Studies in History and Philosophy of Modern Physics 49 (2015) 19–31

Contents lists available at ScienceDirect

Studies in History and Philosophy of Modern Physics

journal homepage: www.elsevier.com/locate/shpsb

Reconceptualising equilibrium in Boltzmannian statistical mechanics and characterising its existence

Charlotte Werndl [a,b,*], Roman Frigg [a]

[a] Department of Philosophy, London School of Economics, Houghton Street, London WC2A 2AE, United Kingdom
[b] Department of Philosophy, University of Salzburg, Franziskanergasse 1, 5020 Salzburg, Austria

ARTICLE INFO

Article history:
Received 27 August 2014
Received in revised form
29 October 2014
Accepted 22 December 2014
Available online 30 January 2015

Keywords:
Equilibrium
Approach to equilibrium
Boltzmann
Statistical mechanics
Ergodicity
Macro-states

ABSTRACT

In Boltzmannian statistical mechanics macro-states supervene on micro-states. This leads to a partitioning of the state space of a system into regions of macroscopically indistinguishable micro-states. The largest of these regions is singled out as the equilibrium region of the system. What justifies this association? We review currently available answers to this question and find them wanting both for conceptual and for technical reasons. We propose a new conception of equilibrium and prove a mathematical theorem which establishes in full generality – i.e. without making any assumptions about the system's dynamics or the nature of the interactions between its components – that the equilibrium macro-region is the largest macro-region. We then turn to the question of the approach to equilibrium, of which there exists no satisfactory general answer so far. In our account, this question is replaced by the question when an equilibrium state exists. We prove another – again fully general – theorem providing necessary and sufficient conditions for the existence of an equilibrium state. This theorem changes the way in which the question of the approach to equilibrium should be discussed: rather than launching a search for a crucial factor (such as ergodicity or typicality), the focus should be on finding triplets of macro-variables, dynamical conditions, and effective state spaces that satisfy the conditions of the theorem.

When citing this paper, please use the full journal title *Studies in History and Philosophy of Modern Physics*

1. Introduction

The core posit of Boltzmannian statistical mechanics (BSM) is that macro-states supervene on micro-states. This leads to a partitioning of the state space of a system into regions of macroscopically indistinguishable micro-states, where by 'macroscopically indistinguishable' we mean indistinguishable with respect to macroscopic properties such as thermodynamic properties. These regions are called *macro-regions*. The largest of these macro-regions is commonly singled out as the system's equilibrium region. What justifies the association of equilibrium with the macro-state corresponding to the largest macro-region?

After briefly introducing the main elements of BSM (Section 2) and illustrating them with three examples, we scrutinise common answers that have been given to this question. We find these wanting both for conceptual and for technical reasons (Section 3). This prompts the search for an alternative answer. This answer cannot be found by revising any of the received approaches, and so we propose a new definition of equilibrium. While previous approaches sought to define equilibrium in terms of micro-mechanical properties, our definition is modelled on the thermodynamic conception of equilibrium, and also incorporates what has become known as the 'minus first law' of thermodynamics (TD) (Section 4).

The new conception of equilibrium is not only free from the conceptual and technical difficulties of earlier notions, but it also provides the spring-board for a general answer to our initial problem. We prove a mathematical theorem which establishes in full generality that the equilibrium macro-region is the largest macro-region (in a requisite sense). The proof is mathematically rigorous and the theorem is completely general in that it makes *no* assumptions either about the system's dynamics or the nature of the interactions between the system's components (Section 5).

We then turn to the question of the approach to equilibrium, to which there exists no satisfactory general answer. In our account, this question is replaced by the question: under what circumstances

* Corresponding author.

http://dx.doi.org/10.1016/j.shpsb.2014.12.002
1355-2198/

does an equilibrium state exists? We point out that for an equilibrium to exist three factors need to cooperate: the choice of macro-variables, the dynamics of the system, and the choice of the effective state space. We then prove a theorem providing fully general necessary and sufficient conditions for the existence of an equilibrium state. This theorem changes the way in which the problem of the approach to equilibrium should be discussed: rather than launching a search for one crucial factor (such as ergodicity or typicality), the focus should be on finding triplets of macro-variables, dynamical conditions, and effective state spaces that satisfy the conditions of the theorem. This gives the discussion of equilibrium a new direction (Section 6).

2. Boltzmannian statistical mechanics

We begin with a brief summary of the apparatus of BSM. This is mainly to introduce notation and state a few crucial results; for detailed introductions to BSM we refer the reader to Frigg (2008) and Uffink (2007). We then introduce three examples that will guide us through our discussion and serve as illustrations of the general claims we make in later sections. The reliance on three different examples is not owed to a preference for abundance. Discussions of BSM have all too often been distorted, and indeed misled, by an all too narrow focus on the dilute gas. Contrasting the dilute gas (our first example) with the Baker's gas and the Kac-ring (our second and third examples) widens the focus and helps illustrate the general claims we make in later sections.

2.1. The framework of Boltzmannian statistical mechanics

A system in statistical mechanics has the mathematical structure of a *measure-preserving deterministic dynamical system* $(X, \Sigma_X, \mu_X, T_t)$. X is the set representing all possible *micro-states*; Σ_X is a σ-algebra of subsets of X; the *evolution function* $T_t : X \to X$, $t \in \mathbb{R}$ (continuous time) or $\mathbb{Z}$ (discrete time), is a measurable function in (t, x) such that $T_{t_1+t_2}(x) = T_{t_2}(T_{t_1}(x))$ for all $x \in X$ and all $t_1, t_2 \in \mathbb{R}$ or $\mathbb{Z}$; μ_X is a measure on Σ_X that it is invariant under the dynamics: $\mu_X(T_t(A)) = \mu_X(A)$ for all $A \in \Sigma_X$ and all t.[1] The *solution* through x, $x \in X$, is the function $s_x : \mathbb{R} \to X$ or $s_x : \mathbb{Z} \to X$, $s_x(t) = T_t(x)$.

At the macro level the system is characterised by a set of *macro-variables* $\{v_1, \ldots, v_l\}$ $(l \in \mathbb{N})$. These variables are measurable functions $v_i : X \to \mathbb{V}_i$, associating a value with each point in X. We use capital letters V_i to denote the values of v_i. A particular set of values $\{V_1, \ldots, V_l\}$ defines a *macro-state* $M_{V_1,\ldots,V_l}$. We only write 'M' rather than '$M_{V_1,\ldots,V_l}$' if the specific values V_i do not matter. For now all we need is the general definition of macro-variables. We will discuss them in more detail in Section 6.1, where we will see that the choice of a set of macro-variables is a subtle matter of considerable importance and that the nature and even existence of an equilibrium state crucially depends on it.

The central philosophical posit of BSM is *supervenience*: macro-states supervene on micro-states. This implies that a system's micro-state uniquely determines its macro-state. This determination relation will be many-to-one. For this reason every macro-state M is associated with a macro-region consisting of all micro-states for which the system is in M. An important yet often neglected issue is on what space macro-regions are defined. The obvious option would be X, but often this is not what happens. In fact, in many cases macro-regions are defined on a subspace $Z \subseteq X$. Intuitively speaking, Z is a subset whose states evolve into the same equilibrium macro-state. In the case of a dilute gas with N particles, for instance, X is the $6N$-dimensional space of all position and momenta, while Z is the $6N-1$ dimensional energy hypersurface. We call X the *full state space* and Z the *effective state space* of the system. The *macro-region* Z_M corresponding to *macro-state* M *relative to* Z can then be defined as the set of all $x \in Z$ for which M supervenes on x. A set of macro-states relative to Z is complete iff (if and only if) it contains all states of Z. The members of a complete set of macro-regions Z_M form a partition of Z (i.e. the Z_M do not overlap and jointly cover Z).

The correct choice of Z depends on the system under investigation, and has to be determined on a case-by-case basis. We return to this point in Section 6.1. There is one general constraint on such a choice, though, that needs to be mentioned now. Since a system can never leave the partition of macro-regions, Z must be mapped onto itself under T_t. Then the sigma algebra can be restricted to Z and one considers a measure on Z which is invariant under the dynamics and where the measure is normalized, i.e. $\mu_Z(Z) = 1$.[2] In this way one obtains the measure-preserving dynamical system $(Z, \Sigma_Z, \mu_Z, T_t)$ with a normalized measure μ_Z. $(Z, \Sigma_Z, \mu_Z, T_t)$ is called the *effective system* (as opposed to the *full system* $(X, \Sigma_X, \mu_Z, T_t)$).

The *Boltzmann entropy* of a *macro-state* M relative to Z is $S_B(M) := k_B \log[\mu_Z(Z_M)]$ (k_B is the Boltzmann constant). The Boltzmann entropy of a *system* at time t, $S_B(t)$, is the entropy of the macro-state the system is in at t relative to Z: $S_B(t) := S_B(M_{x(t)})$, where $x(t)$ is the system's micro-state at t and $M_{x(t)}$ is the macro-state supervening on $x(t)$.

One of the macro-regions is singled out as corresponding to the *equilibrium state* of the system relative to Z. A crucial aspect of the standard presentation of BSM is that equilibrium corresponds to the largest macro-region (measured in terms of μ_Z). In fact, this is often used as a criterion to define equilibrium: the equilibrium state relative to Z is simply the one that is associated with the largest macro-region. Since the logarithm is a monotonic function, the equilibrium state is also the one with the largest Boltzmann entropy.

2.2. Example 1: the dilute gas

Consider a system consisting of N particles in a finite container isolated from the environment. The *micro-state* of the system is specified by a point $x = (q, p)$ in the $6N$-dimensional set of possible position and momentum coordinates Γ. So Γ is the X of the gas. The dynamics of the system is determined by its classical Hamiltonian $H(x)$. Energy is preserved and therefore the motion is confined to the $6N-1$ dimensional energy hypersurface Γ_E defined by $H(x) = E$, where E is the energy value. So Γ_E is the Z of the gas. The solutions of the equations of motion are given by the flow T_t on Γ_E, where $T_t(x)$ is the state into which $x \in \Gamma_E$ evolves after time t has elapsed. Σ_E is the standard Lebesgue-σ-algebra. Γ is endowed with the Lebesgue measure λ, which is preserved under T_t. A measure μ_E on Γ_E can be defined which is preserved as well and is normalised, i.e. $\mu_E(\Gamma_E) = 1$ (cf. Frigg, 2008, p. 104). $(\Gamma_E, \Sigma_E, \mu_E, T_t)$ is the effective measure-preserving dynamical system of the gas.

The macro-states usually considered arise as follows: the state of one particle is determined by a point in its 6-dimensional state space γ, and the state of system of N identical particles is determined by N points in this space. Since the system is confined to a finite container and has constant energy E, only a finite part of γ is accessible. One then partitions the accessible part of γ into cells of equal size $\delta\omega$ whose dividing lines run parallel to the position and momentum axes. The result is a finite partition

[1] At this point the measure of X is allowed to be infinite (hence there is no requirement that the measure is normalized).

[2] The dynamics is given by the evolution equations restricted to Z, and we follow the dynamical systems literature in denoting it again by T_t.