Boltzmann and Gibbs: An attempted reconciliation ☆

D.A. Lavis

Department of Mathematics, King's College, Strand, London WC2R 2LS, UK

Abstract

There are three levels of description in classical statistical mechanics, the microscopic/dynamic, the macroscopic/statistical and the thermodynamic. At one end there is a well-used concept of equilibrium in thermodynamics and at the other dynamic equilibrium does not exist in measure-preserving reversible dynamic systems. Statistical mechanics attempts to situate equilibrium at the macroscopic level in the Boltzmann approach and at the statistical level in the Gibbs approach. The aim of this work is to propose a reconciliation between these approaches and to do so we need to reconsider the concept of equilibrium. Our proposal is that the binary property of the system being or not being in equilibrium is replaced by a continuous property of *commonness*.
© 2005 Elsevier Ltd. All rights reserved.

Keywords: Boltzmann; Gibbs; Approach to equilibrium; Ergodic decomposition

1. Introduction

One of the fundamental problems in the foundations of statistical mechanics is to give an explanation as to why 'equilibrium' statistical mechanics is so successful. That is to say, why the use of the standard Gibbsian methods 'reproduces'

---

thermodynamic results. One offered explanation for this is the standard ergodic approach. As to whether this gives an acceptable explanation Van Lith (2001) offers the impression "that the *communis opinio* in the physics literature is that it does; in the philosophy literature that it doesn't". My impression agrees with hers. However, there is a further twist to the story. When confronted with the question of what is 'actually going on' in a gas of particles (say) when it is in equilibrium, or when it is coming to equilibrium, many physicists are quite prepared to desert the Gibbsian approach entirely and to embrace a Boltzmannian view (Ruelle, 1991; Lebowitz, 1993; Bricmont, 1995; Goldstein, 2001). In particular according to Goldstein:

> "Ludwig Boltzmann explained how irreversible macroscopic laws... originate in the time-reversible laws of microscopic physics. Boltzmann's analysis... is basically correct. The most famous criticisms of Boltzmann's later work on the subject have little merit. Most twentieth century innovations – such as the identification of the state of a physical system with a probability distribution ρ on phase space, of its thermodynamic entropy with the Gibbs entropy of ρ, and the invocation of the notions of ergodicity and mixing for the justification of statistical mechanics – are thoroughly misguided" (Goldstein, 2001, p. 39).

and Lebowitz:

> "Having results for typical microstates rather than averages is not just a mathematical nicety but at the heart of understanding the microscopic origin of observed macroscopic behaviour. We neither have nor do we need ensembles... . What we do need and can expect is typical behaviour" (Lebowitz, 1993, p. 38).

These assertions are reinforced by the opinion of Ruelle[1] that the Boltzmannian approach

> "is now generally accepted by physicists. ... There are some dissenting voices, such as that of Ilya Prigogine, but the disagreement is based on philosophical prejudice rather than physical evidence" (Ruelle, 1991, p. 113).

However, most work in equilibrium statistical mechanics uses the tools developed by Gibbs. Given a particular thermodynamic setup and microscopic model the appropriate probability distribution (microcanonical, canonical, grand-canonical, etc.) is chosen. The entropy is taken to be that of Gibbs and the holy grail for any investigation is an analytic form for the partition function; the notable successes being the solution of the zero-field two-dimensional Ising model by Onsager (1944), of the six-vertex model in 1967 by Lieb (see, Lieb & Wu, 1972) and of the eight-vertex model in 1972 by Baxter (see, Baxter, 1982). There have been many attempts to extend the Gibbs approach to non-equilibrium. As indicated above in the quote

---

[1]I think Ruelle rather overstates the case. In particular the rational subjectivist approach of Jaynes (1983) and the interventionist approach most recently argued for by Ridderbos and Redhead (1998) and Ridderbos (2002) would find some favour.

from Ruelle, the most developed programme for doing this is that of the Brussels–Austin School of the late Count Ilya Prigogine.[2] However, many remain unconvinced of either the actual or potential success of this enterprise. There would seem to be the need to explore the possibility of holding, with Lebowitz, Goldstein, Ruelle et al., to the conviction that Boltzmann got it right about the approach to equilibrium, whilst at the same time with a good conscience continuing to use the standard distributions of Gibbs for everyday equilibrium calculations. We shall attempt to take some steps along this path. In order to do so we need to resolve in some way three questions, to which the current versions of the Gibbs and Boltzmann approaches offer apparently irreconcilable answers.

  (i) What is meant by equilibrium?
 (ii) What is statistical mechanical entropy?
(iii) What is the object of study?

The attempt to produce conciliatory answers to (i) and (ii) will occupy most of this paper. However, we must at the outset deal with (iii). Ensembles are an intrinsic feature of the approach of Prigogine for whom "it is at the level of ensembles that temporal evolution can be predicted" (Prigogine, 1994, p. 8) and for Mackey (1989) a "thermodynamic system *is* [my italics] a system that has, at any given time, states distributed throughout phase space ..., and the distribution of these states is characterized by a density ..."(p. 984). The density referred to is the ensemble density and thus the ensemble becomes the way that a thermodynamic system is *defined*. In contrast to this we shall follow the view of Lebowitz given above that the object of study in statistical mechanics is a *single system*. All talk of 'ensembles' is taken to be just a way of giving a relative frequency flavour to the probabilities of events occurring in that system.

  In Section 2 we define the dynamic microstructure of our system and introduce two gas models which will be used as examples in the later parts of the paper. In Section 3 we describe the different concepts of equilibrium in dynamics, the Boltzmann and Gibbs approaches to statistical mechanics and thermodynamics. In Section 4 we discuss the Boltzmann approach with particular reference to the way in which equilibrium is defined and the Neo–Boltzmannian reliance on the concept of typical behaviour. With respect to the latter we suggest that the expected typical behaviour can be encapsulated by the term 'thermodynamic-like', which signifies the type of fluctuations to be expected in dependent thermodynamic variables, most particularly the entropy, and we discuss the conditions on the system needed for thermodynamic-like behaviour to be typical. At that point it is necessary to consider the usual demarcation made between the system being or not being in equilibrium and we discuss, with the use of the example of the baker's gas with a small number of microsystems, the problematic nature of this property. This leads us to the rather

---

[2]Although the fundamentals of the work of this school have remained unchanged since their inception in the early sixties, the approach has evolved substantially, leading to a great wealth of publications. For a comprehensive review see Bishop (2004).

radical suggestion that it would be better to replace the binary equilibrium property with a continuously changing property which we call *commonness*. In Section 5 we use the example of a gas expanding in a box when a partition is removed to describe the role ascribed to the Gibbs approach in the overall picture we are attempting to develop. Of course, one question which needs to be addressed is the meaning of probability and this is discussed in Section 6. Section 7 contains our proposals for a model incorporating both the Boltzmann and Gibbs viewpoints and our conclusions are presented in Section 8.

2. The microstructure

Consider a system that, at time t, has a microstate given by the vector $\boldsymbol{x}(t)$ in the *phase space* Γ. The time parameter t can be discrete or continuous and the phase space can also be continuous or discrete. Some one-to-one autonomous dynamics $\boldsymbol{x} \rightarrow \phi_t \boldsymbol{x}$, $(t \geqslant 0)$ determines a *flow* in Γ and the set of points $\boldsymbol{x}(t) = \phi_t \boldsymbol{x}(0)$, parameterized by $t \geqslant 0$, gives a *trajectory*. The set of mappings $\{\phi_t\}_{t \geqslant 0}$ is a semi-group. The system is *reversible* if there exists an idempotent operator $\mathfrak{I}$ on the points of Γ, such that $\phi_t \boldsymbol{x} = \boldsymbol{x}'$ implies that $\phi_t \mathfrak{I} \boldsymbol{x}' = \mathfrak{I} \boldsymbol{x}$. Then $\phi_{-t} = (\phi_t)^{-1} = \mathfrak{I} \phi_t \mathfrak{I}$ and the set $\{\phi_t\}$ with $t \in \mathbb{R}$ or $\mathbb{Z}$ is a group. Henceforth we shall assume that the system under discussion is reversible. If Γ consists of a finite number of points then $t \in \mathbb{Z}$ and the system is periodic. If Γ is continuous then $t \in \mathbb{Z}$ or $t \in \mathbb{R}$. In this case, we shall restrict attention to a subset $\Lambda \subseteq \Gamma$ invariant under $\{\phi_t\}$. We denote by $\mathfrak{m}$ a measure on Λ such that (a) $\mathfrak{m}(\Lambda)$ is finite, (b) $\mathfrak{m}$ is absolutely continuous with respect to the Lebesque measure on Λ, (c) $\mathfrak{m}$ is preserved by the flow; that is $\mathfrak{m}(\phi_t \gamma) = \mathfrak{m}(\gamma)$, for all t and all measurable $\gamma \subset \Lambda$. It can now be shown that the Poincaré (1890) recurrence theorem will apply (Ott, 1993, p. 214) to flows for which $\boldsymbol{x}(0) \in \Lambda$.

2.1. Useful examples

To clarify the discussion it is useful to resort to computer simulations of simple models. However, most interesting problems in statistical mechanics concern cooperative systems and, even at equilibrium (see e.g. Baxter, 1982), there are few of these which can be solved exactly. So, of necessity, useful examples are of assemblies of non-interacting microsystems and the literature contains discussions of many 'toy models' of this kind.[3] Here we shall use two.

2.1.1. A perfect gas—continuous time and continuous space
Consider a gas of N particles of unit mass moving in the two-dimensional box $-L/2 \leqslant x \leqslant L/2$, $-L/2 \leqslant y \leqslant L/2$, with elastic walls. We suppose that the initial

---

[3] A model which has featured prominently in such discussions is the spin-echo system. For a recent account of this model see Lavis (2004) and for simulations for a number of deterministic and stochastic systems see Lavis (2003).

positions $(x^{(k)}(0), y^{(k)}(0))$, $k = 1, 2, \ldots, N$ of the particles are such that all the $y^{(k)}(0)$ are distinct and the initial velocities are all in the x-direction. Then all the particles will perform one-dimensional oscillations in the x-direction at constant speeds. If we denote the positions and velocities at time t by $(x^{(k)}(t), v^{(k)}(t))$, $k = 1, 2, \ldots, N$ then $|v^{(k)}(t)|$ is constant and the system is reversible with $\Im(x^{(k)}, v^{(k)}) = (x^{(k)}, -v^{(k)})$. In Fig. 1 we show a typical evolution of the gas from a state where all the particles are congregated near the left-hand end of the box. Calculations for this example are used to illustrate the discussion in Section 7.
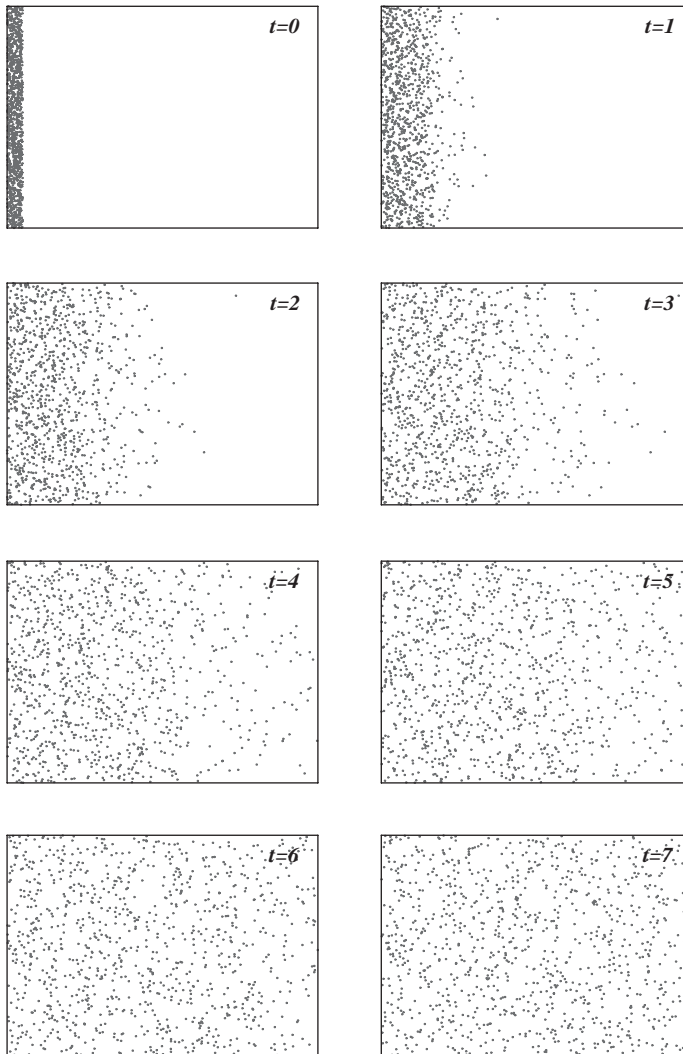

Fig. 1. A perfect gas of $N = 1000$ particles moving horizontally in a box with elastic walls.
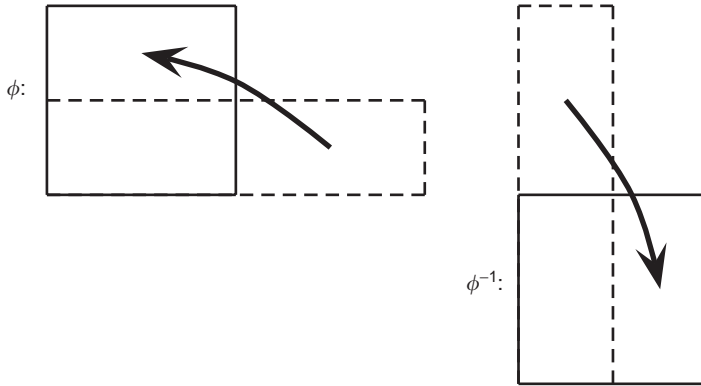
Fig. 2. The baker's transformation and its inverse.

2.1.2. A baker's gas—discrete time and continuous space

This is the transformation, shown in Fig. 2, where a square of side L is stretched to twice its width and then cut in half with the right-hand half used to restore the upper half of the unit square. As the mapping ϕ on the Cartesian coordinates $L \times (x, y)$ of the square, it is given by

$$\phi(x, y) = \begin{cases} (2x, \tfrac{1}{2} y), & \mod 1, \quad 0 \leqslant x \leqslant \tfrac{1}{2}, \\ (2x, \tfrac{1}{2}[y + 1]), & \mod 1, \quad \tfrac{1}{2} \leqslant x \leqslant 1. \end{cases} \tag{1}$$

A convenient way of writing this transformation is to express x and y as binary strings: $x = 0 \cdot x_1 x_2 x_3 \ldots$ and $y = 0 \cdot y_1 y_2 y_3 \ldots$ with $x_k, y_k = 0, 1$. Then the transformation takes the form $\phi(0 \cdot x_1 x_2 x_3 \ldots, 0 \cdot y_1 y_2 y_3 \ldots) = (0 \cdot x_2 x_3 x_4 \ldots, 0 \cdot x_1 y_1 y_2 \ldots)$, with $\phi^{-1}(0 \cdot x_1 x_2 x_3 \ldots, 0 \cdot y_1 y_2 y_3 \ldots) = (0 \cdot y_1 x_1 x_2 \ldots, 0 \cdot y_2 y_3 y_4 \ldots)$. The mapping is reversible with $\phi^{-1} = \mathfrak{I} \phi \mathfrak{I}$ and $\mathfrak{I}(x, y) = (y, x)$. It can also be shown (Lasota & Mackey, 1994, p. 54–56) to be volume-preserving and thus that the Poincaré (1890) recurrence theorem applies. We are interested in a baker's gas of N 'particles'. In Fig. 3 we show a typical evolution from a state where all the particles are congregated near the bottom left corner of the box.

3. Different views of equilibrium

The three levels of description in classical statistical mechanics, referred to in the abstract, each bring with them a definition of equilibrium. In fact by separating the macroscopic (Boltzmann) and statistical (Gibbs) approaches one can identify four distinct concepts of equilibrium.
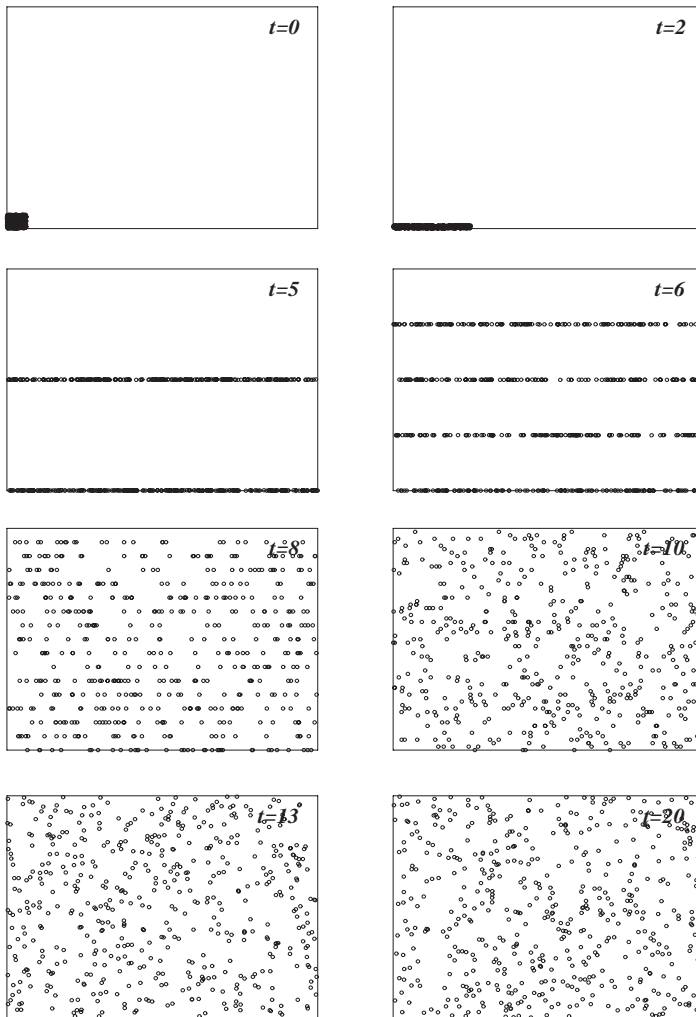
Fig. 3. A gas of $N = 512$ particles moving under the baker's transformation.

A *dynamic system* is in equilibrium at time t if its state is given by a phase point $x(t)$ lying on an attractor.[4] However, of course, measure-preserving reversible dynamic systems do not have attractors.[5]

---

[4]For our purpose this is an adequate definition, although the term is sometimes used only for fixed points, where its use includes repelling (unstable) and marginal equilibrium points.

[5]This means that the remark by Sklar (1993, p. 156) that in statistical mechanics "the equilibrium state exists as the 'attractor' to which the dynamics of non-equilibrium drives systems ..." must not be taken too literally.

To describe the *Boltzmann approach* we need to introduce some *macroscopic variables*. In general, these will be at the observational level for the system, but will encapsulate more detail than the thermodynamic variables.[6]Examples of these would be local variables which quantified spatial inhomogeneities of density or magnetization in a system (see Section 4).[7]Suppose that, as in our examples discussed below, the system consists of N identical microsystems and that we have a set Ξ of macrovariables. A macrostate μ is a $\mathfrak{m}$-measurable subset of Λ. The set of macrostates $\{\mu\}$ is defined so that:

(i) Every $x \in \Lambda$ is in exactly one macrostate denoted by μ_x.
(ii) Each macrostate corresponds to a unique set of values for Ξ.
(iii) μ_x is invariant under all permutations of the microsystems.
(iv) The phase points x and $\mathfrak{I}x$ are in macrostates of the same size.[8]

Given a particular set of macrovariables and macrostates, the *Boltzmann Entropy* is given by

$$S_{\mathrm{B}}(x) = k_{\mathrm{B}} \ln[\mathfrak{m}(\mu_x)]. \tag{2}$$

Equilibrium is defined by reference to the macrostate, which is uniquely given by a set of values of the macrovariables. According to Bricmont (1995, p. 179), "by far the largest volumes correspond to the *equilibrium values* of the macroscopic variables (and this is how 'equilibrium' should be defined)". This is the standard Boltzmannian definition. The system is in equilibrium if its phase point is in the macrostate of largest volume ($\mathfrak{m}$-measure), corresponding to the largest value of S_{B}, and this in turn defines the equilibrium values of the macrovariables.

The starting point for the *Gibbs approach* is to suppose the phase-point x, in some invariant $\Lambda \subset \Gamma$, is distributed according to a probability density function[9] ρ which is invariant under the flow (a solution of Liouville's equation). Equilibrium is defined as the situation where the probability density function is not an explicit function of time and ρ becomes a function of the global constants of motion. The statistical mechanical 'analogues'[10] of thermodynamic quantities are either fixed external parameters, related to phase functions[11] or functionals of ρ. In particular the

---

[6]Ridderbos (2002) refers to them as *supra-thermodynamic variables*.

[7]Another simple example, in the context of coin-flipping, is given by Bricmont (2001, p. 8). In N tosses of a coin a microstate is a particular record of the N outcomes and a macrostate is identified by the macrovariable N_0 and consists of the set of all microstates for which exactly N_0 of the outcomes are heads.

[8]This is necessary to ensure that the macrostate structure reflects the reversibility of the system; which means that the Boltzmann entropy profile along $x \to \phi_t x$ is the same as that along $\mathfrak{I}x \to \mathfrak{I}\phi_t x = \phi_{-t}\mathfrak{I}x$. In fact, of course, in cases like the perfect gas where macrostates are created by course-graining the one-microsystem configuration space and reversibility corresponds to changing the sign of the velocity, x and $\mathfrak{I}x$ are in the same macrostate.

[9]So that the probability of the phase point being in a small region $\delta\Lambda$ around a phase point x is $\rho(x)\mathfrak{m}(\delta\Lambda)$. The meaning we give to probability is discussed in more detail in Section 6.

[10]This is the term used by Gibbs (1902, Chapter 14).

[11]In a way which will be discussed in Section 7.

analogue of thermodynamic entropy is the *Gibbs entropy*

$$S_{\mathrm{G}}[\rho] = -k_{\mathrm{B}} \int_{\Lambda} \rho(\boldsymbol{x}) \ln[\rho(\boldsymbol{x})] \, \mathrm{d}\mathfrak{m}. \tag{3}$$

From a practical point of view this scheme is very satisfactory. However, problems arise when an attempt is made to extend it to non-equilibrium situations, which are now perceived as being represented by time-dependent solutions of Liouville's equation. Specifically:

(i) When $\rho(\boldsymbol{x})$ is replaced in (3) by any time-dependent solution $\rho(\boldsymbol{x};t)$ of Liouville's equation, $S_{\mathrm{G}}[\rho(t)]$ remains invariant with respect to time.
(ii) Given an arbitrary initial condition $\rho(\boldsymbol{x};0)$, the evolving solution $\rho(\boldsymbol{x};t)$ of Liouville's equation will not in general converge to a time-independent (equilibrium) solution as $t \to \infty$.

The Brussels–Austin programme is, as we have indicated above, an attempt to circumvent these problems.

In *classical thermodynamics* the approach to equilibrium is usually described as leaving the system so that it "eventually reaches a state in which no further change is perceptible, no matter how long one waits" (Pippard, 1961, p. 6). The equilibrium state is thus the situation in which (for an isolated system) no perceptible change occurs in any thermodynamic variables. However, Uffink (2001) has argued that the tendency of isolated systems to approach equilibrium is not a consequence of the standard laws of thermodynamics. This has led Brown and Uffink (2001, p. 528) to formulate a principle which they call the *minus first law* that "an isolated system in an arbitrary initial state within a finite fixed volume will spontaneously attain a unique state of equilibrium". With this addition, thermodynamic equilibrium, for a finite isolated system, can be seen to have three important properties:

(a) It is a binary property; a system either is or is not in equilibrium.
(b) A system in equilibrium never evolves away from equilibrium.
(c) A system not in equilibrium evolves into equilibrium.

Indeed the quote from Pippard may seem to allow us to add 'in a finite time' to (c) or at the very least that 'so that the difference from equilibrium after a finite time is imperceptible'.

In this description of four distinct types of equilibrium, dynamic equilibrium has been included for completeness. Although it shares properties (a)–(c) with thermodynamic equilibrium it exists only for dissipative systems. Statistical mechanical equilibrium for the systems considered here, for which the dynamics is non-dissipative, cannot be a consequence of, or related to, the system being in dynamic equilibrium. The Brussels–Austin extension of the Gibbs programme would also satisfy properties (a)–(c). However, as will be discussed in Section 4, the form of equilibrium which features in the Boltzmann approach does not satisfy (b) and satisfies (c) only in a qualified sense (see Section 4).

A classic problem in statistical mechanics (Sklar, 1993, Chapter 9; Callender, 1999) is to understand in what sense thermodynamics can be said to *reduce* to statistical mechanics and thus part of the problem of equilibrium is to relate in some way equilibrium in statistical mechanics and equilibrium in thermodynamics. However, as we have seen there is no unique statistical mechanical equilibrium. The Boltzmann and Gibbs approaches have very different concepts. Indeed, with the usual ensemble interpretation of the probability density function, Gibbsian equilibrium is, unlike Boltzmannian equilibrium, not a property of an individual system. So for us the problem is both to effect a reconciliation between Boltzmannian and Gibbsian equilibrium and to relate them to thermodynamic equilibrium. To do this we propose that equilibrium as a binary property is replaced by something (which we call commonness) which encapsulates degrees of 'equilibriumness'.

4. The Boltzmann approach

Most examples of the use of this approach consider situations where the macrovariables measure inhomogeneities in the distribution of microsystems over the one-microsystem phase space. This space is divided into p cells, labelled $\ell = 1, 2, \ldots, p$, and, for the phase point $x \in \Lambda$, $N_\ell(x)$ is the number of microsystems in cell $\ell = 1, 2, \ldots, p$. Given that the microsystems are identical, the macrostate $\mu_x \in \Lambda$ corresponds to all permutations of microsystems in cells with the same values of the macrovariables. Thus

$$\frac{\mathfrak{m}(\mu_x)}{\mathfrak{m}(\Lambda)} = \frac{N!}{\prod_{\ell=1}^{p}\{N_\ell(x)\}!} \frac{1}{p^N}. \tag{4}$$

For the baker's gas Λ is the union of N squares of side L giving $\mathfrak{m}(\Lambda) = L^{2N}$ and we divide the one-particle phase space (configuration space in this case) into squares of side $L/2^m$ to give $p = 2^{2m}$.[12] In the case of the perfect gas $\mathfrak{m}(\Lambda) = L^N$ and we take $p = 2^m$.

As indicated above, the entropy in the Boltzmann approach is the phase function (2), which will fluctuate with time, the system being construed as being in equilibrium when entropy has its maximum value $(S_B)_{Max}$. The difficulties associated with this definition of equilibrium are discussed in Section 4.1. However, another question arises in relation to this approach. This is usually identified as the problem of the 'approach to equilibrium'.[13] However, it can be extended to the broader problem associated with certain expectations about the form of the entropy profile and the need to give a dynamic account which justifies these expectations. In somewhat

---

[12]The cells are labelled sequentially in rows, so that cell (1,1) in the bottom left corner is labelled $\ell = 1$ and cell $(2^m, 2^m)$ in the top right corner is labelled $\ell = p$.

[13]And part of the discussion is concerned with giving some explanation as to why the universe seems to have started in a very uncommon state. According to Goldstein (2001) this is the "hard part" of the Boltzmann approach. It is beyond the scope of the present discussion.

imprecise language, we expect that in most cases the entropy will behave in a *thermodynamic-like* way. This we take to mean that:

The Boltzmann entropy, for the evolving system, is most of the time close to its maximum value, from which it exhibits frequent small fluctuations and rare large (downward) fluctuations.

The problem of quantifying and justifying these expectations is discussed in Section 4.2.

4.1. Trying to define equilibrium

It is clear that $(S_B)_{\text{Max}}$ will not necessarily be the value of entropy for the largest proportion of microstates. Associated with a macrostate μ, there will be a degeneracy $\Omega(\mu)$, giving the number of macrostates of measure $\mathfrak{m}(\mu)$.[14] We define

$$\Upsilon(\mu) = \frac{\mathfrak{m}(\mu)\Omega(\mu)}{\mathfrak{m}(\Lambda)} \quad \text{where } \mathfrak{m}(\Lambda) = \sum_{\{\mu\}} \mathfrak{m}(\mu) \tag{5}$$

and $(S_B)_\Upsilon$ denotes the Boltzmann entropy for a macrostate giving the largest value of $\Upsilon(\mu)$. With

$$S_\Lambda = S_B(\Lambda) = k_B \ln[\mathfrak{m}(\Lambda)], \tag{6}$$

the *entropy of* Λ, it is clear that the inequalities $S_\Lambda > (S_B)_{\text{Max}} \geqslant (S_B)_\Upsilon$ must be satisfied[15] and, from (4),

$$\frac{S_\Lambda - (S_B)_{\text{Max}}}{Nk_B} \leqslant \frac{S_\Lambda - (S_B)_\Upsilon}{Nk_B} < -\frac{1}{N} \sum_{\{\mu\}} \frac{\mathfrak{m}(\mu)}{\mathfrak{m}(\Lambda)} \ln\left[\frac{\mathfrak{m}(\mu)}{\mathfrak{m}(\Lambda)}\right] < \frac{\ln[\eta(N,p)]}{N}, \tag{7}$$

where

$$\eta(N,p) = \frac{(N+p-1)!}{N!(p-1)!}. \tag{8}$$

In Fig. 4 we show the Boltzmann entropy of the baker's gas for various cases of different numbers of particles and cells, where initially the gas starts with all the particles in cell (1,1).[16] We observe a rapid rise of the scaled entropy from its initial value of zero followed by small fluctuations around a value rather less than the scaled value of unity.[17] It is clear that, in all the cases shown in Fig. 4, $\Upsilon(\mu)$ is not maximal for the largest macrostate, although detailed computations with large values of N are not only difficult, but not particularly revealing.

It is more useful to concentrate on small values of N and p where the results are easy to compute. With $p = 4$ cells ($m = 1$) and $N = 8$ we denote by the vector

---

[14]In the case of our examples $\Omega(\mu)$ will equal the number of distinct permutations of $\{N_1, \ldots, N_p\}$.

[15]Assuming that the Boltzmann entropy is additively scaled so as to be non-negative.

[16]In all graphs of the Boltzmann entropy and Table 1, the minimum entropy is set to zero with the condition $p = V$. Additionally in Fig. 4 the entropy is scaled by its maximum value.

[17]This behaviour is, of course, typical of a wide range of models with both discrete and continuous phase spaces (Lavis, 2003).
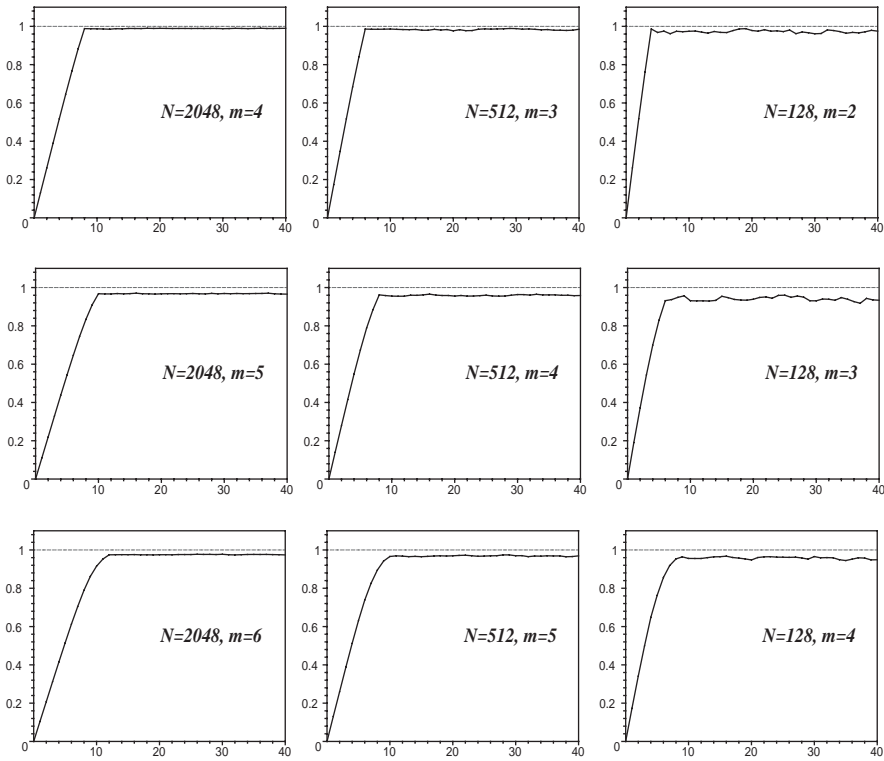
Fig. 4. The scaled Boltzmann entropy as it evolves with time for the baker's gas model. Initially all the particle are randomly distributed in cell (1,1), ($N_1 = N$). The figures in horizontal lines correspond (from top to bottom) to average numbers of particles per cell of 8, 2 and 0.5, respectively.

$n = (n_1, n_2, n_3, n_4)$ an unordered distribution of particles in the four cells, with $\Omega(n)$ being the number of macrostates with this distribution. Data for this case are given in order of decreasing $\Upsilon(n)$ in Table 1. It will be observed that the macrostate of maximum size $n = (2, 2, 2, 2)$ is the eighth in the list. Values derived from Table 1 are $S_A = Nk_B\, 1.3863$, $(S_B)_{Max} = Nk_B\, 0.9790$ and $(S_B)_\Upsilon = Nk_B\, 0.9283$. A typical evolution of the entropy for this case is shown in Fig. 5. The 'plateau' at $(S_B)_\Upsilon$ will be observed.

So if we were to propose a definition of the equilibrium macrostate what would we choose? If we take a quasi-thermodynamic view and suppose that the equilibrium value for the entropy is that into which it settles, with subsequent small fluctuations (and very rare large fluctuations) then for small systems we have shown by simulation that this is not the largest entropy, for which the system would be in the largest macrostate. An obvious strategy would be to broaden the range of entropy for equilibrium. One might, for example, say that the system is in equilibrium if $S_B \in [(S_B)_{Max}, (S_B)_{Max}(1 - \varepsilon)]$, where $\varepsilon = 2[(S_B)_{Max} - (S_B)_\Upsilon]/(S_B)_{Max}$. A similar kind of '$\varepsilon$-equilibrium' strategy, in the context of the Gibbs approach, was proposed by

Table 1
Data for the baker's gas with $N = 8$ and $p = 4$

| $\boldsymbol{n}$	$\Omega(\boldsymbol{n})$	$\mathfrak{m}(\boldsymbol{n})/\mathfrak{m}(\Lambda)$	$S_B(\boldsymbol{n})/(Nk_B)$	$\Upsilon(\boldsymbol{n})$
$(3,2,2,1)$	12	0.0256	0.9283	0.3076
$(4,2,1,1)$	12	0.0128	0.8417	0.1538
$(3,3,1,1)$	6	0.01709	0.8776	0.1025
$(3,3,2,0)$	12	0.0085	0.7910	0.1025
$(4,3,1,0)$	24	0.0043	0.7043	0.1025
$(4,2,2,0)$	12	0.0064	0.7550	0.0769
$(5,2,1,0)$	24	0.0026	0.7271	0.0615
$(2,2,2,2)$	1	0.0385	0.9790	0.0385
$(5,1,1,1)$	4	0.0051	0.7271	0.0205
$(5,3,0,0)$	12	0.0009	0.5032	0.0103
$(6,1,1,0)$	12	0.0009	0.5032	0.0103
$(4,4,0,0)$	6	0.0011	0.5311	0.0064
$(6,2,0,0)$	12	0.0004	0.4165	0.0052
$(7,1,0,0)$	12	0.0001	0.2599	0.0014
$(8,0,0,0)$	4	0.15×10^{-4}	0	0.61×10^{-4}
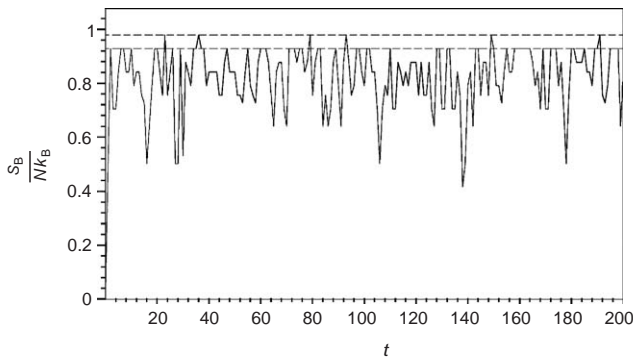
Fig. 5. The scaled Boltzmann entropy for the baker's gas model with $N = 8$ particles and $p = 4$. Initially all the particle are randomly distributed in cell (1,1), with configuration (8,0,0,0) and degeneracy of four. Horizontal lines at $(S_B)_{\text{Max}}$ and $(S_B)_\Upsilon$ are shown.

Van Lith (1999). In our case, of course, this means that there will be small fluctuations of entropy within equilibrium and the possibility of evolution out of equilibrium accompanied by a large downward fluctuation in entropy. In addition to the arbitrary division into macrostates, we have added a demarcation given by the choice of value for ε. This is a simple demonstration of our contention that the quality which we are trying to capture is a matter of degree, rather than the two-valued property of either being *in equilibrium* or *not in equilibrium*. We, therefore, make the following proposal:

> All references to a system being, or not being, in equilibrium should be replaced by references to the commonness of the state of the system, with this property

being given by some suitably-scaled monotonically increasing function of the Boltzmann entropy.

Of course, at this stage, the proposal is confined to the Boltzmann approach. However, we hope to show that its implementation is of relevance to the development of a view of statistical mechanics which incorporates both the Boltzmann and Gibbs approaches and their connection to thermodynamics. It should be emphasized that no connection is implied, at this stage, between commonness and 'commonness of occurrence'. We are not assuming that a microstate with large commonness (a common microstate) belongs to a macrostate which is visited more often by an evolving system. This connection is related to the question of *typical behaviour*.

4.2. Thermodynamic-like behaviour

A favoured term of the Neo–Boltzmannians (but not of Boltzmann himself) is 'typical'. Thus, in the quote from Lebowitz (1993) in Section 1, he refers to 'typical microstates' and 'typical behaviour' and, in a later paper (Lebowitz, 1999), he makes the assertion that

"S_B will *typically* increase in a way which *explains* and describes qualitatively the evolution towards equilibrium of macroscopic systems. Typical, as used here, means that the set of microstates corresponding to a given macrostate [μ] for which the evolution leads to a macroscopic decrease in the Boltzmann entropy during some fixed time period τ, occupies a subset of [μ] whose [measure] is a fraction of [$\mathfrak{m}(\mu)$] which goes very rapidly (exponentially) to zero as the number of [particles] in the system increases" (Lebowitz, 1999, p. S348).[18]

Now in ordinary usage 'typical' is a description applied to an element of a set with respect to some particular property.[19] In most instances of the use of 'typical' by the Neo–Boltzmannians the element of interest is a microstate (phase point) picked from phase space (or an energy surface in phase space). The property is 'imminent evolution towards equilibrium indicated by an increase in Boltzmann entropy' or perhaps more precisely 'the absence of a imminent large decrease in Boltzmann entropy'.

Lebowitz ties this behaviour to the structure of macrostates. This connection could be amplified in the following way. For the macrostate μ let $\mu^{(++)}$ consist of all $x \in \mu$ such that the immediately next and immediately previous macrostates along the trajectory are of measure larger than μ.[20] With $\mu^{(--)}$ and $\mu^{(\pm\mp)}$ defined in a similar way, it follows, from the reversibility of the system and property (iv) of the macrostates listed in Section 3, that $\mathfrak{m}(\mu^{(-+)}) = \mathfrak{m}(\mu^{(+-)})$. The Neo–Boltzmannian

---

[18]We have used square brackets to replace Lebowitz's notation by our own.

[19]Thus $a \in \mathscr{A}$ is typical of $\mathscr{A}$ with respect to the property b if a and most members of $\mathscr{A}$ have the property b. Or alternatively b is typical of the set $\mathscr{A}$ if a randomly chosen member of $\mathscr{A}$ is very likely to have the property b.

[20]For easy of discussion we ignore the possibility of contiguous macrostates of equal measure.

argument is now based on the supposition that

$$\mathfrak{m}(\mu^{(++)}) \gg \mathfrak{m}(\mu^{(\pm\mp)}) \gg \mathfrak{m}(\mu^{(--)}),\tag{9}$$

for all (or most) macrostates μ, with the inequality becoming increasingly pronounced as N increases. Thus, for an arbitrarily chosen macrostate, the measure $\mathfrak{m}(\mu^{(++)} \cup \mu^{(-+)})$ of phase points which lead to an increase in entropy is very much larger than the measure $\mathfrak{m}(\mu^{(+-)} \cup \mu^{(--)})$ of phase points which lead to a decrease in entropy. Then a typical (randomly chosen) point in μ belongs to $\mu^{(++)} \cup \mu^{(-+)}$, leading to an increase in entropy. This is justified by Bricmont (2001) using Bayesian arguments. No explanation is needed for the typical increase in entropy, beyond the application of the law of large numbers in the limit of large N. If typical behaviour differs from this then an explanation must be sought (maybe a re-examination of the designation of macrostates). Within its own terms this Bayesian gloss of typicality is perfectly satisfactory. However, it is predicated on the validity of (9). While this appears to be true for many toy models[21] general rules for its validity are more difficult to formulate.

But most statistical mechanical explanations are concerned with something more than just the immediate increase in entropy following the arbitrary selection of an initial phase point. The object of interest is the evolution along a trajectory and the Bayesian argument does not adequately account for this. Consider the following scenario. For a system a phase point x is chosen. It happens to be in the macrostate μ and according to good Bayesian (or Laplacian) arguments and inequality (9) it will typically lie in $\mu^{(++)} \cup \mu^{(-+)}$. Now suppose, in the course of evolution, the contiguous transition $x \to x'$ occurs to a macrostate μ'. By definition $\mathfrak{m}(\mu') > \mathfrak{m}(\mu)$ and $x' \in \mu'^{(-+)} \cup \mu'^{(--)}$. Now if x' were chosen randomly in $\mu'^{(-+)} \cup \mu'^{(--)}$ we might typically expect, from (9), that $x' \in \mu'^{(-+)}$, leading to a further increase in entropy. However, this is not necessarily the case. The dynamics determines the location of x' in $\mu'^{(-+)} \cup \mu'^{(--)}$. Merely to associate probabilities with measure without random selection is not adequate to close the argument.

The requirement of statistical mechanics is that *a typical trajectory is thermodynamic-like*, according to the definition of this term in Section 4. As yet this definition lacks precision. However, it is certainly the case that if, along a trajectory and for most macrostates μ, the phase point spent an amount of time in μ proportionally related to $\mathfrak{m}(\mu)$ the requirement of thermodynamic-like behaviour would be satisfied. To be more precise, for any $x \in \Lambda$, let $\mathscr{L}_x$ be the trajectory through x, and $\mathsf{T}_x(\mu)$ be the proportion of the time that the phase point on this trajectory is in μ.[22] Now we can define *thermodynamic-like behaviour along the trajectory $\mathscr{L}_x$* as the situation where $\mathsf{T}_x(\mu)$ is approximately proportional to $\mathfrak{m}(\mu)$. The more common a macrostate is the more common will be its occurrence along a

---

[21]It is easy to verify for particular cases of the Kac ring model, which is the example most often cited by Bricmont.

[22]It was shown by Birkhoff (1931) that $\mathsf{T}_x(\mu)$ exists and is independent of the location of x on $\mathscr{L}_x$ for almost all $x \in \Lambda$; that is except possibly for a set of $\mathfrak{m}$-measure zero. From this it follows (see e.g. Lavis, 1977) that $\mathsf{T}_x(\mu)$ is a constant of motion almost everywhere in Λ.

trajectory. For this behaviour to be *typical* it must be true for most trajectories. This will hold in a strong sense if

$$\mathsf{T}(\mu) = \tau(\mu) \quad \text{where } \tau(\mu) = \mathfrak{m}(\mu)/\mathfrak{m}(\varLambda) \tag{10}$$

for all μ and all trajectories apart from a set of measure zero. This is, of course, the condition that the system is *ergodic*. So ergodicity is *sufficient* for (perfect) thermodynamic-like behaviour along all but an atypical set of trajectories of measure zero.

The baker's transformation is a Kolmogorov system (Lasota & Mackey, 1994, p. 91) and is, therefore, both ergodic and mixing.[23] So we have an explanation for the typical thermodynamic-like behaviour exhibited in Figs. 4 and 5. It is also not difficult to describe at least some of the measure-zero atypical trajectories. Suppose that the initial binary strings giving the particle positions were periodic and the periods of all the particles were commensurate. Then the behaviour of the system (detected, for example, by its entropy profile) would have a periodic form at variance with thermodynamic-like behaviour. The perfect gas of Section 2.1.1 is also ergodic (but not mixing) as long as the particle speeds are incommensurate (Farquhar, 1964, pp. 96–98)[24] and the atypical trajectories correspond to a choice of initial speeds which are commensurate and produce periodic, rather than quasi-periodic behaviour.

However, ergodicity is not *necessary*[25] since (10) is stronger than is required for a typical system to behave in a thermodynamic-like way and atypicality can also be broadened to include a set of trajectories of small but non-zero measure. As pointed out by Bricmont (1995, 2001) and verified by numerical calculations (Lavis, 2003) the Kac ring model, which is not ergodic,[26] exhibits clear thermodynamic-like behaviour.

The typicality of thermodynamic-like behaviour, away from the perfect case of an ergodic system, is difficult to formulate. What counts as sufficiently close to the perfect situation of (10) to yield an entropy profile which looks thermodynamic-like and how much atypical behaviour can be tolerated? For the purpose of this work we shall leave these questions open with symbols of approximation standing as markers for a more detailed investigation.[27]

---

[23]In fact it is a Bernoulli system (Arnold & Avez, 1968, pp. 125–126).

[24]The proof of this given by Farquhar, which is a version of the discussion by Khinchin (1949, pp. 58–62), is for two independent particles with periodic boundary conditions and ergodicity is over the reduced manifold defined by the normal integrals of the motion. Our model maps into one with periodic boundary conditions and particle velocities to the right by 'unfolding' the interval $[-L/2, L/2]$ into $[-L, L]$ (Lavis, 2004). Then it can be seen that the only independent normal integrals of the motion are $\dot{x}^{(k)} = v^{(k)}$, $k = 1, 2, \ldots, N$ and the motion is ergodic over the hypercube $x^{(k)} \in [-L, L]$, $k = 1, 2, \ldots, N$.

[25]Bricmont (1995) also argues that ergodicity is not sufficient, but this is based on the use of ergodicity to justify measurements as infinite time-averages.

[26]Its phase space divides into cycles.

[27]Quantification of what counts as thermodynamic-like and then what count as a tolerable degree of atypical non-thermodynamic-like behaviour could both be developed using the sort of ε criteria described above for ε-equilibrium.

Ergodicity on Λ is, of course, equivalent to Λ being metrically indecomposable under the flow, and a non-ergodic situation where we can at least map out the elements of the discussion is when there is a finite or denumerable ergodic decomposition of Λ. That is, to within subsets of zero measure,

$$\Lambda = \bigcup_{\{\kappa\}} \Lambda_\kappa, \tag{11}$$

where each component Λ_κ is invariant and metrically indecomposable under the flow. Ergodic decomposition will, of course, apply to any model (like, for example, the Kac ring) where Λ consists of a finite number of discrete points; then each Λ_κ is a cycle of states. Ergodic decomposition also applies where the system is ε-ergodic (Vranas, 1998). In this case one element of the decomposition Λ_1 (say) is such that $\mathfrak{m}(\Lambda_1) = (1 - \varepsilon)\mathfrak{m}(\Lambda)$ for small $\varepsilon > 0$.[28]

When (11) applies, the time spent in the macrostate μ for all trajectories $\mathscr{L}_x$ with $x \in \Lambda_\kappa$ (apart from at set of measure zero) is

$$\mathsf{T}_\kappa(\mu) = \mathfrak{m}(\mu \cap \Lambda_\kappa)/\mathfrak{m}(\Lambda_\kappa). \tag{12}$$

For such a system it is clear that temporal behaviour will be the same for all trajectories[29] within a Λ_κ. The division must be between those members of the ergodic decomposition in which the behaviour is thermodynamic-like and those in which it is not. So:

(T1) We indicate that the system is behaving in Λ_κ in a way which can be counted as thermodynamic-like by writing

$$\mathsf{T}_\kappa(\mu) \cong \tau(\mu), \quad \forall \mu. \tag{13}$$

(T2) We denote by $\Lambda^{(\mathrm{T})}$ the union of all Λ_κ satisfying (13) with $\Lambda^{(\mathrm{A})} = \Lambda \backslash \Lambda^{(\mathrm{T})}$ and by p_κ the probability[30] that the system is in Λ_κ. Then thermodynamic-like behaviour is typical for the system if

$$\sum_{\Lambda_\kappa \subset \Lambda^{(\mathrm{A})}} \mathsf{p}_\kappa \ll 1. \tag{14}$$

5. The Gibbs approach

Given that we are concerned to incorporate both the Boltzmannian and Gibbsian view of statistical mechanics into our overall picture, we must now examine the role to be ascribed to the Gibbs approach. This we do by means of a simple example.

---

[28]Vranas (1998) has argued that many dynamic systems of interest are ε-ergodic, but as yet the hard evidence for this is lacking.

[29]At all relevant places in the following discussion the phrase: 'except possibly a set of measure zero', is taken to apply.

[30]The meaning of which is discussed in Section 6.

Consider a gas of N particles moving in d dimensions. Then the phase vector $\boldsymbol{x} \in \Gamma$ is $2dN$-dimensional, composed of configuration and momentum vectors $\boldsymbol{q}^{(k)} = (q_1^{(k)}, \ldots, q_d^{(k)})$ and $\boldsymbol{p}^{(k)} = (p_1^{(k)}, \ldots, p_d^{(k)})$, $k = 1, 2, \ldots, N$. Let Λ correspond to the gas being confined to the hypercubic box $\mathscr{B} = \{\boldsymbol{x} | -\frac{1}{2} L \leqslant q_\alpha^{(k)} \leqslant \frac{1}{2} L, \ \forall k, \forall \alpha\}$. Then, if it is left undisturbed 'a sufficient time to attain thermodynamic equilibrium', the phase-point $\boldsymbol{x}$ will, according to the Gibbs prescription, be distributed in Λ according to the appropriated equilibrium probability density function, which we denote by $\rho_G(\boldsymbol{x})$.

Now suppose that the gas is confined by a partition to the part of the box, denoted by $\mathscr{B}^{(-)}$, with $q_1^{(k)} < 0$, $\forall k$. In this situation, if the system is left to attain thermodynamic equilibrium, the appropriate probability density function will be $\rho_G^{(-)}(\boldsymbol{x})$, which differs from $\rho_G(\boldsymbol{x})$ only in respect of the restriction on the configuration space. If the partition is removed at time $t = 0$ the probability density function is

$$\rho(\boldsymbol{x}; 0) = \begin{cases} \rho_G^{(-)}(\boldsymbol{x}), & \boldsymbol{x} \in \Lambda^{(-)}, \\ 0 & \text{otherwise,} \end{cases} \tag{15}$$

where $\Lambda^{(-)}$ is that part of Λ corresponding to all the particles being in $\mathscr{B}^{(-)}$. This is no longer the equilibrium distribution; it will be the initial condition for a non-equilibrium solution $\rho(\boldsymbol{x}; t)$ evolving according to Liouville's equation. However, as we have indicated, the Gibbs entropy for this solution remains constant and $\rho(\boldsymbol{x}; t)$ does not converge to $\rho_G(\boldsymbol{x})$ either in finite time, as we would like for thermodynamics, nor even on an infinite time scale. The most we can obtain is, for a mixing system, when the expectation values of a certain class of phase functions calculated using $\rho(\boldsymbol{x}; t)$ converge, as $t \to \infty$, to their expectation values calculated with $\rho_G(\boldsymbol{x})$. The resolution to this problem suggested by Gibbs (1902, p. 148) (see also Ehrenfest & Ehrenfest-Afanassjewa, 1912) was to coarse-grain the phase-space in the manner in which macrostates were obtained in the Boltzmann approach. In the macrostate μ the probability density function $\rho(\boldsymbol{x})$ is replaced by $\pi(\mu)/\mathfrak{m}(\mu)$ where

$$\pi(\mu) = \int_\mu \rho(\boldsymbol{x}) \, \mathrm{d}\mathfrak{m}, \tag{16}$$

is the probability of the phase point $\boldsymbol{x}$ being in μ. Then from (3)

$$(S_G)_{CG}[\pi] = -k_B \sum_{\{\mu\}} \pi(\mu) \ln[\pi(\mu)] + k_B \sum_{\{\mu\}} \pi(\mu) \ln[\mathfrak{m}(\mu)]. \tag{17}$$

The objections to this approach are well-known[31] and we shall not discuss them here. We shall suggest a different approach.

Suppose that a phase point $\boldsymbol{x}_0 \in \Lambda^{(-)}$. What is the probability of the system being in a small measurable set $\delta\Lambda_0$ around $\boldsymbol{x}_0$? This will clearly depend on the physical circumstances of the system and will differ according to whether the phase point is confined to $\Lambda^{(-)}$ by the physical partition. In that case, if the system is in 'thermodynamic equilibrium' the probability will be $\rho_G^{(-)}(\boldsymbol{x}_0)\mathfrak{m}(\delta\Lambda_0)$. However, if the

---

[31]See Ridderbos (2002) for a recent discussion and Ridderbos and Redhead (1998) and Lavis (2004) for the application of course-graining to the spin-echo system.

partition is not present and the system is in 'thermodynamic equilibrium' the probability will be $\rho_G(\boldsymbol{x}_0)\mathfrak{m}(\delta\Lambda_0)$. Since $\boldsymbol{x}_0$ corresponds to all the particles of the gas being in one end of the box, it will (in our Boltzmann language) be a rather 'uncommon state' and we expect that $\rho_G(\boldsymbol{x}_0)\mathfrak{m}(\delta\Lambda_0)\ll\rho_G^{(-)}(\boldsymbol{x}_0)\mathfrak{m}(\delta\Lambda_0)$. Now consider the case where at time $t=0$ the partition is removed. According to the Gibbs prescription, the probability of $\boldsymbol{x}\in\delta\Lambda_0$ is $\rho(\boldsymbol{x}_0;0)\mathfrak{m}(\delta\Lambda_0)=\rho_G^{(-)}(\boldsymbol{x}_0)\mathfrak{m}(\delta\Lambda_0)$. The removal of the partition has not affected the probability. But compare this situation with that where there has never been a partition present and the system phase point is in $\delta\Lambda_0$. What physically distinguishes the two cases? If, in the course of its dynamic flow, the system is at $\boldsymbol{x}_0$ at $t=0$, its forward evolution will not be affected by whether a partition has just been removed or whether it simply happens to have evolved into this state. So why should a different probability distribution be assigned to these two situations when $t>0$?[32] A consistent approach, consonant with our treatment of the Boltzmann approach, is to suppose that the only meaningful probability density function to be used from the Gibbs approach is the time-independent solution of Liouville's equation determined by the dynamics and the physical constraints on the system. A change of physical constraints will lead to an instant discontinuous change in the probability density function and the Gibbs entropy. An uncommon state (like, for example, the case of all the particles being in one end of the box) will have low probability when calculated using $\rho_G(\boldsymbol{x})$ and low Boltzmann entropy, but the same Gibbs entropy as any other configuration.

6. Probability

In the Gibbs approach of Section 5 we introduced the probability density function $\rho(\boldsymbol{x})$ on the invariant set Λ and in terms of this, in (16), defined the probability $\pi(\mu)$ that the phase point will be in the macrostate μ. What we have not done is define what we mean by probability. Within the Gibbs approach this is most frequently done using ensembles, which means that the probability density is not the property of a single system. But we have already indicated, in Section 1, that our object of interest is a single system. Our aim in this work is to bring some kind of reconciliation between the Gibbs and Boltzmann approaches and, as asserted by Lebowitz (1993) (see the quotation in Section 1), this latter neither has nor needs ensembles.

We need two probabilities, that the phase point is moving on a particular trajectory and that on that trajectory the phase point lies in a particular subset of Λ. For the latter we shall follow Von Plato (1989) in using the *time-average* definition of probability. This analysis is simplified if, as we shall do, the ergodic decomposition (11) is assumed. Then:

(i) The probability $\mathsf{p}_\kappa=\mathsf{Prob}(\boldsymbol{x}\in\Lambda_\kappa)$ was introduced in Section 4.2. It can be taken to mean either the probability that when we choose the initial point $\boldsymbol{x}(0)$ it is in

---

[32]Of course, the trajectories for $t<0$ will differ and the system from which the partition was removed can be regarded as suffering from 'false memory syndrome'.

Λ_κ or the probability that when we investigate the system we find the phase point in Λ_κ. In either case the probability is susceptible to a Bayesian or Laplacian interpretation.

(ii) The conditional probability that $x \in \mu$ given that $x \in \Lambda_\kappa$ is, from (12),

$$\mathsf{Prob}(x \in \mu | x \in \Lambda_\kappa) = \mathsf{T}_\kappa(\mu) = \frac{\mathfrak{m}(\mu \cap \Lambda_\kappa)}{\mathfrak{m}(\Lambda_\kappa)}, \tag{18}$$

where the first equality represents the time-average definition of probability.

Then

$$\mathsf{Prob}(x \in \mu) = \sum_{\{\kappa\}} \mathsf{Prob}(x \in \mu | x \in \Lambda_\kappa)\,\mathsf{p}_\kappa = \langle \mathsf{Prob}(x \in \mu | x \in \Lambda_\kappa)\rangle_{\mathrm{ED}}, \tag{19}$$

where $\langle\cdot\rangle_{\mathrm{ED}}$ denotes the expection value over the ergodic decomposition.

In each ergodic component there is a unique probability density function invariant under the flow given by

$$\rho_\kappa(x) = \begin{cases} 1/\mathfrak{m}(\Lambda_\kappa), & x \in \Lambda_\kappa, \\ 0 & \text{otherwise} \end{cases} \tag{20}$$

and

$$\rho(x) = \sum_{\{\kappa\}} \rho_\kappa(x)\,\mathsf{p}_\kappa = \langle \rho_\kappa(x)\rangle_{\mathrm{ED}}. \tag{21}$$

Then, from (16), (20) and (21),

$$\pi(\mu) = \mathsf{Prob}(x \in \mu) = \sum_{\{\kappa\}} \frac{\mathfrak{m}(\mu \cap \Lambda_\kappa)}{\mathfrak{m}(\Lambda_\kappa)}\,\mathsf{p}_\kappa, \tag{22}$$

as given by (18) and (19).

For any functions f, integrable on Λ, and $\mathscr{G}$, summable on the macrostates, the expectation value, of f with respect to ρ and $\mathscr{G}$ with respect to π, are respectively

$$\langle f \rangle_\rho = \int_\Lambda \rho(x)f(x)\mathrm{d}\mathfrak{m}, \quad \langle \mathscr{G} \rangle_\pi = \sum_{\{\mu\}} \pi(\mu)\,\mathscr{G}(\mu). \tag{23}$$

Clearly

$$\langle f \rangle_\rho = \langle\langle f|\mu\rangle\rangle_\pi \quad \text{where } \langle f|\mu\rangle = \frac{1}{\pi(\mu)} \int_\mu \rho(x)f(x)\mathrm{d}\mathfrak{m}. \tag{24}$$

The time-average of $\mathscr{G}$ along a trajectory in Λ_κ is

$$\overline{\mathscr{G}(\mu)}^\kappa = \sum_{\{\mu\}} \mathsf{T}_\kappa(\mu)\,\mathscr{G}(\mu) \tag{25}$$

and, from (18), (19) and (22),

$$\langle \overline{\mathscr{G}(\mu)}^{\kappa} \rangle_{\text{ED}} = \sum_{\{\kappa\}} \frac{\mathsf{p}_{\kappa}}{\mathfrak{m}(\Lambda_{\kappa})} \sum_{\{\mu\}} \mathfrak{m}(\mu \cap \Lambda_{\kappa}) \mathscr{G}(\mu) = \langle \mathscr{G} \rangle_{\pi}. \tag{26}$$

When the system behaviour in Λ_{κ} is thermodynamic-like it follows, from (10) and (13), that

$$\overline{\mathscr{G}(\mu)}^{\kappa} \cong \sum_{\{\mu\}} \frac{\mathfrak{m}(\mu)}{\mathfrak{m}(\Lambda)} \mathscr{G}(\mu). \tag{27}$$

A phase function f which happens to be constant over each of the macrostates μ has $f(\boldsymbol{x}) = \langle f | \mu \rangle, \forall \boldsymbol{x} \in \mu$, giving $\langle f \rangle_{\rho} = \langle f \rangle_{\pi}$. In particular, from (22),

$$\langle S_{\text{B}} \rangle_{\rho} = \langle S_{\text{B}} \rangle_{\pi} = k_{\text{B}} \sum_{\{\kappa\}} \frac{\mathsf{p}_{\kappa}}{\mathfrak{m}(\Lambda_{\kappa})} \sum_{\{\mu\}} \mathfrak{m}(\mu \cap \Lambda_{\kappa}) \ln[\mathfrak{m}(\mu)]. \tag{28}$$

Relating the Gibbs entropy to the Boltzmann entropy will necessitate some rational choice for the probabilities $\{\mathsf{p}_{\kappa}\}$. One view of p_{κ} is that it is the probability that, if we make a random choice for the the initial system point $\boldsymbol{x}(0)$, then it lies in Λ_{κ}. If we assume that all points of Λ are equally likely then on Bayesian/Laplacian grounds and consonant with the approach of Bricmont (2001) we should choose

$$\mathsf{p}_{\kappa} = \mathfrak{m}(\Lambda_{\kappa})/\mathfrak{m}(\Lambda). \tag{29}$$

Then, from (20), (21) and (22)

$$\rho(\boldsymbol{x}) = 1/\mathfrak{m}(\Lambda), \quad \pi(\mu) = \mathfrak{m}(\mu)/\mathfrak{m}(\Lambda), \tag{30}$$

which is the microcanonical distribution. From (3) and (6),

$$S_{\text{G}} = S_{\Lambda} = k_{\text{B}} \ln[\mathfrak{m}(\Lambda)] \tag{31}$$

and, from (28),

$$\langle S_{\text{B}} \rangle_{\rho} = \langle S_{\text{B}} \rangle_{\pi} = S_{\text{G}} + k_{\text{B}} \sum_{\{\mu\}} \pi(\mu) \ln[\pi(\mu)]. \tag{32}$$

When the system behaviour in Λ_{κ} is thermodynamic-like, it follows, from (26) and (27), that

$$\overline{\mathscr{G}(\mu)}^{\kappa} \cong \langle \overline{\mathscr{G}(\mu)}^{\kappa} \rangle_{\text{ED}} = \langle \mathscr{G} \rangle_{\pi} \tag{33}$$

and, from (14), thermodynamic-like behaviour is typical if

$$\mathfrak{m} \left(\bigcup_{\Lambda_{\kappa} \subset \Lambda^{(\text{A})}} \Lambda_{\kappa} \right) \ll \mathfrak{m}(\Lambda). \tag{34}$$

7. Proposals

In this section we shall make the following assumptions:

(a) That the system has the ergodic decomposition (11).[33]
(b) That the probabilities $\{p_\kappa\}$ for the system phase point being in members of the decomposition are given by (29).
(c) That (34) applies and thus that thermodynamic-like behaviour is typical.

Assumption (a) is a purely dynamic assumption which may or may not be true for any particular system. The statistical assumption (b) is again based on the dynamics and has the same status as the assumption one might make about a die being unbiased when making a sequence of trials. Assumption (c) has a dynamic aspect but is also dependent on some quantification of whether behaviour is or is not thermodynamic-like.

As indicated in Section 1, two of the obstacles to overcome, in reconciling the Boltzmann and Gibbs approaches, are the different definitions of equilibrium and of entropy. With regard to equilibrium we propose, as indicated in Section 4.1, that the two-valued condition of the system being or not being in equilibrium is replaced by a continuous property called *commonness*, based on a designation of macrostates related to a set of macrovariables. A microstate is more or less common according to whether the macrostate in which it is situated is of greater or smaller size. The Boltzmann entropy is a measure of the size of macrostates and thus provides a measure of commonness. Gibbs methods with a time-independent probability density function are to be retained as the practical means for obtaining an 'analogue' of thermodynamics. The reason for using a time-independent probability density function to calculate thermodynamic properties is not that the system is in equilibrium but that the underlying dynamic is autonomous.[34]

The question of entropy can be subsumed into a more general account of the relationship between statistical mechanical and thermodynamic variables. As we have indicated in Section 3, the standard Gibbs perception of the relationship between statistical mechanical and thermodynamic variables is that they fall into three classes. Those for which

(C1) the corresponding statistical mechanical variable is related to a phase function.
(C2) the statistical mechanical variable and thermodynamic variable are identical and equal to an external parameter of the system.
(C3) the corresponding statistical mechanical variable is a functional of the probability density function.

The primary example of C3 is the entropy given by the Gibbs formula (3).[35] In the interests of producing a smooth relationship between the Boltzmann and Gibbs

---

[33]Of course, the analysis includes the ergodic case where the decomposition has one member with $p_1 = 1$.
[34]A non-autonomous dynamic system would not yield a time-independent solution to Liouville's equation.
[35]The class into which a variable falls depends, of course, on the environmental conditions of the system. Thus for a thermally open system the temperature is an independent parameter, whereas for a thermally closed system (like the isolated systems considered here) the temperature is a functional of ρ.

approaches we shall make two modifications to the C1–C3 scheme. The first, which is part of the standard Boltzmann approach, is to related thermodynamic entropy to a phase function (which is also a macrovariable), namely the Boltzmann entropy $S_B(\boldsymbol{x})$; so entropy is now of type C1. The second is to propose a particular relationship, in the case of C1 variables, between thermodynamic variables and phase functions.

In Section 3 we introduced the set Ξ of macrovariables defined on the macrostates $\{\mu\}$. The scheme we now propose is that a thermodynamic variable, F in class C1, is related to a phase function $f(\boldsymbol{x})$, defined on Λ, via a macroscopic function $\mathscr{F}(\mu) \in \Xi$. The statistical mechanical variable is the macrovariable $\mathscr{F}$. In the case of entropy the phase function and macrovariable are identical and equal to the Boltzmann entropy. For other thermodynamic quantities this is not the case. We are, thus, proposing the three-part scheme $f \rightsquigarrow \mathscr{F} \rightsquigarrow F$, where we need to explicate the relationships denoted by '$\rightsquigarrow$'.

$\underline{\mathscr{F} \rightsquigarrow F}$: The value $\mathscr{F}(\mu)$ is the result of a measurement of f, course-grained, to effectively give the same value throughout the macrostate μ. This perception together with the definition of macrostates and the Boltzmann entropy in Section 3 serves as the demarcation between the microscopic and macroscopic realms. Of course, this demarcation is to some extent arbitrary, but it is equally so for any macroscopic physical theory.[36] The value of the thermodynamic quantity F, along the trajectory $\mathscr{L}_{\boldsymbol{x}_0}$ passing through $\boldsymbol{x}_0$, we take to be equal to the average of the result of a large number of measurements of $\mathscr{F}$ taken at arbitrarily chosen times. So we can effectively define

$$F = \sum_{\{\mu\}} \mathsf{T}_{\boldsymbol{x}_0}(\mu) \mathscr{F}(\mu). \tag{35}$$

F is a constant of motion, but not constant on Λ except when the system is ergodic. When Λ has the ergodic decomposition (11), F has the set of values

$$F_\kappa = \sum_{\{\mu\}} \mathsf{T}_\kappa(\mu) \mathscr{F}(\mu) = \overline{\mathscr{F}(\mu)}^\kappa. \tag{36}$$

From (33), when the behaviour in Λ_κ is thermodynamic-like,

$$F_\kappa \cong \langle \mathscr{F} \rangle_\pi. \tag{37}$$

Measurement of the thermodynamic variable F will typically lead to a value close to $\langle \mathscr{F} \rangle_\pi$.

$\underline{f \rightsquigarrow \mathscr{F}}$: We now need to be more precise about what we mean by the macrovariable $\mathscr{F}$. Ideally, of course, the identification

$$\mathscr{F}(\mu) = \langle f | \mu \rangle, \tag{38}$$

---

[36]See e.g. the definition of fluid density in Landau and Lifshitz (1959, p. 1).

in the notation of (24), would be desirable, because then it follows from (37), that, for thermodynamic-like behaviour,

$$F_\kappa \cong \langle f \rangle_\rho. \tag{39}$$

This, to within the tolerance indicated by '$\cong$', is the standard identification between thermodynamic variables and the expectation values of phase functions in Gibbs theory. As we have observed in Section 6, (38) is exactly true for the Boltzmann entropy because it is, by definition, constant over each macrostate. We should not expect this to be true for all phase functions. But, of course, not all phase functions have a correspondence with thermodynamic variables. In particular we shall assume[37] $f(\boldsymbol{x})$ is (a) continuous on each macrostate and (b) invariant under permutations of the N microsystems. We have already assumed, in Section 3 condition (iii), that each macrostate is invariant under permutations of the N microsystems. So the macrostate μ can be divided into n_μ non-overlapping subsets $\lambda_\mu^{(j)}$, $j = 1, 2, \ldots, n_\mu$ each of identical $\mathfrak{m}$-measure. Representing by λ_μ a generic member of this set, it follows, from (24), that

$$\langle f | \mu \rangle = \frac{1}{\pi(\lambda_\mu)} \int_{\lambda_\mu} \rho(\boldsymbol{x}) f(\boldsymbol{x}) \, d\mathfrak{m}. \tag{40}$$

When macrostates are constructed from cells in the one-microsystem phase space, n_μ is the combinatorial factor in (4) and $\mathfrak{m}(\lambda_\mu) = \mathfrak{m}(\Lambda)/p^N$. This latter becomes small as p becomes large. In this case, since $f(\boldsymbol{x})$ is continuous over $\Lambda(\mu)$, it is likely to vary very little.[38] Although the macrostate μ may be a very large part of Λ it will be made of a large number of cells in each of which $f(\boldsymbol{x})$ has the same nearly constant value. Thus[39]

$$\mathscr{F}(\mu) \approx f(\boldsymbol{x}), \quad \text{for any point } \boldsymbol{x} \in \mu \tag{41}$$

and

$$\langle \mathscr{F} \rangle_\pi \approx \langle f \rangle_\rho, \quad \text{as the size } \mathfrak{m}(\lambda_\mu) \text{ becomes small for all } \mu. \tag{42}$$

So to recapitulate the steps of the argument:

 (i) Thermodynamic variables are time-averages of phase functions f course-grained (to produce statistical mechanical variables) over macrostates.
(ii) The thermodynamic variable is approximated along (typical) thermodynamic-like trajectories by the average of f with respect to the probability density function ρ.

---

[37] These are two of the properties assumed for a *finite-range observable* function by Lanford (1973, pp. 2–3), but, as he says, his use of the word 'observable' is not intended to have "any profound significance".

[38] One might at this stage, augment the conditions on f, given above, to the full set defining a finite-range observable (Lanford, ibid). Then, for large N, the results of Lanford and Khinchin (1949, pp. 62–69) could be used to assert that the dispersion of f around $\langle f \rangle_\rho$ is small.

[39] The symbol '$\approx$' used here to indicated approximate equality, should be distinguished from '$\cong$' which is used in this paper with the special sense of 'approximately equal in a system for which thermodynamic-like behaviour is typical'.

This establishes the relationship between a Boltzmannian view of thermodynamic variables as time averages and the Gibbsian view in which they are phase averages with respect to the appropriate probability density function.

We now consider in more detail the case of entropy. It is often said that in "equilibrium [the Gibbs entropy] agrees with Boltzmann and Clausius entropies (up to terms that are negligible when the number of particles is large) and everything is fine" (Bricmont, 1995, p. 188). Apart from the caveats we have entered in Section 4.1 with respect to the concept of equilibrium, this assertion is based on two propositions (i) that the appropriate probability density function for the Gibbs approach is given by the microcanonical distribution (30) and (ii) that for large N the 'equilibrium' macrostate effectively fills the whole of Λ. The first of these we have established, within our approach, for a system with an ergodic decomposition and with the choice of probabilities (29). In the case where macrostates are defined using the cell method described in Section 4, (ii) is not difficult to establish. Using Stirling's formula with (8),

$$\frac{\ln[\eta(N,p)]}{N} \simeq \begin{cases} \dfrac{(p-1)\ln(N)}{N} & \text{with fixed } p, \\ (1+\xi)\ln(1+\xi) - \xi\ln(\xi) & \text{with } p = \xi N, \end{cases} \quad \text{as } N \to \infty. \quad (43)$$

So in the case where p, the number of cells is fixed, it follows from (7) that

$$(S_{\mathrm{B}})_{\mathrm{Max}}/N \to S_{\mathrm{G}}/N, \quad (S_{\mathrm{B}})_{\varUpsilon}/N \to S_{\mathrm{G}}/N, \quad \text{as } N \to \infty. \quad (44)$$

This is in agreement with the results shown in the diagonal figures in Fig. 4 where we see that the value at which the entropy settles down is closer to the maximum entropy for larger values of N. In the second case $1/\xi$ is the number of microsystems per cell. With both N and p large the convergence is effected by increasing the number of microsystems per cell ($\xi \to 0$). This is illustrated by the figures in the vertical columns of Fig. 4. It also follows from (33), (32) and (7) that

$$\overline{S_{\mathrm{B}}(\mu)}^{\kappa}/N \cong \langle S_{\mathrm{B}} \rangle_{\rho}/N \to S_{\mathrm{G}}/N \quad \text{as } N \to \infty. \quad (45)$$

As an illustration of these results we take the perfect gas and suppose that it is first confined by a partition to the region $x < 0$. The Gibbs entropy, given by (6) and (31), is $S_{\mathrm{G}}^{(-)} = k_{\mathrm{B}} N \ln(L/2)$ and $S_{\mathrm{G}} = k_{\mathrm{B}} N \ln(L)$, for the respective cases where the gas is confined by the partition and where it is free to evolve over the whole box. The Boltzmann entropy in the latter case is given by (2) and (4), with $V = L$ and in the former case by the same formula but restricted by the condition that $N_\ell = 0$ for $\ell > p/2$.[40] We consider the situation in which the system starts in a minimum entropy macrostate (with all the particles in the first strip). It is allowed to evolve over the time interval [0,50] in the region $x \leqslant 0$. The partition is then removed and it is allowed to evolve in the whole box in the time interval [50,100]. This is shown in Fig. 6 where $(S_{\mathrm{G}} - (S_{\mathrm{B}})_{\mathrm{Max}})/(N k_{\mathrm{B}})$ in each range is well within the upper bound given by (7) and (43).

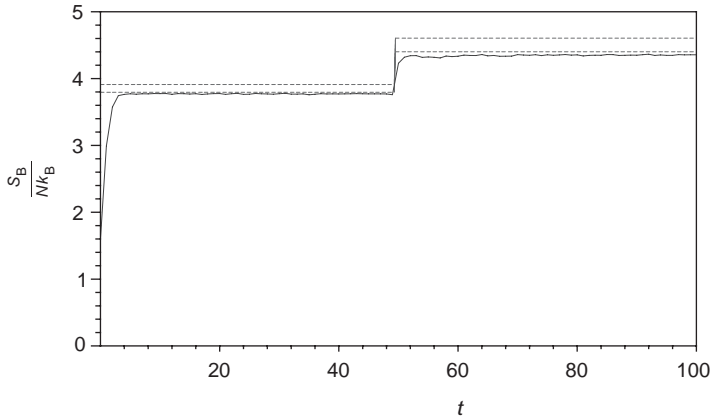[40]This expression is, of course, exactly the same as that for a box of length $L/2$ partitioned into $p/2$ strips.

Fig. 6. The scaled Boltzmann entropy for the perfect gas model with $N = 1000$, $p = 100$. The gas is allowed to evolved in the time interval $0 \leqslant t \leqslant 50$ confined to the left half of the box and then in the interval $50 \leqslant t \leqslant 100$ into the whole box. The upper and lower horizontal lines in each range correspond to S_G and $(S_B)_{Max}$ respectively.

8. Conclusions

In this programme we have used ergodicity and ergodic decomposition and as indicated in Section 1 there is deep (and justified) suspicion of the use of ergodic arguments to support the foundations of statistical mechanics, particularly among philosophers of physics. Having given a comprehensive review of the problems of ergodic arguments, Earman and Rédei (1996, p. 75) offer the opinion "that ergodic theory in its traditional form is unlikely to play more than a cameo role in whatever the final explanation of the success of equilibrium statistical mechanics turns out to be". In its 'traditional form' the ergodic argument goes something like this: (a) measurement processes on thermodynamic systems take a long time compared to the time for microscopic processes in the system and thus can be effectively regarded as infinite time averages; (b) in an ergodic system the infinite time average can be shown, for all but a set of measure zero, to be equal to the macrostate average with respect to an invariant normalized measure which is unique.[41] The traditional objections to this argument are also well-known: (i) Measurements may be regarded as time averages, but they are not *infinite* time averages. (If they were one could not, by measurement, investigate a system not in equilibrium. In fact, traditional ergodic theory does not distinguish between systems in equilibrium and not in equilibrium.) (ii) Ergodic results are all to within sets of measure zero and one cannot equate such sets with events with zero probability of occurrence. (iii) Rather few systems have been shown to be ergodic. So one must look for a reason for the success of

---

[41]In the sense that it is the only invariant normalized measure absolutely continuous with respect to the Lebesgue measure.

equilibrium statistical mechanics for non-ergodic systems and when it is found it will make the ergodicity of ergodic systems irrelevant as well.

Our use of ergodicity differs substantially from that described above and it thus escapes wholly or partly the strictures applied to it. In respect of the question of equilibrium/non-equilibrium we argue that the reason this does not arise in ergodic arguments is that equilibrium does not exist. The phase point of the system, in its passage along a trajectory, passes through common (high entropy) and uncommon (low entropy) macrostates *and that is all.* So, although in our definition of a thermodynamic variable we have extended a large finite number of measurements to an infinite set of measurements, we cannot be charged with 'blurring out' the period when the system was not in equilibrium. The charge against ergodic arguments related to sets of measure zero is applicable only if one wants to argue that the procedure always works; that is that non-thermodynamic-like behaviour never occurs. But we have, in this respect taken a Boltzmann view. We need thermodynamic-like behaviour to be typical, but we admit the possibility of atypical behaviour occurring with small but not-vanishing probability.[42] While the class of systems admitting a finite or denumerable ergodic decomposition is likely to be much larger than that of the purely ergodic systems,[43] there remains the difficult question of determining the conditions under which the temporal behaviour along a trajectory, measured in terms of visiting-times in macrostates, approximates, in most members of the ergodic decomposition, to something recognizable as thermodynamic-like behaviour.

Acknowledgements

References

Arnold, V. I., & Avez, A. (1968). *Ergodic problems of classical mechanics*. New York: Benjamin.
Baxter, R. J. (1982). *Exactly solved models in statistical mechanics*. New York: Academic Press.
Birkhoff, G. D. (1931). Proof of the ergodic theorem. *Proceedings of the national academy of sciences of the United States of America*, vol. 17 (pp. 656–660).
Bishop, R. C. (2004). Nonequilibrium statistical mechanics Brussels–Austin style. *Studies in History and Philosophy of Modern Physics*, *35*, 1–30.
Bricmont, J. (1995). Science of chaos or chaos in science? *Physicalia*, *17*, 159–208.

---

[42]For our model systems we have indicated in Section 4.2 the way in which atypical behaviour can occur and Lavis (2003) demonstrates simulations of cases of this kind.

[43]As indicated above it will include any system with a finite number of phase points.

Bricmont, J. (2001). Bayes, Boltzmann and Bohm: Probabilities in physics. In J. Bricmont, D. Dürr, M. C. Galvotti, G. Ghirardi, F. Petruccione, N. Zanghi (Eds.), *Chance in physics: Foundations and perspectives* (pp. 3–21). Berlin: Springer.

Brown, H., & Uffink, J. (2001). The origins of time asymmetry in thermodynamics: The minus first law. *Studies in History and Philosophy of Modern Physics*, *32*, 525–538.

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

Earman, J., & Rédei, M. (1996). Why ergodic theory does not explain the success of equilibrium statistical mechanics. *British Journal for the Philosophy of Science*, *47*, 63–78.

Ehrenfest, P., Ehrenfest-Afanassjewa, T. (1912). *The conceptual foundations of the statistical approach in mechanics*. Cornell University Press. (English translation).

Farquhar, I. E. (1964). *Ergodic theory in statistical mechanics*. New York: Interscience.

Gibbs, J. W. (1902). *Elementary principles in statistical mechanics*. Yale University Press. (reprinted 1960).

Goldstein, S. (2001). Boltzmann's approach to statistical mechanics, In J. Bricmont, D. Dürr, M. C. Galvotti, G. Ghirardi, F. Petruccione, N. Zanghi (Eds.), *Chance in physics: Foundations and perspectives*, (pp. 39–54). Berlin: Springer.

Jaynes, E. T. (1983). In R. D. Rosenkratz (Ed.), *Papers on probability*, *statistics and statistical physics*. Dordrecht: Reidel.

Khinchin, A. I. (1949). *Mathematical foundations of statistical mechanics*. New York: Dover.

Landau, L. D., & Lifshitz, E. M. (1959). *Fluid mechanics*. Oxford: Pergamon.

Lanford, III, O. E. (1973). Entropy and equilibrium states in classical statistical mechanics. In J. Ehlers, K. Hepp, H. A. Weidenmüller (Eds.), *Statistical mechanics and mathematical problems* (pp. 1–113). Berlin: Springer.

Lasota, A. L., & Mackey, M. C. (1994). *Chaos, fractals and noise*. Berlin: Springer.

Lavis, D. A. (1977). The role of statistical mechanics in classical physics. *British Journal for the Philosophy of Science*, *28*, 255–279.

Lavis, D. A. (2003). Some examples of simple systems: Version 2. Available at: www.mth.kcl.ac.uk/˜dlavis/papers/examples-ver2.pdf.

Lavis, D. A. (2004). The spin-echo system reconsidered. *Foundations of Physics*, *34*, 669–688 Preprint available at: arXiv.org/abs/cond-mat/0311527.

Lebowitz, J. L. (1993). Boltzmann's entropy and time's arrow. *Physics Today*, *46*, 32–38.

Lebowitz, J. L. (1999). Statistical mechanics: A selective review of two central issues. *Reviews of Modern Physics*, *71*, S346–S357.

Lieb, E. H. & Wu, F. W. (1972). Two-dimensional ferroelectric models. In C. Domb, M. S. Green (Eds.), *Phase transitions and critical phenomena*. vol. 1 (pp. 332–490). New York: Academic Press.

Mackey, M. C. (1989). The dynamic origin of increasing entropy. *Reviews of Modern Physics*, *61*, 981–1015.

Onsager, L. (1944). Crystal statistics. I: Two-dimensional model with order-disorder transition. *Physical Review*, *65*, 117–149.

Ott, E. (1993). *Chaos in dynamical systems*. Cambridge: Cambridge University Press.

Pippard, A. B. (1961). *The elements of classical thermodynamics*. Cambridge: Cambridge University Press.

Poincaré, H. (1890). On the three-body problem and the equations of dynamics. In S. G. Brush (Ed.), *Kinetic theory*, vol. 2, 1996, (pp. 194–202). Oxford: Pergamon Press. (English reprint).

Prigogine, I. (1994). *Les Lois du Chaos*. Flammarion.

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

Ridderbos, T. M., & Redhead, M. L. G. (1998). The spin-echo experiments and the second law of thermodynamics. *Foundations of Physics*, *28*, 1237–1270.

Ruelle, D. (1991). *Chance and chaos*. Princeton: University Press.

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

Uffink, J. (2001). Bluff your way in the second law of thermodynamics. *Studies in History and Philosophy of Modern Physics*, *32*, 305–394.

Van Lith, J. (1999). Reconsidering the concept of equilibrium in classical statistical mechanics. *Philosophy of Science, 66*, S107–S118.

Van Lith, J. (2001). *Stir in stillness*. Ph.D. Thesis, University of Utrecht.

Von Plato, J. (1989). Probability in dynamical systems. In J. E. Fenstad, I. T. Frolov, R. Hilpinen (Eds.), *Logic, methodology and philosophy of science*, vol. VIII (pp. 427–443). Amsterdam: Elsevier.

Vranas, P. B. M. (1998). Epsilon-ergodicity and the success of equilibrium statistical mechanics. *Philosophy of Science, 65*, 688–708.