results can be found in Werndl and Frigg (2017a) and (2019a).

[3]See Dizadji-Bahmani (2011) for a discussion.

2 A Primer on BSM and GSM

SM studies physical systems like a gas in a container, a magnet on a laboratory table and a liquid in jar. Described mathematically, these systems have the structure of a *measure-preserving dynamical system*, i.e. a quadruple $(X, \Sigma_X, \phi_t, \mu)$. Here X is the state space of the system, i.e. a set containing all possible micro-states the system can be in. For a gas with n molecules X has $6n$ dimensions: three dimensions for the position of each particle and three dimensions for the momentum of each particle. Σ_X is a σ-algebra on X and μ is a measure on (X, Σ_X) (it is required to be invariant under the dynamics, meaning that $\mu_X(T_t(A)) = \mu_X(A)$ for all $A \in \Sigma_X$ and all t). The dynamics of the model is given by an *evolution function* $\phi_t : X \to X$, where $t \in \mathbb{R}$ if time is continuous and $t \in \mathbb{Z}$ if time is discrete. ϕ_t is assumed to be measurable in (t, x) and to satisfy the requirement $T_{t_1+t_2}(x) = T_{t_2}(T_{t_1}(x))$ for all $x \in X$ and all $t_1, t_2 \in \mathbb{R}$ or $\mathbb{Z}$. If at a certain point of time t_0 the system is in micro-state x_0, then it will be in state $\phi_t(x_0)$ at a later time t. For systems that are governed by an equation of motion such as Newton's equation, ϕ_t corresponds to the solutions of this equation. The *solution* (or *trajectory*) through a point x in X is the function $s_x : \mathbb{R} \to X$, $s_x(t) = T_t(x)$ (and *mutatatis mutandis* for discrete time).

At the macro-level the system is characterised by a set of *macro-variables*, which are measurable functions $v_i : X \to \mathbb{V}_i$, associating a value with each point in state space. Examples of macro-variables include volume, internal energy, and magnetisation. Mathematically speaking, macro-variables are real-valued functions on the state space, i.e. $f : X \to \mathbb{R}$. For example, if f is the magnetisation of the system and the system is in micro-state x, then $f(x)$ is the magnetisation of the system.

Both BSM and GSM share this characterisation of systems. What they disagree about is how statistical assumptions are introduced into SM and about what the observables of the theory are. We now introduce each theory in turn and describe how they differ.

In BSM a system is in a particular *macro-state* at any given time. A macro-state is defined by the values of a set of *macro-variables* $\{v_1, \ldots, v_l\}$ ($l \in \mathbb{N}$). This paper uses capital letters V_i to denote the values of v_i. A macro-state is then defined by a particular set of values $\{V_1, \ldots, V_l\}$. That is, the model is in macro-state $M_{V_1,\ldots,V_l}$ iff $v_1 = V_1, \ldots, v_l = V_l$.[4] A central posit of BSM is that macro-states *supervene* on micro-states, implying that at a system's micro-state uniquely determines its macro-state. This determination relation is normally many-to-one. Therefore, every macro-state

[4]Sometimes it is also useful to define macro-states by interval ranges, i.e. by the macro-variables taking values in a certain range or interval. One can then say that the model is in macro-state $M_{[A_1,B_1],\ldots,[A_l,B_l]}$ iff $V_1 \in [A_1, B_1], \ldots, V_l \in [A_l, B_l]$ for suitably chosen intervals.

M is associated with a *macro-region* X_M consisting of all micro-states for which the system is in M. For a complete set of macro-states the macro-regions form a partition of X (i.e. the different X_M do not overlap and jointly cover X).

One of these macro-states is the equilibrium macro-state of the system. Intuitively speaking, a system is in equilibrium when its properties do not change. This intuition is built into thermodynamics, where a system is said to be in equilibrium when all change has come to a halt and the thermodynamic properties of the system remain constant over time (Fermi 2000, 4). However, such a definition of equilibrium cannot be implemented in SM because measure-preserving dynamical systems exhibit Poincaré recurrence and time reversal invariance. As a consequence, when the time evolution of a system unfolds without any outside influence, the system will eventually return arbitrarily close to the micro-state in which it started. Hence a system starting outside equilibrium (for instance, when the gas was confined to one half of the container) will eventually return to that macro-state. So in SM no system will remain in any state forever. Obviously, this precludes a definition of equilibrium as the state which the system never leaves once it has reached it. Different formulations of BSM offer different prescriptions of how exactly to define equilibrium in SM. We base our discussion on the *long-run residence time definition of equilibrium*, which aims to come as close to the thermodynamic definition of equilibrium as the mathematical constraints imposed by measure-preserving dynamical systems permit (Werndl and Frigg 2015); we briefly comment on the typicality version of BSM below and point out that our main conclusion can equally be reached from this alternative point of view (and hence does not depend on which reading of BSM one adopts).

To give a formal statement of this definition, we first need the concept of the long-run fraction of time $LF_A(x)$ that a system spends in a subset A of X:[5]

$$LF_A(x) = \lim_{t\to\infty} \frac{1}{t} \int_0^t 1_A(T_\tau(x))d\tau, \tag{1}$$

where $1_A(x)$ is the characteristic function of A: $1_A(x) = 1$ for $x \in A$ and 0 otherwise. Note that long-run fractions depend on the initial condition.

The notion of 'most of the time' can be read in two different ways, giving rise to two different notions of equilibrium. The first introduces a lower bound of 1/2 for the fraction of time; it then stipulates that whenever a model spends more than half of the time in a particular macro-state, this is the equilibrium state of the model. Mathematically, let α be a real number in $(\frac{1}{2}, 1]$, and let ε be a very small positive

[5]We state the definitions for continuous time. The corresponding definitions for discrete time are obtained simply by replacing the integrals by sums.

real number. If there is a macro-state $M_{V_1^*,\ldots,V_l^*}$ satisfying the following condition, then that state is the system's α-ε-*equilibrium state*:

> There exists a set $Y \subseteq X$ such that $\mu_X(Y) \geq 1-\varepsilon$, and all initial states $x \in Y$ satisfy $LF_{X_{M_{V_1^*,\ldots,V_l^*}}}(x) \geq \alpha$. A system is in equilibrium at time t iff its micro-state at t, x_t, is in $X_{M_{V_1^*,\ldots,V_l^*}}$.

According to the second reading, 'most of the time' refers to the fact that the model spends more time in the equilibrium state than in any other state (and this can be less than 50% of its time). Mathematically, let γ be a real number in $(0,1]$ and let ε be a small positive real number. If there is a macro-state $M_{V_1^*,\ldots,V_l^*}$ satisfying the following condition, then that state is the system's γ-ε-*equilibrium state*:

> There exists a set $Y \subseteq X$ such that $\mu_X(Y) \geq 1-\varepsilon$ and for all initial conditions $x \in Y$: $LF_{X_{M_{V_1^*,\ldots,V_l^*}}}(x) \geq LF_{Z_M}(x) + \gamma$ for all macro-states $M \neq X_{V_1^*,\ldots,V_l^*}$. Again, a system is in equilibrium at time t iff its micro-state at t, x_t, is in $X_{M_{V_1^*,\ldots,V_l^*}}$.

It should come as no surprise that these two notions are not equivalent. More specifically, an α-ε-equilibrium is strictly stronger than a γ-ε-equilibrium in the sense that the existence of the former implies the existence of the latter but not vice versa.

These definitions are about the *time* a model spends in the equilibrium state. Hence it is not immediately clear what they imply about the *size* of the equilibrium macro-regions. It turns out that equilibrium regions, thus defined, are the largest macro-regions. More specifically, a macro-region is called β-*dominant* if its measure is greater or equal to β for a particular $\beta \in (\frac{1}{2}, 1]$. A macro-region is called δ-*prevalent* if its measure is larger than the measure of any other macro-region by a margin of at least $\delta > 0$. The following theorems can then be proved (Werndl and Frigg 2015b, 2017b):

> *Dominance Theorem*: If $M_{\alpha\text{-}\varepsilon\text{-}eq}$ is an α-ε-equilibrium, then the following holds for $\beta = \alpha(1-\varepsilon)$: $\mu_X(X_{M_{\alpha\text{-}\varepsilon\text{-}eq}}) \geq \beta$.[6]

> *Prevalence Theorem*: If $M_{\gamma\text{-}\varepsilon\text{-}eq}$ is a γ-ε-equilibrium, then the following holds for $\delta = \gamma - \varepsilon$: $\mu_X(X_{M_{\gamma\text{-}\varepsilon\text{-}eq}}) \geq \mu_X(X_M) + \delta$.[7]

It is a consequence of these definitions that a system is not always in equilibrium and that it does fluctuate away from equilibrium. This is a radical departure from thermodynamics. It is therefore worth pointing out that this is not merely a concession

[6] We assume that ε is small enough so that $\alpha(1-\varepsilon) > \frac{1}{2}$.

[7] We assume that $\varepsilon < \gamma$.

to the demands of measure-preserving dynamical systems. Having no fluctuations at all is also physically undesirable. There are experimental results that show that equilibrium is not the immutable state that classical thermodynamics presents us with because systems exhibit fluctuations away from equilibrium (MacDonald 1962; Wang et al. 2002). Hence adopting a notion of equilibrium that allows for fluctuations increases the empirical adequacy of the theory.

Before turning to GSM, we would like comment briefly on an alternative version of BSM, namely the typicality approach. In his seminal 1877 paper Boltzmann introduced what is today known as the *combinatorial argument*. It is a consequence of this argument that the equilibrium macro-region of an ideal gas is by far the largest macro-region. This prevalence of the equilibrium macro-region can be described in terms of *typicality*: equilibrium micro-states are typical in X because they occupy a region that is much larger than the region occupied by non-equilibrium states. Typicality then affords an explanation of the thermodynamic behaviour of a gas. As Goldstein puts it, because the phase space 'consists almost entirely of phase points in the equilibrium macrostate', '[f]or a non-equilibrium phase point $[x]$ of energy E, the Hamiltonian dynamics governing the motion $[x(t)]$ would have to be ridiculously special to avoid reasonably quickly carrying $[x(t)]$ into [the equilibrium macro-region] and keeping it there for an extremely long' (2001, 43-44).[8] So the typicality approach reaches the same conclusion as the long-run residence time approach, namely that the system spends most of its time in equilibrium. The two approaches differ in how they reach the conclusion – in fact, they reach them in reverse order. The typicality approach takes as its point of departure fact that the equilibrium macro-region is the largest macro-region and argues that this has the consequence that the system spends most of its time in equilibrium; the long-run residence time approach defines equilibrium as the state in which the system spends most of its time and then establishes that this implies that the equilibrium macro-region is also the largest region (either in the sense of prevalence or in the sense of dominance). At this point we are not concerned with the pros and cons of these approaches; nor are we concerned with their relative advantages vis-à-vis each other. What matters at this point is that both see equilibrium as the state in which the system spends most of its time, and hence the main point that we are making in this paper – that BSM is the fundamental theory while GSM is an effective theory – can equally be made from both perspectives.

The core object that is studied in GSM is a probability density (or distribution)

[8]The combinatorial argument is introduced in Boltzmann's (1877); for discussions of this argument see Albert's (2000), Frigg's (2008) and Uffink's (2007). The typicality approach originates in Goldstein's (2001) and Lebowitz's (1993a, 1993b). Discussions of typicality and its use in GSM can be found in Frigg's (2009, 2010), Frigg and Werndl's (2011), Uffink's (2007), Volchan's (2007), and Wilhelm's (forthcoming).

$\rho(x,t)$ over X.[9] The density $\rho(x,t)$ describes the probability of finding the state of a system in a region $R \subseteq X$ at time t:

$$p_t(R) = \int_R \rho(x,t)dx. \tag{2}$$

On physical grounds the probability density must be conserved, meaning that for every region $R(t)$ of X that is moving forward under the time evolution ϕ_t the probability must be constant. If the time evolution is generated by Hamiltonian equations of motion this is the case if, and only if, the Liouville's equation holds (Tolman 1938).

Gibbs (1902, 8) introduces what he refers to as the *condition of statistical equilibrium*: a probability density is in statistical equilibrium iff it is stationary. That is, if it does not change under the dynamics of the system: $\rho(x,t) = \rho(x)$ for all t. There are usually a large number of stationary density functions for a given ϕ_t. Hence the question arises which of these is the best to characterise a given physical situation. According to Gibbs, the so-called *microcanonical distribution* describes the equilibrium of a physical system which is completely isolated from its environment. It is the constant distribution on the system's energy hypersurface $H(x) = E$. The *canonical distribution* should be used when the system is in contact with a heat bath. It is given by $e^{-H(x)/kT}/\zeta_T$, where H is the system's Hamiltonian, T is the temperature, k is the Boltzmann constant, and ζ_T is the so-called *partition function*. For a discussion how to justify the choice of these distributions, see Frigg and Werndl's (2019) and Myrvold's (2016).

How do Gibbsian probability densities connect to observations on physical systems? That is, what does an experimentalist observe when measuring, say, the magnetisation of a sample of iron? To reply to this question we first introduce the *phase average* $\langle f \rangle$ of a macro-variable f:

$$\langle f \rangle = \int_X f(x)\rho(x,t)dx. \tag{3}$$

When the system is in statistical equilibrium, it follows that $\langle f \rangle$ is time-independent. Standardly a connection between the Gibbsian probability density and observable results is established by appealing to the *averaging principle* (AP). According to

[9] In Gibbs' (1902) original presentation $\rho(x)$ is described as representing an ensemble, an infinite collection of independent systems that are governed by the same laws of motion but are in different initial states. There are alternative presentations that endeavour to avoid reference to ensembles; they regard GSM simply as probabilistic algorithm. What follows does not depend on these interpretational issues and hence we set this question aside. Various different interpretations of GSM are discussed in Frigg and Werndl (2019).

this principle, when observing the physical quantity associated with f on a system in equilibrium, then the observed equilibrium value of f is the phase average $\langle f \rangle$. When reviewing textbooks of statistical mechanics it becomes clear that many textbooks explicitly state and endorse this principle (for a more detailed discussion of the principle, see Werndl and Frigg's (2019a), and references therein). For example, Chandler calls AP '[t]he primary assumption of statistical mechanics' (1987, p. 58), and Pathria and Beale state that they regard AP as the 'the most important result' in SM (2011, p. 31). For this reason our discussion in Section 4 is based on a version of GSM that incorporates AP. However, there are alternative interpretations of GSM that do not accept AP. We will comment on how our arguments carry over to these alternative versions at the end of Section 5.

These brief summaries of BSM and GSM should make it clear how distinct the two theories are. In particular, their characterisations of equilibrium are entirely different. BSM first introduces macro-states and then defines the equilibrium macro-state as the macro-state in which the system spends most of its time. In doing so, it explicitly allows systems to fluctuate away from the equilibrium state every now and then. In GSM, by contrast, equilbrium is a property of a probability distribution. More specifically, it is defined as a stationary probability distribution. Observable equilibrium properties are equated with the phase averages of macro-variables, and these phase averages stay constant over time if the distribution is in equilibrium.

So we seem to find ourselves in the disconcerting situation that when we talk about 'statistical mechanics' it is unclear whether we mean BSM or GSM (or both), and the two theories are different in important respects.

As a first reaction one might try to downplay the problem by arguing that despite their theoretical differences, the formalisms are empirically equivalent, at least as far as equilibrium properties are concerned. This immediately raises another question: what does it mean for BSM and GSM to be empirically equivalent? The Boltzmannian notion of equilibrium is formulated to mirror the thermodynamic notion of equilibrium; the Gibbsian notion of statistical equilibrium connects to thermodynamic equilibrium through the averaging principle. Hence it is natural to think that Gibbsian phase averages, Boltzmannian equilibrium values, and thermodynamic equilibrium values should all coincide. This suggests that the following is a necessary condition for the empirical equivalence of BSM and GSM:

$$F \approx \langle f \rangle \tag{4}$$

holds for *all* macro-variables f in all systems that fall within the scope of both theories, where F is the Boltzmannian equilibrium value of f and '$\approx$' means that the two

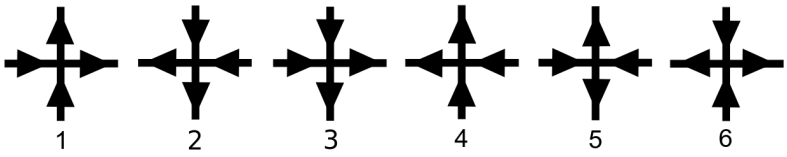

Figure 1: The configurations of the six-vertex model.

values are approximately equal. We call this the *mechanical averaging principle* and refer to Equation (4) as the *mechanical averaging equation* ('mechanical' because the principle connects two mechanical quantities, namely equilibrium values in BSM and equilibrium values in GSM). The question now is whether this principle holds true.

3 When the Mechanical Averaging Equation Fails

BSM and GSM turn out not to be empirically equivalent. Boltzmannian equilibrium values and Gibbsian phase averages do agree for paradigmatic examples such as the dilute gas with macro-variables that assign the same value to all states that are in the Maxwell-Boltzmann distribution (or in a distribution that is very close to the Maxwell-Boltzmann distribution). Yet there are important cases where F and $\langle f \rangle$ are substantially different. This shows that the mechanical averaging equation does not hold generally, implying that the mechanical averaging principle fails. We will now present the six vertex model with the internal energy macro-variable as an example in which Gibbsian and Boltzmannian calculations come apart for any finite number of particles N. Further examples where the Boltzmannian equilibrium values and the Gibbsian phase averages disagree are discussed in Werndl and Frigg's (2019a).

Consider a two-dimensional quadratic lattice with N grid points that lies on a two-dimensional torus (which allows us to neglect border effect because on a torus every grid point has exactly four nearest neighbours). The 'vertices' are the grid points and each vertex is connected to its four nearest neighbours by an edge. On each edge there is either an arrow pointing towards or away from the vertex. A rule known as the 'ice-rule' is now imposed: the arrows have to be distributed such that at each vertex in the lattice there are exactly two outward and two inward pointing arrows. As shown in Figure 1, at every vertex there are exactly six configurations of the arrows that satisfy the ice-rule. The ice rule is satisfied by water ice and several crystals including potassium dihydrogen phosphate (Baxter 1982; Lavis and Bell 1999).

The micro-state of the six-vertex model $\kappa = (\kappa_1, \ldots, \kappa_N)$ is specified by assigning to each vertex in the model one of the six types of configurations of the arrows permitted

by the ice rule. Each of the six configurations has a certain energy ϵ_i, $1 \leq i \leq 6$. Denote by $\epsilon(\kappa_j)$ the energy or the j^{th} vertex (thus all $\epsilon(\kappa_j)$ range over the ϵ_i). Then the energy of the state κ is:

$$E(\kappa) = \sum_{j=1}^{N} \epsilon(\kappa_j). \tag{5}$$

We make the common assumption that the energy of the different configurations is $\epsilon_1 = \epsilon_2 = 0$ and $\epsilon_3 = \epsilon_4 = \epsilon_5 = \epsilon_6 = 1$ (cf. Lavis and Bell 1999, 299). The canonical distribution

$$\rho(\kappa) = e^{-E(\kappa)/kT}/\zeta_T \tag{6}$$

with

$$\zeta_T = \sum_{\kappa} e^{-E(\kappa)/kT} \tag{7}$$

is usually taken to be the probability distribution. There are many versions of the six-vertex model, but most versions work with a stochastic dynamics that is assumed to be an irreducible Markov model. Note that the canonical distribution is historically associated with Gibbs. Yet this should not mislead us to believe that we treat the model in a Gibbsian way right from the start. The canonical distribution *per se* is neither Gibbsian nor Boltzmannian and is simply used here as a probability distribution that can figure in either BSM or GSM.[10]

Let us now consider the *internal energy* defined in Equation (5) as the relevant macro-variable. In the *Boltzmannian* perspective the state space consist of all possible states κ which satisfy the ice rule. $E = 0$ is the lowest energy value and it defines a macro-state M_0 with the corresponding macro-region $\bar{X}_{M_0} = \{\kappa^*, \kappa^+\}$, where κ^* is the state for which all vertices are in the first configuration, and κ^+ is the state for which all vertices are in the second configuration. Suppose now that the number of vertices N is a sufficiently large but finite number. Due to the fact that for sufficiently low values of the temperature T the probability mass is concentrated on the two lowest energy states, it is the case that $\bar{X}_{M_0}$ is the largest macro-region for sufficiently low temperatures. For an irreducible Markov process the model spends more time in the largest macro-region than in any other of the macro-regions. For this reason M_0 is a Boltzmannian γ-0-equilibrium, and the Boltzmannian equilibrium value is $E = 0$.

[10] As noted in the Introduction, BSM can be formulated with stochastic rather than a deterministic time evolution. In most basic terms this means that one replaces ϕ_t with a stochastic algorithm, and such algorithms can be formulated using the canonical distribution. For details see Werndl and Frigg (2017a).

In the Gibbsian treatment $\rho(\kappa)$ is the stationary equilibrium distribution and the observable f is of course the internal energy. We know that the internal energy assumes its lowest value $E = 0$ only for the two specific micro-states κ^* and κ^+, and that it will assume higher values for all other micro-states (and we know that for any $T > 0$ there will be a non-zero probability assigned to these higher energy states). Consequently, the phase average $\langle E \rangle$ is greater than 0; thus it is higher than the Boltzmannian equilibrium value, implying that the mechanical averaging equation fails.

It is not difficult to see that this difference can be significant: choose a T such that $\{\kappa^*, \kappa^+\}$ is the largest macro-region while its probability is less than 0.5.[11] In such a case the Boltzmannian equilibrium value is still $E = 0$. For the six vertex model the second lowest macro-value is $E = \sqrt{N}$, which corresponds to the energy of micro-states where all columns except one are taken up by states with the first or the second configuration, and the states in the exceptional row are all states of the third or fourth configuration.[12] Consequently $\langle E \rangle$ is higher than $\sqrt{N}/2$, implying that the Gibbsian phase average and the Boltzmannian equilibrium value will differ by at least $\sqrt{N}/2$. This is clearly not a negligible difference, in particular for large N, and hence the mechanical averaging equation fails. This is underscored by the fact that the Boltzmannian macro-value that is closest to the value obtained from Gibbsian phase averaging is higher or equal to $\sqrt{N}$. And a Boltzmannian macro-value of higher or equal to $\sqrt{N}$ is different from 0, the Boltzmannian macro-value in equilibrium.[13]

[11] This is possible because the higher the temperature, the more uniform the probability distribution. Consequently, for sufficiently high values of T, the largest macro-region is different from $\{\kappa^*, \kappa^+\}$. Because of the continuity of the canonical distribution, there has to be a T such that $\{\kappa^*, \kappa^+\}$ is the largest equilibrium macro-region while its probability is less than 0.5.

[12] This is the smallest possible departure from states with zero energy: it can be shown that the number of downward pointing arrows is the same for all rows. From this is follows that there has to be a perturbation in each row. Therefore $\sqrt{N}$ is the second lowest value of the internal energy (Lavis and Bell 1999, Chapter 10).

[13] We note that the the internal energy macro-variable considered here is an *extensive* macrovariable, i.e. it depends on the number of constituents of the system. If one instead considers the energy density, an *intensive* macro-variable, then one finds that difference between the Gibbsian and Boltzmannian equilibrium calculations tends toward zero as $N \to \infty$ because $\sqrt{N}/2N \to 0$. This illustrates the point that whether or not Gibbsian and Boltzmannian calculations agree crucially depends on the macro-variable. Yet we note that the problems we are discussing cannot be dissolved simply by restricting attention to intensive variables. Extensive variables are important (cf. Lieb and Yngvason 1999), and, as discussed in Werndl and Frigg (2019a), in some cases there differences between Gibbsian and Boltzmannian values both for intensive and extensive variables.

4 GSM as an Effective Theory

We argued that GSM and BMS are not empirically equivalent, which forces upon us the question of how the two approaches relate to one another and of which prediction is correct if they disagree. The answer to these question, we submit, lies in the realisation that BSM and GSM are not alternative theories that are on par with each other. Rather, GSM is an *effective theory* while BSM is the fundamental theory. This implies that in situations where Boltzmannian and Gibbsian equilibrium values come apart, the Boltzmannian values are the correct values.

To add specificity to the claim that GSM is an effective theory we have to characterise effective theories in more detail. Physicist James Wells says that effective theories 'are theories because they are able to organise phenomena under an efficient set of principles, and they are effective because it is not impossibly complex to compute outcomes' (2012, 1). The ability of effective theories to provide results comes at a cost of incompleteness. As Wells puts it, '[t]he only way a theory can be effective is if it is manifestly incomplete [...] Any good Effective Theory systematises what is irrelevant for the purposes at hand. In short, an Effective Theory enables a useful prediction with a finite number of input parameters.' (*ibid.*) His examples of effective theories are Galileo's law of falling bodies, the harmonic oscillator, classical gravity, and effective theories of particle masses.

Our suggestion is that GSM should be added to this list because it fits Wells' criteria. First, by offering a characterisation of equilibrium in terms of stationary distributions, GSM offers an *organisation* of phenomena under the umbrella of small set of principles. Second, GSM offers actionable principles and tractable methods to calculate equilibrium values of large array of materials, which makes it an efficient tool for *computations*. Third, GSM is *incomplete* in a number of ways. As noted in Section 2, GSM is unconcerned with the dynamics of the model. The role of the system's dynamics in GSM is reduced to ensuring that the proposed equilibrium distribution is stationary, and no other properties of the dynamics play any role in the theory. GSM considers neither equations of motion nor dynamical laws; it completely disregards trajectories; no time averages along trajectories are studied; and the initial conditions are left unspecified.[14] The system's dynamics is considered immaterial to understanding equilibrium as long as it – somehow – produces the stationary distribution that enters into the calculations. The system's Hamiltonian is used in formulating the most common Gibbsian distributions – the microcanonical and the

[14]Notions of this kind are sometimes considered in attempts to justify the Gibbsian formalism, but they are not part of the formalism itself. For a discussion of justificatory endeavours see, for instance, Frigg's (2008).

canonical distributions – but a Hamiltonian is does not, by itself, provide dynamical information. The Hamiltonian of the system becomes relevant to the dynamics only when combined with equations of motions or when used in the formulation of a stochastic process, which is not part of the Gibbsian formalism. The Gibbsian phase averages are the same for all time evolutions that are such that ρ is invariant over time, no matter how different they may otherwise be. Finally, GSM is explicit about what it omits, and it thereby *systematises* what it regards as irrelevant.

These features of GSM stand in stark contrast with BSM, where dynamical considerations occupy centre stage. As we saw in Section 2, GSM introduces macro-states with corresponding macro-regions, and then defines equilibrium in explicitly dynamical terms (namely as the macro-state whose macro-region is such that, in the long run, the system's state spends most of its time in that macro-region). Under the assumption (adopted in this paper) that the world is classical and that systems are governed by classical laws of motion, the dynamics considered in BSM is the true dynamics at the fundamental level: the unabridged and unidealised dynamics with all interactions between all micro-constituents of the system. Equilibrium results from macro-states that are defined in terms of macro-variables that supervene on the true micro-dynamics of the system, and where a system fluctuates away from equilibrium it does so as a result of the true underlying dynamics. The theory gives a full account of what happens in a classical world – nothing is left out and nothing is averaged over. BSM therefore is the fundamental theory.

Since a true and complete fundamental theory cannot be wrong, the results of BSM are the correct results in cases where BSM and GSM disagree. Consider the example of magnet with the macro-variable of total magentization m as discussed in Werndl and Frigg's (2019). Such a system can be represented by the Ising model. Experiments show a phase transition as the temperature is varied. Calculations in the Boltzmannian framework show this phase transition as one would expect. Yet the Gibbsian account fails to describe the phase transition successfully because it yields that $\langle m \rangle = 0$ for all temperature values. Hence the results of BSM are in agreement with the experimental results but the Gibbsian results obtained by phase averaging are not.

The flip side of fundamentality is often intractability, and BSM is no exception. If one wants to find out whether a Boltzmannian equilibrium exists, and if so, determine the equilibrium state, then one has to explicitly specify the macro-state structure of the model and determine the macro-states' macro-regions; one has to know enough about the underlying dynamics to be able to calculate the long-run fractions of time that a model spends in each macro-region; and one has to be able to estimate the measure

of the set of initial conditions that lie on trajectories that do not have well-behaved long-run fractions of time. This feasible only in very special cases (for instance, if the dynamics is ergodic); in general it is a dead end because it requires more information than we have. The omissions in effective theories are designed to eliminate these intractabilities and deliver workable recipes. So the omissions and simplifications are what make effective theories effective!

Relegating a host of things to the realm of irrelevance comes at a cost. Wells is careful in pointing out that whenever we recognise a theory as an effective theory we will also have to 'confront a theory's flaws, its incompletenesses, and its domain of applicability as an integral part of the theory enterprise' (*ibid.*). In other words, effective theories have limited domains of applicability, and using the theory correctly requires scientists to know where the limits are. For this reason, Wells notes that '[t]he most useful Effective Theories are ones where we know well their domains of applicability' (*ibid.*).

So if we view GSM as an effective theory, we have to be able to delimit its range of application. That is, we have to be able to say when it yields trustworthy results and when its procedures fail to deliver. This is the task to which we turn now.

5 The Boundaries of Effectiveness

Recall that for BSM and GSM to agree on a system's equilibrium properties it must be the case that $F \approx \langle f \rangle$ (where F is the Boltzmannian equilibrium value), and agreement is a necessary condition for BSM to be effective. This can be the case under several different conditions. We will now discuss four conditions that are individually sufficient for $F \approx \langle f \rangle$ to hold: the Khinchin condition, the requirement that fluctuations be small, the averaging equivalence theorem and the cancelling out theorem. However, note that these conditions are not necessary and there could be other conditions that guarantee that the phase average equals the Boltzmannian equilibrium value.

It is trivial that phase averages and Boltzmannian equilibrium values are identical if the macro-variables under considerations take the same value everywhere, i.e. $f(x) = c$ for all x in X and a constant c. Then $F = \langle f \rangle = c$. Clearly, such macro-variables are not interesting; yet they raise a useful question: how far does one have to move away from this uninteresting case to obtain a useful condition while still retain the basic idea? The Khinchin condition provides an answer to this question.[15]

[15]The condition owes its name to Khinchin (1949), who engaged in a systematic study of functions that satisfy strong symmetry requirements and therefore have small fluctuations for systems with a large number of constituents. There are at least two versions of the condition. We here focus

First of all we need the notion of a *fluctuation*. For an arbitrary micro-state x consider the difference between the value $f(x)$ (the true value of the observable if the model is in state x) and the phase average:

$$\Delta f(x) = f(x) - \langle f \rangle. \tag{8}$$

$\Delta f(x)$ is the fluctuation when the system is in micro-state x, and $|\Delta f(x)|$ is the *magnitude* of the fluctuation. The *Khinchin condition* then requires that there is a subset $\bar{X}$ of X with $\mu(\bar{X}) = 1 - \delta$ for a very small $\delta \geq 0$ such that $|\Delta f(x)| = 0$ for all x in $\bar{X}$.

If the condition is satisfied, we have $F = \langle f \rangle$, which can be seen as follows. Assume that a Boltzmannian equilibrium exists and that F is the Boltzmannian equilibrium value of f. According to the Khinchin condition then there are only a few states (of at most measure δ) whose macro-values differ from $\langle f \rangle$. It is clear that these 'exceptional' states cannot form the Boltzmannian equilibrium macro-state because the macro-region corresponding to the Boltzmannian macro-state is the largest macro-region. We conclude that the set of micro-states for which $f(x) = F$ has to be the macro-region of the Boltzmannian macro-state, and for the states in that region it follows that $F = \langle f \rangle$. Therefore we conclude that if the Khinchin condition is satisfied, then BSM equilibrium value and the Gibbsian phase average agree. The paradigmatic example for the Khinchin condition is the dilute gas with macro-variables that assign the same value to all states that are in the Maxwell-Boltzmann distribution, or in a distribution that is very close to the Maxwell-Boltzmann distribution (Ehrenfest and Ehrenrest-Afanassjewa 1959).

An alternative approach is to focus on the statistics of fluctuation patterns and then to show that GSM reproduces the fluctuation patterns of BSM under certain conditions. More specifically, let us start by looking at fluctuations from a Gibbsian perspective. The core idea here is to use the probabilities of GSM as given in Equation (2) to calculate the probability that a fluctuation of a certain magnitude occurs. In more detail: given an interval $\delta := [\delta_1, \delta_2]$, where δ_1 and δ_2 are real numbers such that $0 \leq \delta_1 \leq \delta_2$, Equation (2) can then be used to arrive at the probability for a fluctuation of a magnitude between δ_1 and δ_2 to occur:

$$p(\delta) = \int_D \rho(x) dx, \tag{9}$$

on the first version, which is the one appealed to, amongst others, by Wallace (2015, 289), Lavis (2005, 267-268); Malament and Zabell (1980, 344-345), and Vranas (1998, 693). The second version is appealed to in Ehrenfest and Ehrenfest-Afanassjewa's (1959, 46-52) and discussed in Werndl and Frigg's (2019c).

where $D = \{x \in X : \delta_1 \leq |\Delta(t)| \leq \delta_2\}$.

It is important to interpret the scope of this equation correctly. Sometimes the probabilities in Equation (2) are interpreted as holding universally. That is, ρ is seen as providing the correct probabilities for the state of a system to be in region R at time t for *all* R in X and for *any* time t. Under such an interpretation the fluctuation probabilities in Equation (9) are then seen as universal in the sense that for any magnitude and for any time t, $p(\delta)$ gives the probability for a fluctuation of a certain magnitude to occur at t. Yet universality of this kind is a very strong demand and fails in general. A careful study at GSM reveals that at least one of two conditions have to be met in order for this to be the case (for more on those two conditions, see Frigg and Werndl 2019). First, the *masking condition* requires either that the system has access to all parts of phase space, or, if that is not the case, that f must be such that the proportion of states for which f assumes a particular value is the same in each invariant subset of X. Second *f-independence* (roughly) states that the dynamics of the system must be such that the probability of finding a specific value of f in two consecutive yet sufficiently temporally distant measurements have to be (approximately) independent of each other. The Gibbsian ρ can be used to calculate correct fluctuation probabilities only if at least one of these conditions is satisfied. These conditions limit the scope of GSM in determining fluctuations. Since both conditions are strong conditions on the dynamics and the macro-variables, their satisfaction cannot be taken for granted and they limit the scope of GSM in determining fluctuations.

Let us now turn to BSM and first focus on the masking condition and explain how, from the perspective of BSM, the fluctuation probabilities of Equation (9) turn out to be correct. The starting point for the masking condition is to observe the behaviour of the same system over time and to consider the fluctuations that arise in this way. What this amounts to is to track the system over an infinite period of time when the system starts in a particular initial condition and its state evolves under the dynamics of the system. According to the masking condition, either the system can access all parts of X or the proportion of states for which f assumes a particular value is the same in each invariant subset of X. From this it follows immediately that from a Boltzmannian perspective the fluctuations that arise in the same system over an infinite period of time are equal to the probabilities assigned to the fluctuations by the measure ρ, implying that Equation (9) holds. For example, suppose that a system spends, say, β of its time in a certain macro-state for which the function f assumes the value F'. Then for this macro-state the magnitude of the fluctuation away from the phase average is $|F' - \langle f \rangle|$. Suppose that δ_0 is the interval that consists only of $|F' - \langle f \rangle|$. Then the probability $p(\delta_0)$ has to be β.

Let us now turn to the second condition, i.e. f-independence, and again explain how in the Boltzmannian framework the fluctuation probabilities of Equation (9) turn out to be the correct ones. For an observable f with a finite number of macro-states, suppose that the dynamics of the system is such that for two points of time t_1 and t_2 that are sufficiently far apart f-independence holds, meaning that the probability of finding a specific value of f in the two measurements are approximately independent of each other. From a Boltzmannian perspective this immediately implies that, given a specific macro-value at t_1, the probability of finding the system in a macro-value at t_2 is given by the probability measure ρ, and hence again Equation (9) holds. For example, suppose that the measure assigns β to a certain macro-state for which the function f assumes the value F'. For this macro-state the magnitude of the fluctuations away from the phase average is $|F' - \langle f \rangle|$. Suppose δ_0 is the interval consisting only of $|F' - \langle f \rangle|$. It then follows that, assuming that the system was in a certain macro-state at t_1, the probability of obtaining the fluctuation δ_0 at t_2 (where t_1 and t_2 are sufficiently far apart) is given by the probability $p(\delta_0) = \beta$.

The third condition that is sufficient to guarantee the equality of the Boltzmannian equilibrium value and phase averages is given by the Average Equivalence Theorem (Werndl and Frigg 2017).[16] The conditions of this theorem are referred to as the 'Average Equivalence Conditions'.

> **Average Equivalence Theorem (AET)**. Assume that a model is composed of $N \geq 1$ components. That is, the state $x \in X$ is given by the N coordinates $x = (x_1, \ldots, x_N)$; $X = X_1 \times X_2 \ldots \times X_N$, where $X_i = X_{oc}$ for all i, $1 \leq i \leq N$ (X_{oc} is the one-component space). Let μ_X be the product measure $\mu_{X_1} \times \mu_{X_2} \ldots \times \mu_{X_N}$, where $\mu_{X_i} = \mu_{X_{oc}}$ is the measure on X_{oc}.[17] Suppose that the macro-variable K is the sum of the one-component variable, i.e. $K(x) = \sum_{i=1}^{N} \kappa(x_i)$ (the assumptions made here is that all sums of possible values of the one-component variable are possible values of the macro-variable). Then it follows that the value corresponding to the largest macro-region as well as the value obtained by Gibbsian phase averaging is $\frac{N}{k}(\kappa_1 + \kappa_2 + \ldots \kappa_N)$.

For a proof of the theorem, see Werndl and Frigg's (2017). Note that if the theorem is used to make claims about Boltzmannian equilibrium (as we do here), dynamical assumptions have to come in by assuming that a Boltzmannian equilibrium exists.

[16] In Werndl and Frigg's (2017) the theorem was referred to as 'equilibrium equivalence theorem'. This name turned out to be potentially misleading because the theorem concerns the largest macro-region and *not* the a Boltzmannian equilibrium *per se*. For this reason we now use the label 'average equivalence theorem'.

[17] It is assumed here that N is a multiple of k, i.e. $N = k * s$ for some $s \in \mathbb{N}$.

More specifically, if a Boltzmannian equilibrium state exists, it follows from the dominance/prevalence theorems that the Boltzmannian equilibrium value is equal to the value of the largest macro-region. From the Average Equivalence Theorem it then follows that this Boltzmannian equilibrium value is equal to the Gibbsian phase average.

The most important assumptions of the theorem are (i) that the macro-variable is a sum of variables on the one-component space (where all sums of possible values of the one-component variable are possible values of the macro-variable); (ii) that the macro-variable on the one-component space corresponds to a partition with cells of equal probability; and (iii) that the measure on state space is the product measure of the measure on the one-component space. To some extent these assumptions are restrictive; still, a number of standard applications of SM fall within the scope of the theorem. One instance of the AET is the baker's gas; another instance of the AET is the Kac ring with the standard macro-state structure; and yet another example is the ideal gas with N particles (see Werndl and Frigg's (2019) for details).[18]

The fourth sufficient condition to guarantee the equality of the Boltzmannian equilibrium value and the Gibbisan phase average is given by the Cancelling Out Theorem (the conditions of this theorem will be refereed to as the 'Cancelling Out Conditions'). Let us first state the theorem (for the proof, see Werndl and Frigg 2019):

> **Cancelling Out Theorem (COT).** Consider a deterministic model with Boltzmannian equilibrium macro-state M_{equ} with equilibrium value V_{equ} and other macro-states $M_1, \ldots, M_q$, $q \in \mathbb{N}$, with corresponding macro-values $V_{M_1}, \ldots, V_{M_q}$. Further, suppose that for any macro-state $M_i \neq M_{equ}$ there is a macro-state M_j such that (i) $\mu_X(X_{M_i}) = \mu_X(X_{M_j})$ and (ii) $V_{M_i} + V_{M_j} = 2V_{M_{equ}}$. Then the Boltzmannian equilibrium value as well as the value obtained by phase averaging is V_{equ}.

[18] One might think that there is a similarity between AET and the weak law of large numbers (LLN), which states that given independent and identially distributed random variables (which we consider in the AET theorem) for any $\varepsilon > 0$ (cf. Meester 2003, Section 4.1):

$$\mu\left(\left\{x : \left|\frac{\sum_{i=1}^{N}\kappa(x_i)}{N} - \frac{(\kappa_1 + \kappa_2 + \ldots \kappa_N)}{k}\right| < \varepsilon\right\}\right) \geq 1 - \frac{\sigma^2}{\varepsilon^2 N}. \tag{10}$$

This similarity is superficial, and the theorems are different. First the LLN does *not* say whether the values of the *extensive* macro-variables we consider in the AET, $\sum_{i=1}^{N}\kappa(x_i)$, are close to $N(\kappa_1 + \kappa_2 + \ldots \kappa_N)/k$. All one obtains from the LLN is that their values are within $N\varepsilon$, but $N\varepsilon$ can be very large. Second, AET and LLN are results about *different* macro-variables. AET is a result about the macro-variable $\sum_{i=1}^{N}\kappa(x_i)$ or, if it is divided by N, about $\sum_{i=1}^{N}\kappa(x_i)/N$. By contrast, LLN is a statement about the probability of states that are close or equal to $(\kappa_1 + \ldots + k_k)/k$. Hence it can tell us something about the *different* macro-variables that are defined by assigning the same macro-value to all states that are close or equal to $(\kappa_1 + \ldots + k_k)/k$.

Intuitively, the theorem states the following: if the state space is divided up in such a way that next to the largest macro-region (corresponding to the Boltzmannian equilibrium) there are always two macro-states of equal measure such that their average equals the Boltzmannian equilibrium value, then it follows that the Boltzmannian equilibrium value equals the Gibbsian phase average, and GSM can serve as an effective theory. Needless to say, the assumptions of the theorem are to some extent restrictive because it requires the macro-state structure to be of a special kind. Still, a number of standard cases in SM such as the six-vertex model with the polarisation macro-variable for sufficiently high temperatures or the Ising model with the magnetisation macro-variable for sufficiently high temperatures fall under the scope of the theorem (cf. Cipra 1987; Lavis and Bell 1999, Chapter 3).

As noted in Section 2, there is an alternative interpretation of GSM that does not include AP. On such an interpretation the theoretical core of GSM contains only ρ, while Equation (4), the mechanical averaging equation, has the status of a pragmatic rule that is adopted *only when it provides correct results*. When this equation fails, GSM is simply silent about the correct equilibrium values. This move immunises GSM against arriving at calculations that disagree with the calculations of BSM, but it does so at the cost of further restricting the scope of GSM. This is not *per se* objectionable, but it changes nothing fundamentally in our argument. On this alternative interpretation GSM is still an effective theory with a limited range of applicability (and the limits are identical to the limits of the standard interpretation). The only difference is that in cases in which there would be disagreement, GSM is now seen not as giving wrong results but rather as providing no results at all.[19]

6 Conclusion

We argued that GSM is an effective theory while BSM is a fundamental theory. This clarifies the relation between the two approaches. We presented an account of effective theories and showed that GSM fits the relevant criteria. The range of application of effective theories is limited, which raises the question under what circumstances the calculations of GSM are correct. We presented four conditions under which this happens; these are individually sufficient but not necessary. This means that other,

[19]Furthermore, as argued in Frigg and Werndl's (2019), there is no single reasonable interpretation of Gibbs that can make sense of all the successful applications of Gibbs. Reasonable interpretations of Gibbs such as the fluctuation account can always only explain some of the applications of GSM. That there is no single reasonable interpretation of Gibbs that can account for all successful applications of GSM further strenghtens the view that GSM is an effective theory.

yet unknown, conditions could exist. It is, however, an open question whether such conditions do exist and whether there is complete list of such conditions.

Classifying GSM as an effective theory has implications for foundational debates. It implies that GSM does not address foundational questions and that such questions should not be discussed in that theory. Asking whether GSM provides a correct fundamental description of the world, or, if the answer to this question is negative, trying to revise GSM so that it does provide such a description, is a mistaken endeavour. Effective theories do not offer fundamental descriptions; they are calculatory devices of instrumental value; no more and no less.

There have been (and still are) sustained attempts to turn GSM into fundamental theory. Probably the most important programme of that kind aims to extend GSM to non-equilibrium processes. It is common to characterise the approach to equilibrium as a process of increasing entropy. This means that if a system is prepared in a non-equilibrium macro-state of low entropy, then, as soon as the constraints are lifted and the system evolves freely, the entropy should increase until it reaches a maximum. This does not happen in GSM because the Gibbs entropy, defined as $\int_X \rho \ln(\rho) dx$, is a constant of motion. This undermines attempts to describe the approach to equilibrium as a process of increasing entropy. This problem is the starting point of a research programme that aims to revise GSM in such a way that the Gibbs entropy increases over time. Coarse-graining combined with a mixing dynamics, interventionism, and attempts to redefine Gibbsian equilibrium in way that avoids reference to stationary distributions are but the most prominent proposals in that programme.[20]

For those who regard GSM as an effective theory this research programme gets started on the wrong foot. If the Gibbs entropy does not change over time, we should conclude that GSM does not offer an effective description of non-equilibrium processes and limit its range of applicability to equilibrium situations rather than trying to turn GSM into a correct description of non-equilibrium processes. Such a programme would be justified only if it turned GSM into an effective theory of non-equilibrium processes. But at least so far this does not seem to have happened. Non-equilibrium versions of GSM are not effective non-equilibrium theories. Not only do they not offer manageable algorithms to compute outcomes (thereby violating Wells' first criterion); they often also are not empirically adequate (spin echo experiments are a case in point).

Foundational questions concerning GSM remain important when they concern the

[20]For a review and discussion of these proposals see Frigg's (2008), Sklar's (1993) and Uffink's (2007).

empirical adequacy of the theory or its connection to the fundamental theory, BSM. One such question is the one we addressed in the previous section, namely under what circumstances Gibbsian and Boltzmannian equilibrium values coincide. Another is the problem of the justification of maximum entropy methods. The choice of the outcome distribution is often guided by maximum entropy considerations, and there is a legitimate question why these considerations work, and why they output distributions that provide correct equilibrium values; see Uffink's (1995, 1996) for a discussion. But these are questions concerning the instrumental efficiency of the theory and are not aimed at turning GSM into a fundamental theory.

Acknowledgments

We would like to thank Valia Alori for inviting us to participate in this project. Thanks also to an anonymous referee for helpful comments. In the process of researching this paper we benefited from discussions with Jeremy Butterfield, David Lavis, Stephan Hartmann, Wayne Myrvold, Patricia Palacios, Jos Uffink and Giovanni Valente.

References

Albert, David 2000. Time and chance, Cambridge, MA, London: Harvard University Press.

Baxter, Rodney J. 1982. Exactly Solved Models in Statistical Mechanics, London: Academic Press.

Boltzmann, Ludwig 1877. "Über die Beziehung zwischen dem zweiten Hauptsatze der mechanischen Wäermetheorie und der Warhscheinlichkeitsrechnung resp. den Sätzen über des Wärmegleichgewicht," Wiener Berichte 76, 373–435.

Chandler, David 1987. Introduction to Modern Statistical Mechanics, New York: Oxford University Press.

Davey, Kevin. 2009. "What Is Gibbs's Canonical Distribution?" Philosophy of Science 76 (5), 970–783.

Dizadji-Bahmani, Foad 2011. "The Aharonov Approach to Equilibrium," Philosophy of Science 78 (5): 976–988.

Ehrenfest Paul and Tatjana Ehrenfest-Afanassjewa 1959. The Conceptual Foundations of the Statistical Approach in Mechanics, Ithaka: Cornell University Press.

Fermi, Enrico. 2000. Thermodynamics, Mineola: Dover.

Frigg, Roman 2008. "A Field Guide to Recent Work on the Foundations of Statistical Mechanics," in Rickles, ed. The Ashgate Companion to Contemporary Philosophy of Physics, London: Ashgate, 99–196.

— 2009. "Typicality and the Approach to Equilibrium in Boltzmannian Statistical Mechanics," Philosophy of Science 76, 2009, 997-1008.

— 2010, "Why Typicality Does Not Explain the Approach to Equilibrium," in Mauricio Suárez (ed.) Probabilities, Causes and Propensities in Physics. Synthese Library. Dordrecht: Springer, 2010, 77-93.

— 2010. "Probability in Boltzmannian Statistical Mechanics," in: Ernst and Hüttemann, eds. Time, Chance and Reduction. Philosophical Aspects of Statistical Mechanics. Cambridge: Cambridge University Press, 92–118.

Frigg, Roman and Charlotte Werndl 2011. "Demystifying Typicality, " Philosophy of Science 79(5), 2012, 917-929.

— 2019. "Can Somebody Please Say What Gibbsian Statistical Mechanics Says?" forthcoming in The British Journal for Philosophy of Science.

Gibbs, Josiah W. 1902. Elementary Principles in Statistical Mechanics, New York: Scribner.

Goldstein, Sheldon (2001), Boltzmann's Approach to Statistical Mechanics, in: In Jean Bricmont, Detlef Dürr, Maria Carla Galavotti, Gian Carlo Ghirardi, Francesco Petruccione and Nino Zanghì (eds.), Chance in Physics: Foundations and Perspectives (pp. 39-54). Berlin and New York: Springer, 39-54.

Isihara, A. 1971. Statistical Physics, Cambridge MA: Academic Press.

Khinchin, Alexandr. I. 1949. Mathematical Foundations of Statistical Mechanics. Mineola/New York: Dover Publications.

Landau, Lev D. and Evgeny M. Lifshitz (1980). Statistical Physics, Oxford: Pergamon Press.

Lavis, David 2005. "Boltzmann and Gibbs: An Attempted Reconciliation," in Studies in History and Philosophy of Modern Physics 36: 245–73.

— and Bell, G.M. 1999. *Statistical Mechanics of Lattice Systems, Volume 1: Closed Form and Exact Solutions*. Ellis Horwood.

Lebowitz, Joel L. 1993a, "Boltzmann's entropy and time's arrow," Physics Today, September Issue,32-38.

Lebowitz, Joel L. 1993b, "Macroscopic laws, microscopic dynamics, time's arrow and Boltzmann's entropy," Physica A 194, 1-27.

Lieb, Elliott H. and Yngvason, Jakob (1999). The Physics and Mathematics of the Second Law of Thermodynamics. *Physics Reports* 310, 1-96.

MacDonald, Donald K.C 1962. Noise and Fluctuations. An Introduction, New York: Wiley.

Malament, David B. and Zabell, Sandy .L. 1980. "Why Gibbs Phase Averages Work. The Role of ErgodicTheory," in Philosophy of Science 56: 339–49.

Meester, Ronald 2003. *A Natural Introduction to Probability Theory*. Springer.

Myrvold, Wayne C. 2016. "Probabilities in Statistical Mechanics," in Hájek and Hitchcock, eds. The Oxford Handbook of Probability and Philosophy, Oxford University Press, 573–600.

Pathria, Raj K. and Paul D. Beale 2011. Statistical Mechanics, Oxford: Elsevier.

Sklar, Lawrence. 1993. Physics and Chance. Philosophical Issues in the Foundations of Statistical Mechanics, Cambridge: Cambridge University Press.

Tolman, Richard C. 1938. The Principles of Statistical Mechanics, Mineola, NY: Dover Publications.

Uffink, Jos. 1995. "Can the Maximum Entropy Principle be Explained as a Consistency Requirement?" Studies in History and Philosophy of Modern Physics 26 (3), 223–261.

— 1996. "The Constraint Rule of the Maximum Entropy Principle," Studies in History and Philosophy of Modern Physics 27 (1), 47–79.

— 2007. "Compendium of the Foundations of Classical Statistical Physics," in Butterfield and Earman, eds. Philosophy of Physics, Amsterdam: North Holland, 923–1047.

Volchan, Sérgio B. 2007, "Probability as Typicality," Studies in History and Philosophy of Modern Physics 38, 801-814

Vranas, Peter B. 1998. "Epsilon-ergodicity and the Success of Equilibrium Statistical Mechanics." Philosophy of Science 65, 688–708.

Wallace, David 2015. "The Quantitative Content of Statistical Mechanics". Studies in History and Philosophy of Modern Physics 52, 285–293.

Wang, Genmiao, Edie Sevick, Emil Mittag, Debra, J., Searles, and Denis J. Evans, 2002. "Experimental Demonstration of Violations of the Second Law of Thermodynamics for Small Systems and Short Time Scales," in Physical Review Letters 89, 050601.

Wells, James D. 2012. Effective Theories in Physics. From Planetary Orbits to Elementary Particle Masses. Heidelberg and New York: Springer.

Werndl, Charlotte and Roman Frigg 2015. "Reconceptionalising Equilibrium in Boltzmannian Statistical Mechanics," Studies in History and Philosophy of Modern Physics 49, 19–31.

— 2017a. "Boltzmannian Equilibrium in Stochastic Systems," in Massimi and Romeijn, eds. Proceedings of the EPSA15 Conference. Berlin and New York: Springer, 243–254.

— 2017b "Mind the Gap: Boltzmannian versus Gibbsian Equilibrium", Philosophy of Science 84(5), 2017, 1289-1302.

— 2019a. "When Does Gibbsian Phase Averaging Work?" Manuscript.

— 2019b. "When Does a Boltzmannian Equilibrium Exist?," forthcoming in Bendingham, Maroney, and Timpson, eds. Quantum Foundations of Statistical Mechanics, Oxford University Press.

— 2019c. "Ehrenfest and Ehrenfest-Afanassjewa on Why Boltzmannian and Gibbsian Calculations Agree," forthcoming in Uffink, Valente, Werndl and Zuchowski, eds. Tatjana Afanjassewa and Her Legacy: Philosophical Insights from the Work of an Original Physicist and Mathematician, Berlin: Springer.

Wilhelm, Isaac (fortcoming): "Typical," forthcoming in The British Journal for Philosophy of Science.