On the Definition of Equilibrium

Itamar Pitowsky

Department of Philosophy, The Hebrew University,

Mt. Scopus, Jerusalem 91905, Israel

e-mail: itamarp@vms.huji.ac.il and

web site: http://edelstein.huji.ac.il/staff/pitowsky/

Abstract

Boltzmann's approach to statistical mechanics is widely believed to be conceptually superior to Gibbs' formulation. However, the microcanonical distribution often fails to behave as expected: The ergodicity of the motion relative to it can rarely be established for realistic systems; worse, it can often be proved to fail. Also, the approach involves idealizations that have little physical basis. Here we take Khinchin's advice and propose a definition of equilibrium that is more realistic: The definition reflects the fact that the system is made of a great number of particles, and implies that all *measurable* macroscopic observables have steady values.

The statistical mechanical arrow of time is often defined in terms of the tendency of "isolated mechanical system consisting of a very large numbers of molecules, to approach thermal equilibrium in which all 'macroscopic' variables have reached steady values" (Uhlenbeck and Ford, 1963). An explanation of such tendency, in case it exists, is a step towards a definition of the arrow of time in statistical mechanics. However, we should be clear first on what is the meaning of equilibrium, or more precisely a state of equilibrium. For the sake of simplicity I shall work in the framework of classical statistical mechanics, so by *state* I mean a probability measure on phase space. My intention is not to eliminate the need for an independent notion of probability by reducing it to long term relative frequency. Nevertheless, the connection of probability to relative frequency will become clear.

The need for a realistic notion of equilibrium stems from various difficulties, some conceptual, and some technical. In the equilibrium state all macroscopic variables reach steady values, meaning that their observed values remain most of the time near their phase-space average (weighted by the state). The standard account, however, suffers from considerable problems:

1. Following Boltzmann and the Ehrenfests (1912) the microcanonical distribution is taken by many as the conceptually superior tool for the development of statistical mechanics (for example, Pauli, 1973). However, the microcanonical distribution often fails to behave as expected: The ergodicity of the motion relative to it can rarely be established for realistic systems; worse, it can often be proved to fail.

2. The common interpretation of 'macroscopic observable' is too liberal and includes many functions on phase-space which cannot possibly be measured (Khinchin, 1949).

3. The microcanonical distribution corresponds to completely and ideally closed systems. Here we encounter a schizophrenic attitude: The so-called foundations of the theory rely on a theoretical notion of an equilibrium state that hardly and rarely delivers the goods; but all the calculations are done with the Gibbs' canonical (or macro-canonical) distributions for systems that are not isolated.

4. The excuse given for using Gibbs' (as opposed to Boltzmann's) formulation in practice is that all ensembles (alledgedly) become equivalent in the thermodynamic limit (see for example, Hecht, 1990). But if the sheer number of particles plays such a substantial role in the explanation of thermal equilibrium, why not follow Khinchin's advice and put it from the start in the definition of equilibrium (Khinchin, 1949; see also Batterman, 1998).

The aim is therefore to define equilibrium in a more realistic way, that will not only apply to closed systems, but hopefully include the canonical ensembles among others. A definition which implies that in equilibrium all reasonable macroscopic observables have steady values, and reflect the fact that the system is made of a great number n of particles. The approach incorporates the spirit of the Malament and Zabell's (1980) proposal, and the epsilon-ergodicity view (Vranas, 1998) but is different from both in one important respect: A greater stress is placed on the role of the large number of particles in the phenomenon of equilibrium. It is my belief that equilibrium is to a great extent a combinatorial phenomenon, in the sense that what a dynamical system lacks in its mixing properties is greatly compensated by averaging over the huge number of degrees of freedom.

Let $x_1, x_2, ..., x_n$ be the positions of the particles, and their momenta be $p_1, p_2, ..., p_n$; the x_i's and p_j's change as a function of time according to the laws of motion and initial conditions. Let $\rho(x_1, x_2, ..., x_n, p_1, p_2, ..., p_n)$ be a probability density defined over the phase space associated with the system. We shall make three preliminary simplifying assumptions:

a. The probability density $\rho(x_1, x_2, ..., x_n, p_1, p_2, ..., p_n)$ is integrable with respect to the Lebesgue measure on $\mathbb{R}^{(6n)}$; that is, the probability measure is absolutely continuous with respect to the Lebesgue measure. In particular, this means that we deal with probability measures defined on the whole of phase space, like Gibbs's measure. Our definition however can be translated to probability measures which are supported on a hypersurface, but the representation becomes cumbersome. More on that in appendix 1.

b. The density ρ is stationary.

c. ρ is symmetric with respect to the interchange of particles. In other words:

$$\rho(x_{\pi(1)}, x_{\pi(2)}, ..., x_{\pi(n)}, p_{\pi(1)}, p_{\pi(2)}, ..., p_{\pi(n)}) = \rho(x_1, x_2, ..., x_n, p_1, p_2, ..., p_n) \tag{1}$$

for every permutation π of $1, 2, ..., n$. This assumption simplifies the presentation, but it is also unessential. In appendix 2 I indicate how the present treatment can be extended to non-symmetric cases.

Given a density ρ we can define its marginals:

$$\rho^{(1)}(x_1,p_1) = \tag{2}$$
$$\int\int ... \int \rho(x_1,x_2,...,x_n,p_1,p_2,...,p_n) d^3x_2 d^3x_3 ... d^3x_n d^3p_2 d^3p_3 ... d^3p_n$$
$$\rho^{(2)}(x_1,x_2,p_1,p_2) =$$
$$\int\int ... \int \rho(x_1,x_2,...,x_n,p_1,p_2,...,p_n) d^3x_3 d^3x_4 ... d^3x_n d^3p_3 d^3p_4 ... d^3p_n$$

and so on for all values $\rho^{(j)}$.

Let k be a fixed natural number such that $n \gg k$. We shall assume that k is of the order of magnitude of $\log n$. Roughly speaking, k is the order of the highest correlations that yield observable macroscopic phenomena, so that our estimation for k is very liberal. Let $A, B \subset \mathbb{R}^{(3)}$ be regions in 3-dimensional space, we assume that these are small regions, but still of macroscopic size. The letter A will designate regions in position space, and B in momentum space. Consider the expression

$$D(A,B,t) = \frac{1}{n} \sum_{i=1}^{n} \chi_B(p_i(t)) \chi_A(x_i(t)) \tag{3}$$

Where $\chi_A(x) = 1$ when $x \in A$, and 0 otherwise, and similarly for B. $D(A,B,t)$ is the average number of particles in region A whose momenta lie in the region B at time t. Similarly, if $A_1, B_1, A_2, B_2, ..., A_j, B_j \subset \mathbb{R}^{(3)}$ we denote

$$D(A_1,B_1,A_2,B_2,...,A_j,B_j,t) = \tag{4}$$
$$= \frac{(n-j)!}{n!} \sum \chi_{B_1}(p_{m_1}(t)) \chi_{A_1}(x_{m_1}(t)) ... \chi_{B_j}(p_{m_j}(t)) \chi_{A_j}(x_{m_j}(t))$$

where the sum is taken over all j-tuples $(m_1, ..., m_j)$ *with the m_i's all different*, and $1 \leq m_i \leq n$.

Definition: *ρ is a k-level equilibrium state if for almost all initial conditions $(x_1(0), ..., x_n(0), p_1(0), ..., p_n(0)) \in \mathbb{R}^{(6n)}$ and every family of regions $A_1, B_1, A_2, B_2, ..., A_j, B_j \subset \mathbb{R}^{(3)}$ we have*

$$\lim_{\tau\to\infty}\frac{1}{\tau}\int\limits_0^\tau D(A_1,B_1,A_2,B_2,...,A_j,B_j,t)dt \tag{5}$$
$$=\int_{B_1}\cdots\int_{B_j}\int_{A_1}\cdots\int_{A_j}\rho^{(j)}(x_1,...,x_j,p_1,...,p_j)d^3x_1...d^3x_jd^3p_1...d^3p_j$$

for all $1\leq j\leq k$[1].

Here 'almost all' refers to the Lebesgue measure on $\mathbb{R}^{(6n)}$. The role played by the Lebesgue measure is essentially identical to the role that it plays in the standard approaches to classical statistical mechanics. The attempt to eliminate the concept of probability by reducing it to relative frequency is circular. This is true in the present context, as it is true in the theory of probability in general. The Lebesgue measure plays here the role of an a-priory scale relative to which 'probability zero' is defined. The justification of this move- for example, by reference to the invariance properties of the Lebesgue measure- lies outside the scope of this article (see Guttmann, 1999).

There are two main differences between our definition and the requirement of ergodicity. Firstly, it involves averages over all particles. Consider, for example, the case $j=1$. In this case the conditions Eq (5) is

$$\lim_{\tau\to\infty}\frac{1}{\tau}\int\limits_0^\tau\left[\frac{1}{n}\sum_{i=1}^{n}\chi_B(p_i(t))\chi_A(x_i(t))\right]dt=\int_B\int_A\rho^{(1)}(x,p)d^3xd^3p \tag{6}$$

for almost all initial conditions. This means that the long run frequency of the average number of particles entering region A, while having their momenta in region B is the ρ-measure of $A\times B\times\mathbb{R}^{(6n-6)}$. This condition seems to be implied by the very notion of equilibrium (given the symmetry of ρ). Now, compare Eq (6) with the stronger definition of equilibrium, namely ergodicity. Take just one single particle, say particle number 1, then if the probability measure associated with ρ is ergodic we have:

$$\lim_{\tau\to\infty}\frac{1}{\tau}\int\limits_0^\tau\chi_B(p_1(t))\chi_A(x_1(t))dt=\int\limits_B\int\limits_A\rho^{(1)}(x,p)d^3xd^3p \tag{7}$$

[1] We can also consider a somewhat relaxed version of the definition according to which the difference between the phase-space and time averages in Eq (5) remains below a given threshold at all sufficiently long times.

for almost all initial conditions. This means that the long run frequency with which particle number 1 enters region A while having its momentum in region B is the ρ-measure of $A \times B \times \mathbb{R}^{(6n-6)}$, and similarly for each and every particle alone. The same remains approximately true (up to an error ε) when we relax the condition of full ergodicity and move to epsilon-ergodicity. This is really an unnecessary idealization. For all we know our initial condition may lead particle 1 to spend the entire time lying idly in one corner, and this will not influence our judgment that the gas is in a state of thermal equilibrium. However, in both the ergodic and epsilon-ergodic approaches such orbits in phase space are counted among the pathological non-equilibrium microstates. Similar remarks apply to the higher marginals $\rho^{(j)}$.

An even more extreme situation arises when we consider some non-symmetrical ρ's (appendix 2). Consider the case of a gas in which particle number 1 is of one kind a, and all the other particles are of another b, and there are no a-b interactions. Clearly, if n is large the statistical behavior of this gas is indistinguishable from that of a uniform b-gas. Now, consider a strange probability density ρ which is like the canonical distribution for the b-particles, but is supported on the left half of the container for the single a-particle (regardless of the fact that a's actual trajectory may frequently enter the right hand side). Clearly no observable difference will arise between this strange density and the canonical one ρ_c. So what is wrong with ρ? Well, the characteristic function of $A \times \mathbb{R}^{(6n-3)}$, where A is the left half of the container, fails to stabilize to its ρ-value. It has ρ-measure $= 1$, and ρ_c-measure $= 0.5$. However, the characteristic function of $A \times \mathbb{R}^{(6n-3)}$ is not a thermodynamic observable, it can be observed only if particle 1 can be individually traced. Why should we put requirements on its long term stabilization? This is a part of Khinchin's point which has not been incorporated by Malament and Zabell, or by the epsilon-ergodicity it inspired. The observables that do give rise to thermodynamic magnitudes are averages, and functions of averages over properties of particles, particle pairs, triples, and so on, until we get to k-tuples; where the $(k+1)$-fold correlations fall below (or even well below) observational capacity[2].

Indeed, the second and more important difference between the present approach and

[2] According to our definition for this case (appendix 2) the difference between ρ and ρ_c is $\leq O(\frac{1}{n})$ for all the averages over particle properties, and zero for all averages over j-tuples, for $j \geq 2$.

ergodicity is the 'cutoff' at $k \ll n$. To see the meaning of this assumption notice that we can represent the probability density ρ as

$$\begin{aligned}
&\rho(x_1, x_2, ..., x_n, p_1, p_2, ..., p_n) = \\
&= \prod_{j=1}^{n} \rho^{(1)}(x_j, p_j)(1 + \sum_{0 \leq i < j \leq n} \frac{c^{(2)}(x_i, x_j, p_i, p_j)}{\rho^{(1)}(x_i, p_i)\rho^{(1)}(x_j, p_j)} + \\
&+ \sum_{0 \leq i < j < k \leq n} \frac{c^{(3)}(x_i, x_j, x_k, p_i, p_j, p_k)}{\rho^{(1)}(x_i, p_i)\rho^{(1)}(x_j, p_j)\rho^{(1)}(x_k, p_k)} + ...)
\end{aligned} \tag{8}$$

The functions $c^{(2)}$, $c^{(3)}$,... are the pair, triple, etc. correlation densities:

$$\begin{aligned}
&c^{(2)}(x_1, x_2, p_1, p_2) = \\
&\rho^{(2)}(x_1, x_2, p_1, p_2) - \rho^{(1)}(x_1, p_1)\rho^{(1)}(x_2, p_2)
\end{aligned} \tag{9}$$

and more generally

$$\begin{aligned}
&c^{(j)}(x_1, x_2, ..., x_j, p_1, p_2, ..., p_j) = \\
&\sum_{\phi \subseteq \{i_1, ..., i_m\} \subseteq \{1, ..., j\}} (-1)^{j-m} \rho^{(m)}(x_{i_1}, x_{i_2}, ..., x_{i_m}, p_{i_1}, p_{i_2}, ..., p_{i_m}) \times \\
&\times \prod_{l \notin \{i_1, ..., i_m\}} \rho^{(1)}(x_l, p_l)
\end{aligned} \tag{10}$$

The assumption implicit in our definition of equilibrium is that no $c^{(j)}$, with $j > k$, has a measurable influence on any macroscopic observable, so that the sum on the right hand side of Eq (8) can effectively be stopped at $c^{(k)}$. The quantities $c^{(j)}$ can be interpreted in terms of correlations between fluctuations (Pitowsky, 2001). For example, consider

$$C(A_1, B_1, A_2, B_2) = \int_{B_1} \int_{B_2} \int_{A_1} \int_{A_2} c^{(2)}(x_1, x_2, p_1, p_2) d^3x_1 d^3x_2 d^3p_1 d^3p_2 \tag{11}$$

and suppose that $C(A_1, B_1, A_2, B_2) > 0$. Then, if the number of particles in region A_1 with momenta in region B_1 fluctuates above its average $\sim n \int\limits_{B_1} \int\limits_{A_1} \rho^{(1)}(x, p) d^3x d^3p$, it is likely that so does the number of particles in region A_2 with momenta in region B_2. Although the fluctuations themselves are small, and diminish as $n \to \infty$, the correlations between them remain finite in a non-ideal gas. However, one can expect that in a state of equilibrium, away from a phase transition, only a small number of $c^{(j)}$'s have a measurable contribution, although it is possible that none of them is strictly zero. Now, if ρ is assumed to be ergodic then Eq (5) is valid for every $1 \leq j \leq n$. Consequently, the correlations $c^{(j)}$ of all levels

$1 \leq j \leq n$ are assumed to stabilize in the long run. This is not a reasonable physical assumption; just think about the case $j = 0.5n$, or even $j = 0.1n$. Epsilon-ergodicity does not solve this problem either; it also entails that all $c^{(j)}$'s are stable as long as restrict the variables x_i, p_i to the large invariant subset on which the restriction of ρ is ergodic.

Let $F(x_1, ..., x_j, p_1, ..., p_j)$ be a real, continuous, integrable function of $6j$ variables, $j \leq k$. Put

$$\widetilde{F}(x_1, ..., x_n, p_1, ..., p_n) = \frac{(n-j)!}{n!} \sum F(x_{m_1}, ..., x_{m_j}, p_{m_1}, ..., p_{m_j}) \tag{12}$$

where the sum is taken over all j-tuples $(m_1, ..., m_j)$ with the m_i's all different, and $1 \leq m_i \leq n$.

Proposition: *If ρ is a k-level equilibrium state then for almost all initial conditions* $(x_1(0), ..., x_n(0), p_1(0), ..., p_n(0)) \in \mathbb{R}^{(6n)}$:

$$\lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau \widetilde{F}(x_1(t), ..., x_n(t), p_1(t), ..., p_n(t)) dt = \tag{13}$$

$$\int ... \int F(x_1, ..., x_j, p_1, ..., p_j) \rho^{(j)}(x_1, ..., x_j, p_1, ..., p_j) d^3x_1 ... d^3x_j d^3p_1 ... d^3p_j$$

The outline of the proof is given in appendix 3. This means that the long time average of $\widetilde{F}$ is the same as the phase-space average of $\widetilde{F}$ (which is the same as that of F). Therefore we can claim that "all macroscopic variables have reached steady values". The reason for this is that the thermodynamic observables of the gas all have the form $\widetilde{F}$, or functions of such $\widetilde{F}$'s; that is, averages over properties of at most k-tuples of particles, and functions of such averages. The reason to stop at k is obvious: If the $c^{(j)}$'s are very near zero for $j \geq k$, then it follows from Eq (8) that only marginals up to $\rho^{(k)}$ effectively enter the phase space averages.

The actual value of the 'cutoff' k is an empirical matter. The advantage of this definition is that it is likely that the usual probability densities, the canonical distributions in particular, are k-level equilibrium states for the relevant k; at least as long as the temperature is bounded away from its critical values. It is unlikely that this claim can be proved generally; more probably it will have to be established piecemeal, case by case. The main body of evidence that the canonical distribution satisfies Eq (5), at least within a small negligible error, comes from results on the thermodynamic limit (Minlos, 2000). For various systems (including various discrete lattice gasses which we did not discuss) the limit density ρ_∞ exists. It is often ergodic, so that one could expect it to have a regular average behavior for finite n. More importantly, the correlation functions of ρ_∞ decay fast with the distance between phase space points, which probably justifies the cutoff at $k \ll n$ for a finite n.

The disadvantage is that the definition is contingent to a limited extent on the specifics of the situation (for example, k may vary from one material to another, but there maybe a value of k large enough to cover all cases). The notion of equilibrium looses some of its universality, but it seems to me a small price to pay.

Acknowledgments: *I wish to thank the organizers of the Arrows of Time Symposium for the invitation, and in particular Lee Gohlike for the generous hospitality. This research is supported by grant number 879/02 of the Israel Science Foundation.*

Appendix

1. The case of a singular ρ: If the system is closed, and the density ρ is supported on a hypersurface Γ, we can use the following translation of the formulas: $\int\limits_B \int\limits_A \rho^{(1)}(x,p)d^3xd^3p$ translates to $\int\int_{\Gamma(A,B)} \rho(\gamma)d\gamma$, where $\Gamma(A,B) = \left(A \times B \times \mathbb{R}^{(6n-6)}\right) \cap \Gamma$, and $d\gamma$ is the Lebesgue measure on Γ. Similar translations apply to the other integrals. The assumption that ρ is symmetric implies that Γ is also invariant under permutations of the coordinates of the particles. The translation procedure applies also to the case of the microcanonical measure, that is, $d\gamma$ itself with the usual normalization procedure.

2. The case of a non-symmetric ρ: Suppose, for example, that the gas is a mixture of two types of molecules a and b, whose numbers are n_a and n_b respectively, with $n_a + n_b = n$. In this case the density ρ is invariant under the action of the group $S_{n_a} \times S_{n_b}$ of permutations of the a-particles, and permutations of the b particles. We consider the marginals $\rho^{(j_a,j_b)}$ with $0 \le j_a \le n_a$, and $0 \le j_b \le n_b$ by integrating over $n_a - j_a$ position and momentum variables of the type a particles, and $n_b - j_b$ variables of type b. The k-level equilibrium is defined in terms of the marginals $\rho^{(j_a,j_b)}$ for which $j = j_a + j_b \le k$. Rather than stating the definition in the most general case, consider the equivalent of Eq (5) in the case $j = 2$:

$$\lim_{\tau\to\infty} \frac{1}{\tau} \int\limits_0^\tau D(A_1, B_1, A_2, B_2, t)dt = \tag{14}$$

$$\int_{B_1}\int_{B_2}\int_{A_1}\int_{A_2} \left[\frac{n_a(n_a-1)}{n(n-1)}\rho^{(2,0)} + \frac{2n_a n_b}{n(n-1)}\rho^{(1,1)} + \frac{n_b(n_b-1)}{n(n-1)}\rho^{(0,2)}\right] d^3x_1 d^3x_2 d^3p_1 d^3p_2$$

Denote $\lambda = \frac{n_a}{n}$ so that $1 - \lambda = \frac{n_b}{n}$. Since n is large the coefficients of the marginals in the integrand in Eq (14) are very close to λ^2, $2\lambda(1-\lambda)$, and $(1-\lambda)^2$ respectively. The generalization to any $j \leq k$ is straightforward. In particular, since $k \ll n$ we can express the coefficients of the marginals in all those cases by powers of λ and $1-\lambda$. All the observations of the symmetric case remain intact with the obvious necessary adjustments. The case of three or more types of molecules is handled in the same manner. The extension to these cases is possible provided the mixture is made of a relatively small number of different types of particles.

3. Outline of the proof of the proposition: Denote by $\mathbf{s}$ a variable ranging over $\mathbb{R}^{(6j)}$. We shall use a sequence of approximations:

1. Since $F(\mathbf{s}) = F(x_1, ..., x_j, p_1, ..., p_j)$ is continuous and integrable we can find a continuous function F_1, such that $\|F - F_1\|_\infty$ (the supremum norm) is arbitrarily small; and such that $E = \{\mathbf{s};\ |F_1(\mathbf{s})| > 0\}$ is bounded, so that $E \subset I_1 \times I_2 \times ... \times I_{6j}$ where the $I_l \subset \mathbb{R}$ are finite intervals.

2. We can divide each interval I_l into finitely many small disjoint intervals $I_l = I_l^1 \cup I_l^2 \cup ... \cup I_l^{r_l}$, such that the variation of F_1 on any Cartesian product of the form $E(\mathbf{m})$ $=E(m_1, ..., m_{6j}) = I_1^{m_1} \times I_2^{m_2} \times ... \times I_{6j}^{m_{6j}}$ is as small as we wish. Choose arbitrary fixed $\mathbf{s_m} \in E(\mathbf{m})$ and define $G(\mathbf{s}) = \sum_{\mathbf{m}} F_1(\mathbf{s_m})\chi_{E(\mathbf{m})}(\mathbf{s})$. Then $\|F_1 - G\|_\infty$ is as small as we want.

3. $\widetilde{G}$ satisfies the claim, since G is a finite sum of the indicator functions of the hypercubes $E(\mathbf{m})$, but $\widetilde{G}$ is arbitrarily close to $\widetilde{F}$ in the supremum norm.

References

Batterman, R. (1998) 'Why Equilibrium Statistical Mechanics Works: Universality and the Renormalization Group' *Philosophy of Science* **65**, 183-208.

Ehrenfest, P. and Ehrenfest, T. (1912) *The Conceptual Foundations of the Statistical Approach in Mechanics* (New York: Dover 1990).

Guttmann, Y. M. (1999) *The Concept of Probability in Statistical Physics* (Cambridge: Cambridge University Press).

Hecht, C. E. (1990) *Statistical Thermodynamics and Kinetic Theory* (New York: Dover).

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

Malament, D. B. and Zabell, S. L. (1980) 'Why Gibbs phase averages work- The role of ergodic theory' *Philosophy of Science* **47**, 339-349.

Minlos, R. A. (2000) *Introduction to Mathematical Statistical Mechanics* (Providence, Rhode Island: American Mathematical Society).

Pauli, W. (1973) *Statistical Mechanics-Pauli Lectures on Physics Volume 4* (New York: Dover)

Pitowsky, I. (2001) 'Local Fluctuations and Local Observers in Equilibrium Statistical Mechanics' *Studies in the History and Philosophy of Modern Physics* **32**, 595-607.

Uhlenbeck, G. E. and Ford, G. W. (1963) *Lectures in Statistical Mechanics* (Providence: American Mathematical Society).

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