

Individualist and Ensemblist Approaches to the Foundations of Statistical Mechanics

Sheldon Goldstein

Departments of Mathematics, Physics, and Philosophy – Hill Center
Rutgers, The State University of New Jersey
110 Frelinghuysen Road
Piscataway, NJ 08854-8019, USA
E-mail: oldstein@math.rutgers.edu

March 15, 2019

Abstract

I will contrast the two main approaches to the foundations of statistical mechanics: the individualist (Boltzmannian) approach and the ensemblist approach (associated with Gibbs). I will indicate the virtues of each, and argue that the conflict between them is perhaps not as great as often imagined.

1 Introduction

It is well known that the foundations of quantum mechanics are highly controversial. So too are the foundations of statistical mechanics, though it is not quite so widely recognized that this is so. Perhaps the main source of the controversy and confusion is the role played by probability in the subject. In fact probability plays a variety of roles in the foundations of statistical mechanics, roles whose differences are not always well appreciated.

The very notion of probability brings with it a strong suggestion of uncertainty, ignorance, and strength of belief, and hence of subjectivity. Thus one could easily form the impression that subjectivity plays a crucial role in our understanding of statistical mechanics, and in the thermodynamics—including in particular its second law, about entropy increase—to which statistical mechanics gives rise.

Indeed, John von Neumann [1, p. 400], who made profound contributions to the foundations of quantum statistical mechanics, has written the following:

The time variations of the entropy are then based on the fact that the observer does not know everything, that he cannot find out (measure) everything which is measurable in principle.

Von Neumann's statement falls within the framework of what has been called the *ensamblist* view on the foundations of statistical mechanics, often associated with Josiah Willard Gibbs [2]. This view is in contrast with the *individualist* view associated with Ludwig Boltzmann.

Here are a couple of statements by Schrödinger in which he expresses his strong support for the Boltzmannian view:

The spontaneous transition from order to disorder is the quintessence of Boltzmann's theory . . . This theory really grants an understanding and does not . . . reason away the dissymetry of things by means of an a priori sense of direction of time . . . No one who has once understood Boltzmann's theory will ever again have recourse to such expedients. It would be a scientific regression beside which a repudiation of Copernicus in favor of Ptolemy would seem trifling. [3]

. . .no perception in physics has ever seemed more important to me than that of Boltzmann—despite Planck and Einstein. [4]
(Schrödinger)

And here is a statement of Boltzmann [5] himself:

I have . . . emphasized that the second law of thermodynamics is from the molecular viewpoint merely a statistical law. Zermelo's paper shows that my writings have been misunderstood; . . . Poincaré's theorem, which Zermelo explains at the beginning of his paper, is clearly correct, but his application of it to the theory of heat is not. . . . Thus, when Zermelo concludes, from the theoretical fact that the initial states in a gas must recur – without having calculated how long a time this will take – that the hypotheses of gas theory must be rejected or else fundamentally changed, he is just like a dice player who has calculated that the probability of a sequence of 1000 one's is not zero, and then concludes that his dice must be loaded since he has not yet observed such a sequence!

Note that in this quote Boltzmann makes several references to probability and related notions, suggesting an ensamblist character. Nonetheless, as should become clear,

the statement lies entirely within the individualist view about the foundations of statistical mechanics.

In what follows I shall focus only on classical statistical mechanics. While the case of quantum statistical mechanics would involve some important differences, for the purpose of what I wish to convey here these differences would not be relevant. I shall also choose units for which Boltzmann's constant $k = 1$.

2 Individualist and ensemblist approaches to our understanding of thermal equilibrium

So what exactly does it mean for a system, say a gas in a box, to be in thermal equilibrium? There are two views about this, corresponding to two different attitudes towards the foundations of statistical mechanics: individualist and ensemblist.

For an individualist, a system is in thermal equilibrium if it is in an *appropriate pure state*, given by a suitable point in phase space. For an ensemblist, it is in thermal equilibrium if it is in an *appropriate statistical state or ensemble*, given by a suitable probability measure on phase space.

2.1 Equilibrium

We shall focus here on the notion of equilibrium for the simple case of a classical gas in a box $\Lambda \subset \mathbb{R}^3$ consisting of N atoms of a single type. We shall assume that the internal structure of the atoms can be ignored, so that each atom may be regarded as a point particle.

2.1.1 Individualist

A pure state then corresponds to the positions $\mathbf{q}_1, \dots, \mathbf{q}_N$ and the momenta $\mathbf{p}_1, \dots, \mathbf{p}_N$ of the particles, which together form the *microstate* of the system, a point

$$X = (\mathbf{q}_1, \dots, \mathbf{q}_N, \mathbf{p}_1, \dots, \mathbf{p}_N)$$

in the phase space

$$\Gamma = \Lambda^N \times \mathbb{R}^{3N} = (\Lambda \times \mathbb{R}^3)^N$$

of the gas. (We shall usually assume that the particles all have mass $m = 1$, so that we needn't distinguish between velocities and momenta.)

An important quantity associated with the system is its energy, given by a function $H(X)$ on its phase space. The energy is typically a sum of a kinetic energy term, depending only on the momenta, and a potential energy term, depending only on the positions. However, we shall not be too concerned here with the form of H .

Since the energy is a constant of the motion we often regard it as given, and we then usually focus on the energy surface

$$\Gamma_E = \{X : H(X) = E\}$$

corresponding to energy E , instead of the entire phase space Γ .

In addition to the energy E , the particle number N , and the volume $V = |\Lambda|$, where $|\cdot|$ denotes volume, there are other macrovariables—for example, those that convey the density profile—that are relevant to the notion of thermal equilibrium: The phenomenon that lies behind the notion of a system being in equilibrium is the fact that the system tends to evolve into a state in which such macrovariables no longer seem to change in a noticeable way. An example of such macrovariables, relevant especially for a low density gas, will be described in Section 2.3.

Give an appropriate choice of macrovariables one lumps together microstates for which these variables have more or less the same value, to obtain a partition

$$\Gamma_E = \bigcup_{\nu} \Gamma_{\nu} \tag{1}$$

of the energy surface into a collection of macrostates Γ_{ν} . We can think of this partition as corresponding to the relation of macroscopic equivalence. The points in the same macrostate Γ_{ν} are more or less indistinguishable from a macroscopic perspective.

In Figure 1 we have depictions of two rather different partitions of the energy surface into macrostates. Both involve a special macrostate, indicated by “Eq”, bigger than all the others. This *dominant* macrostate, for a gas in a box, is in fact so very dominant that it occupies almost the entire energy surface. The depiction on the right is thus much more appropriate than the one on the left. But it too is unrealistic since only the dominant macrostate would be visible if the macrostates were drawn to scale.

The special dominant macrostate is, ipso facto, the *equilibrium macrostate*. It happens to be characterized by the uniformity and other properties associated with equilibrium. According to what has just been said, we have that

$$\frac{|\Gamma_{eq}|}{|\Gamma_E|} \approx 1. \tag{2}$$

For an individualist a system is *in thermal equilibrium* if its microstate X is in the equilibrium macrostate, $X \in \Gamma_{eq}$.

2.1.2 Ensemblist

Suppose we more or less know that our system, our gas in a box, is in equilibrium. This may be because we’ve waited long enough or otherwise somehow come to know

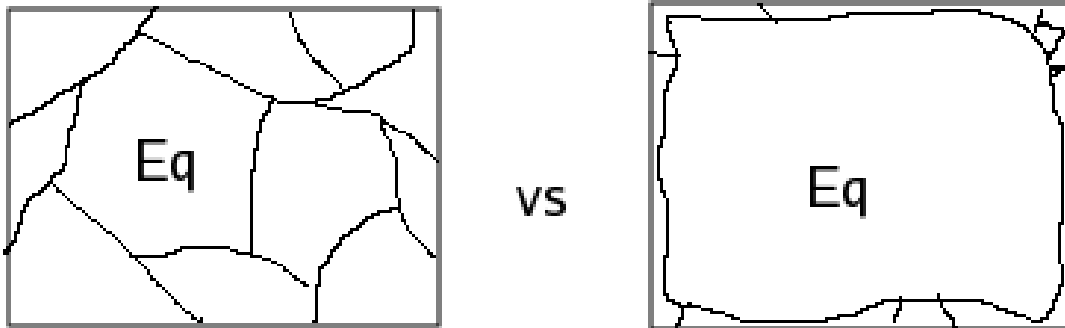


Figure 1: Two depictions of the partition of Γ_E into macrostates

that the system’s macrovariables have their equilibrium values. In this situation we presumably do not know, nor need to know, the exact microstate X of the system, about which we are almost surely quite uncertain. We might thus imagine that a more useful description of a system in thermal equilibrium would be provided, not by its exact microstate X , but by an *ensemble*, a probability distribution over possible microstates.

This is indeed the case, for an individualist as well as for an ensembler, as will be explained in Section 3.3. But an ensembler goes much further. For an ensembler the very meaning of a system’s being in thermal equilibrium is essentially connected with the use of ensembles. For an ensembler a system is in thermal equilibrium if its microstate X is in some sense random, with probability distribution given by one of the fundamental statistical mechanical ensembles, for example the microcanonical ensemble (24), corresponding to a “uniform” distribution over the energy surface Γ_E , or the canonical ensemble (25); see Section 3.3.

2.2 Evolution to equilibrium

2.2.1 Ensembler

Insofar as equilibrium and approach to equilibrium are concerned, an ensembler regards the state of a system as given by a probability distribution—for a system in equilibrium one of the standard ensembles, and for non-equilibrium some other probability distribution, a non-equilibrium ensemble. One of the advantages of the ensembler perspective is that it affords a simple clean *mathematical* formulation of what it means for a system to approach equilibrium, namely that in the infinite-time limit, time-evolving non-equilibrium ensembles μ_t approach an equilibrium ensemble

$$\mu_{eq}, \quad \mu_t \longrightarrow \mu_{eq} \quad (3)$$

as $t \rightarrow \infty$. For a sufficiently chaotic dynamical system, this sort of behavior, for suitable initial non-equilibrium states, can actually occur. The technical term for this sort of behavior is *mixing*.

While mixing is a clear mathematical notion, its physical relevance to the problem at hand is somewhat obscure. What, after all, does *the* probability distribution μ_t of our system at a given time refer to? What in fact is its actual probability distribution? I'm aware of no plausible answer to this question.

2.2.2 Individualist

Approach to equilibrium for an individualist is mathematically not so simple. An individualist counterpart to (3), such as

$$X_t \longrightarrow \Gamma_{eq}$$

in the sense that $X_t \in \Gamma_{eq}$ (or is near Γ_{eq}) as $t \rightarrow \infty$, is typically impossible. This is because of the Poincaré recurrence theorem, which implies that almost all microstates in a macrostate will return, under the time evolution, to that macrostate, in fact infinitely often. It will typically not be the case that the system, initially far from equilibrium, evolves so as to be in, or near, equilibrium for all sufficiently large times. In fact, even when the system begins in, or near, equilibrium, one expects the system to fluctuate far from equilibrium after a sufficiently long (probably an incredibly long) time.

That this is so did not bother Boltzmann, and it should not bother us either. As Boltzmann [6] wrote:

The applicability of probability theory to a particular case cannot of course be proved rigorously. . . . Despite this, every insurance company relies on probability theory. . . . It is completely incomprehensible to me how anyone can see a refutation of the applicability of probability theory in the fact that some other argument shows that exceptions must occur now and then over a period of eons of time; for probability theory itself teaches just the same thing.

What can occur, and what approach to equilibrium for an individualist does mean, is this:

$$X_t \in \Gamma_{eq} \text{ for } \textit{by far most} \text{ (sufficiently large) } t \quad (4)$$

(even when the system is initially not in equilibrium). A sufficient condition for this is that the classical dynamics of the system be mixing or, in fact, merely ergodic,

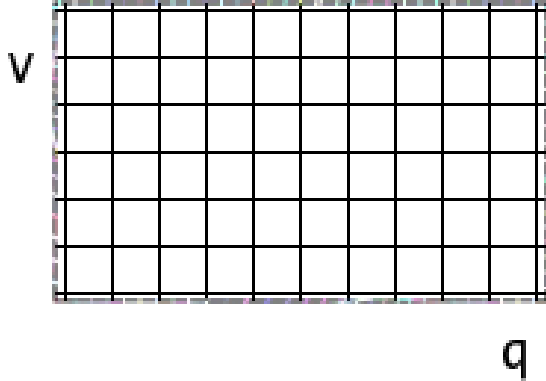


Figure 2: Cells of the one-particle phase space $\Lambda \times \mathbb{R}^3$

but this is hardly necessary. While (4) is difficult to prove for any realistic system, Figure 1 and equation (2) suggest that it should generically hold.

The sort of behavior expressed by (4) is clearly physically relevant, and is, indeed, exactly the sort of behavior needed to account for the observed phenomenon of approach to thermal equilibrium.

2.3 The individualist approach in the simplest case

In his analysis in 1872 of the approach to equilibrium of a low density gas, Boltzmann chose as macrovariables the *1-particle empirical distribution* $f_{emp}(\mathbf{q}, \mathbf{v})$, conveying the density of particles in the 1-particle phase space $\Lambda \times \mathbb{R}^3$. This is based on a partition of that space into cells Δ , see Figure 2.3, that are macroscopically small but microscopically large, and expresses the (normalized) number of particles per unit volume in each cell: For $X = (\mathbf{q}_1, \mathbf{v}_1, \dots, \mathbf{q}_N, \mathbf{v}_N)$,

$$f_{emp}(\mathbf{q}, \mathbf{v}) \equiv f_X(\mathbf{q}, \mathbf{v}) = \frac{n_X(\Delta(\mathbf{q}, \mathbf{v}))/N}{|\Delta(\mathbf{q}, \mathbf{v})|}, \quad (5)$$

where $\Delta(\mathbf{q}, \mathbf{v})$ is the cell containing the point $(\mathbf{q}, \mathbf{v}) \in \Lambda \times \mathbb{R}^3$, $|\Delta(\mathbf{q}, \mathbf{v})|$ is its volume, and $n_X(\Delta)$ is the number of particles of the microstate X (at locations $(\mathbf{q}_i, \mathbf{v}_i)$) in Δ .

I've written f_X above in order to emphasize that the empirical distribution here is a function of the microstate X (not a real-valued function but one whose values are functions of \mathbf{q} and \mathbf{v}). Since X is time dependent so is the empirical distribution f_X :

$$f_{emp}(\mathbf{q}, \mathbf{v}, t) \equiv f_{X(t)}(\mathbf{q}, \mathbf{v}). \quad (6)$$

The exact evolution of $f_{emp}(\mathbf{q}, \mathbf{v}, t)$ is extremely complicated. However, Boltzmann argued—and non-rigorously showed—that, at low density, this evolution is *typically* well approximated (for times that are not too very large) by a much simpler one given by $f_t(\mathbf{q}, \mathbf{v})$,

$$f_{X(t)}(\mathbf{q}, \mathbf{v}) \approx f_t(\mathbf{q}, \mathbf{v}), \quad (7)$$

that obeys *Boltzmann's equation*

$$\frac{\partial f_t}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{q}} f_t = Q(f_t), \quad (8)$$

where Q is Boltzmann's collision kernel, which happens to be quadratic in f and about which I shall say no more.

The qualification “typically” above (7) refers to the fact that—as a consequence of time-reversal invariance or of Poincaré recurrence—there are rare exceptional (initial) microstates X for which the empirical distribution is not well approximated by a solution to Boltzmann's equation (8). In 1872 Boltzmann did not realize that this qualification was necessary. By 1877 he did.

As consequences of Boltzmann's equation:

- $f_t(\mathbf{q}, \mathbf{v}) \rightarrow f_{eq}(\mathbf{q}, \mathbf{v}) \sim e^{-\frac{1}{2}m\mathbf{v}^2/kT}$ as $t \rightarrow \infty$, corresponding to the approach to the equilibrium macrostate. (This has not been entirely proven with complete mathematical rigor, but there is little doubt that it is true for all practical purposes.)
- The Boltzmann H -function

$$H(f_t) = \int f_t(\mathbf{q}, \mathbf{v}) \log f_t(\mathbf{q}, \mathbf{v}) \mathbf{d}\mathbf{q} \mathbf{d}\mathbf{v}$$

is nonincreasing as t increases, and remains constant for $f = f_{eq}$, which minimizes $H(f)$. This suggested to Boltzmann that the entropy S of the second law of thermodynamics, a quantity which is supposed never to decrease for an isolated thermodynamic system, should be identified microscopically with

$$S(X) = -NH(f_X). \quad (9)$$

2.4 A wrong turn, suggesting an ensembler approach

Suppose we regard $f = f_X$ not as an empirical distribution but as a probability distribution, expressing something like, say, ignorance. Here is a natural and probably not uncommon thought: For a gas at low density the particles should be more or less independent, since the interactions between them should be rather weak. In this case the probability distribution ϱ for the entire gas should be the product distribution

$$\varrho(X) \sim \prod_i f(\mathbf{q}_i, \mathbf{v}_i).$$

For such a distribution the right hand side of (9) is in fact the *Gibbs entropy* S_G :

$$S_G(\varrho) = - \int \varrho(X) \log \varrho(X) dX, \quad (10)$$

where in general

$$\varrho(X) \equiv \varrho(\mathbf{q}_1, \mathbf{v}_1, \dots, \mathbf{q}_N, \mathbf{v}_N).$$

When ϱ is of the above product form, we have that

$$S_G(\varrho) = -NH(f).$$

This would seem to suggest that the appropriate extension of Boltzmann's entropy formula (9) for a low density gas to higher density should be given by the Gibbs entropy S_G . And insofar as S_G is a function of a probability distribution on the phase space of the gas, an ensembler approach to the foundations statistical mechanics would then seem to be appropriate.

2.5 The individualist approach: entropy and macrostates

The Gibbs entropy (10) is a very important concept, useful in a variety of disciplines, from physics to mathematics to information theory and much else. It is not, however, the appropriate extension of Boltzmann's entropy formula (9) to higher density. After all, f_X is an empirical distribution, and thus is determined by the microstate X of an individual system. Hence the quantity (9) is as well.

Between 1872 and 1877 Boltzmann realized that the appropriate extension of (9) is the *Boltzmann entropy* S_B , given by

$$S_B(X) = \log |\Gamma_X|, \quad (11)$$

where Γ_X is the macrostate containing the system's microstate X . This of course depends upon a choice of macrostates Γ_ν , see (1), which in turn depends upon a sensible choice of macrovariables, the precise details of which do not much matter in practice.

Boltzmann chose for his analysis the macrovariables $f_X(\mathbf{q}, \mathbf{v})$. These constitute a rather large collection of macrovariables: a function of X for each cell Δ in Figure 2.3, given by $X \mapsto f_X(\mathbf{q}, \mathbf{v})$ for $(\mathbf{q}, \mathbf{v}) \in \Delta$. The corresponding macrostates are then of the form

$$\Gamma_f = \{X \in \Gamma_E \mid f_X(\mathbf{q}, \mathbf{v}) \approx f(\mathbf{q}, \mathbf{v})\}$$

for suitable choices of prescribed functions f and a suitable sense of \approx . For such macrostates we have that

$$\Gamma_X = \Gamma_{f_X}.$$

and that

$$S_B(X) = \log |\Gamma_{f_X}|.$$

Boltzmann proceeded then to compute the volume of these macrostates, finding that, at low density,

$$|\Gamma_f| \sim e^{-NH(f)}, \quad (12)$$

and thus that

$$\log |\Gamma_f| \approx -NH(f).$$

Thus, for a low density gas,

$$S_B(X) \approx -NH(f_X)$$

and (11) is indeed the appropriate extension of (9).

Since for a macroscopic system the number of particles N is very large, say

$$N \sim 10^{20},$$

it is strongly suggested by (12) that by far most of Γ_E is $\Gamma_{f_{eq}}$, as indicated in Figure 1, so that $\Gamma_{f_{eq}} \equiv \Gamma_{eq}$ is indeed dominant, i.e., the equilibrium macrostate. In fact, (12) suggests that the volume of the totality of nonequilibrium microstates, relative to $|\Gamma_{eq}|$, is of the incredibly small order $10^{-10^{20}}$ (for $N = 10^{20}$).

This corresponds to a vast distortion between the microscopic and the macroscopic levels of description, see Figure 3. On the right we have the f -space, the space of possible macrostates, associated with the values of the macrovariables f . In this space there is one special point, one special choice of f , namely f_{eq} represented by the central dot. The arrows on the right represent the trajectories of the Boltzmann equation, all approaching f_{eq} . That they should do so is made almost inevitable by the corresponding microscopic image on the left, almost all points of which correspond to the central dot on the right.

With Figure 3 in mind one should carefully read the following statement of Boltzmann [5]:

One should not forget that the Maxwell distribution is not a state in which each molecule has a definite position and velocity, and which is thereby attained when the position and velocity of each molecule approach these definite values asymptotically. . . . It is in no way a special singular distribution which is to be contrasted to infinitely many more non-Maxwellian distributions; rather it is characterized by the fact that by far the largest number of possible velocity distributions have the characteristic properties of the Maxwell distribution, and compared to these there are only a relatively small number of possible distributions that deviate significantly from Maxwell's. Whereas Zermelo says that the number

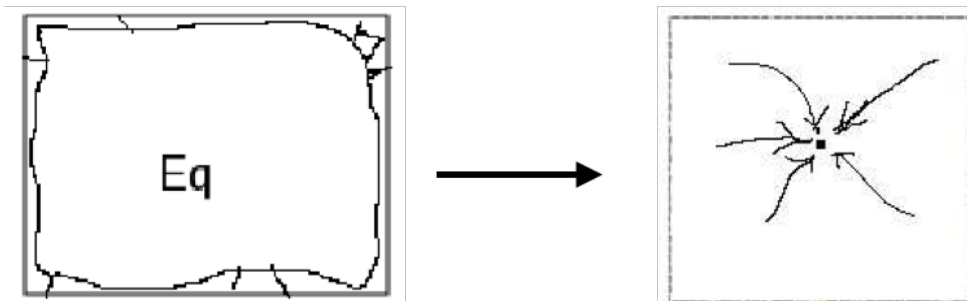


Figure 3: On the left we have the microscopic picture, consisting of the microstates $X \in \Gamma_E$, almost all of which are equilibrium microstates, in $\Gamma_{f_{eq}}$. On the right we have the macroscopic picture, consisting of the macrostates f and comprising the f -space, with the dot in the center representing f_{eq} . The arrow connecting the pictures represents the link between the two levels of description, given by the map $X \mapsto f_X$ that carries almost the entire image on the left to the central dot on the right (in the $N \rightarrow \infty$ limit).

of states that finally lead to the Maxwellian state is small compared to all possible states, I assert on the contrary that by far the largest number of possible states are “Maxwellian” and that the number that deviate from the Maxwellian state is vanishingly small.

By “Maxwell distribution” here Boltzmann is referring to f_{eq} , and by “velocity distributions” to microstates X .

2.6 The value of the ensemblist approach

The ensemblist and individualist approaches to the foundations of statistical mechanics are obviously quite different. It should be clear that I believe that the individualist approach is the correct one. Nonetheless, I also believe that while the ensemblist approach is conceptually confused, its machinery is quite useful.

For many practical purposes the two approaches agree. For example, if ϱ is uniform on the macrostate Γ_X , it is easy to see that $S_B(X) = S_G(\varrho)$, see (11).

Moreover, ensembles often provide an individualist with an extremely efficient means to compute the equilibrium values of macrovariables of interest. For example, consider the total kinetic energy K (and suppose it is not among the macrovariables used to specify an equilibrium macrostate—if it were its value would ipso facto be explicitly given with the specification of the macrostate). Then the equilibrium value of K will, with negligible error, be given by the microcanonical average of K : For typical (i.e., by far most) $X \in \Gamma_{eq}$

$$K(X) \approx \langle K \rangle_E \tag{13}$$

where $\langle \cdot \rangle_E$ denotes the microcanonical average—the average with respect to the microcanonical ensemble μ_E , see (24), corresponding to the appropriate uniform distribution over the energy surface Γ_E . This is because, just as Γ_{eq} is almost all of Γ_E , by law of large numbers type considerations K should be almost constant as it varies over the energy surface. Thus its average value should be more or less the same as its typical value.

There is a respect in which the ensemblist machinery is of even greater value: it provides a sharp clean model for the thermodynamic formalism. While the exact relationship among the values of macrovariables for a gas in a box are quite complicated, these values are extremely well approximated by ensemblist values related in a much simpler way, as described by the thermodynamics formalism. This will be discussed in the next section.¹

3 Thermodynamic ensembles and the thermodynamic formalism

Let us now turn to the crucial issue of how it is that the thermodynamic ensembles—the Gibbsian approach—efficiently captures the relationships embodied in the thermodynamic formalism. We begin with a few words about that formalism.

3.1 Representations

All equilibrium thermodynamic relations are given by a *fundamental relation* or equation of the form

$$S = S(X_1, \dots, X_m),$$

expressing the entropy in terms of a collection of extensive macrovariables X_i . For example, for the simple sort of gas of atoms in a box that we’ve been considering, we would have

$$S = S(U, V, N), \tag{14}$$

where U is the energy, V the volume, and N the particle (atom) number of the gas. This is the only case that we shall consider here. The entropy function is assumed in

¹For a somewhat similar example, consider Boltzmann’s equation, describing an evolution that is far simpler than the exact evolution of the empirical distribution f_X , and which has a sort of ensemblist flavor. (Indeed, Boltzmann’s equation is often regarded as describing the evolution, in a suitable limit, of the one-particle marginal distribution arising from a probability distribution on the N -particle phase space.)

thermodynamics (axiomatically) to have certain properties, such as concavity, which is required for stability, but we shall not discuss these properties here. We shall discuss instead the partial derivatives of S with respect to its variables, i.e., U , V , and N . These partial derivatives yield important “intensive” variables:

$$\partial S/\partial U = 1/T, \quad \partial S/\partial V = -p/T, \quad \text{and} \quad \partial S/\partial N = -\mu/T,$$

where T is the temperature, p the pressure, and μ the so-called chemical potential.

What I have just described, involving the entropy as a function of basic extensive variables, is called the *entropy representation*. There is another (equivalent) representation, called the *energy representation*, in which the roles of entropy and energy are exchanged. The fundamental relation in the energy representation is

$$U = U(S, V, N) \tag{15}$$

obtained by solving $S = S(U, V, N)$ for U in terms of S , with the other variables fixed. (It follows from the properties of S that U is convex.) $U(S, V, N)$ contains exactly the same information as $S(U, V, N)$ (since, in fact, one can be obtained from the other).

The partial derivatives in the energy representation are simpler:

$$\partial U/\partial S = T, \quad \partial U/\partial V = -p, \quad \text{and} \quad \partial U/\partial N = \mu. \tag{16}$$

Equivalently, $dU = TdS - pdV + \mu dN$. That $dU = TdS - pdV$ when $dN = 0$ should be clear, if we recall the original defining equation for the thermodynamic entropy, namely $dS = dQ/T$, where Q refers to heat.

3.2 Subrepresentations

It is often useful to use, for example, temperature as a variable, instead of energy. If one expresses the energy as

$$U = U(T, V, N) \tag{17}$$

(assuming this is possible), this relationship need not provide complete thermodynamic information: it need not be possible to reconstruct the fundamental relation (14) or (15) from (17).

But one can reconstruct the fundamental relation from its *Legendre transform*, with respect to any combination of its independent variables in both the energy and entropy representations. (Physicists tend to mean by Legendre transform the negative of what mathematicians mean by it, and I will be using here the physicist’s convention.)

The most important subrepresentations are for the energy representation. Two of the most important of these—important also for our purposes here—are the free

energy representation and the pressure, or grand canonical, representation. Once again, both of these involve fundamental relations—from which, say, the energy representation can be reconstructed, in fact by a Legendre transform.

The free energy representation expresses the free energy $F = U - TS$ as a function of T , V , and N :

$$F = F(T, V, N) \tag{18}$$

is the fundamental relation in this representation. The function (18) is the Legendre transform of (15) with respect to S , since $\partial U/\partial S = T$. We then have that

$$\partial F/\partial T = -S, \quad \partial F/\partial V = -p, \quad \text{and} \quad \partial F/\partial N = \mu. \tag{19}$$

To obtain the grand canonical (or pressure) representation we form the Legendre transform of the energy (15) with respect to both S and N , obtaining the grand canonical potential

$$\mathcal{G}(T, V, \mu) = U - TS - \mu N \tag{20}$$

regarded as a function of T , V , and $\mu = \partial U/\partial N$. We have that

$$\partial \mathcal{G}/\partial T = -S, \quad \partial \mathcal{G}/\partial V = -p, \quad \text{and} \quad \partial \mathcal{G}/\partial \mu = -N. \tag{21}$$

This representation is called the pressure representation because the grand canonical potential per unit volume is minus the pressure p :

$$\mathcal{G} = -pV. \tag{22}$$

This follows from the extensivity of the energy (15) (i.e., that it is homogeneous of degree 1), which yields the Euler relation:

$$U = TS - pV + \mu N. \tag{23}$$

(To see this take the derivative with respect to λ of both sides of

$$U(\lambda S, \lambda V, \lambda N) = \lambda U(S, V, N)$$

and evaluate at $\lambda = 1$. (23) then follows from (16).)

This representation is called the grand canonical representation because it is related to an important ensemble, the grand canonical ensemble, which I shall soon describe. In fact, we have three basic ensembles connected to three basic representations: the microcanonical ensemble is connected to the entropy representation, the canonical ensemble is connected to the free energy representation, and the grand canonical ensemble is connected to the pressure representation.² We shall now turn to this connection.

²In the canonical ensemble the energy can vary. Similarly, in the grand canonical ensemble also N can—corresponding to the fact that \mathcal{G} is directly a function of μ rather than N .

3.3 Ensembles and representations

In addition to the microcanonical ensemble

$$d\mu_E = \frac{1}{Z_E} d\sigma_E / |\nabla H| \quad (24)$$

corresponding to the “uniform” distribution on the energy surface Γ_E , and the canonical ensemble

$$d\mu_\beta = \frac{1}{Z_\beta} e^{-\beta H} dX, \quad (25)$$

a probability measure, on the N -particle phase space $\Gamma = \Gamma^{(N)}$, with unconstrained energy H , there is also the grand canonical ensemble,

$$d\mu_{\beta,\mu} = \frac{1}{Z_{\beta,\mu}} e^{-\beta(H-\mu N)} dX, \quad (26)$$

where now dX refers to the Lebesgue measure on

$$\bigcup_{N=1}^{\infty} \Gamma^{(N)},$$

for which the particle number N is also unconstrained.

Here μ is identified with the chemical potential, with corresponding *fugacity*

$$z = e^{\beta\mu}$$

in terms of which (26) becomes

$$d\mu_{\beta,\mu} = \frac{1}{Z_{\beta,\mu}} z^N e^{-\beta H} dX,$$

i.e., the measure on $\bigcup \Gamma^{(N)}$ with density

$$\rho_{\beta,\mu}^{(N)} = \frac{1}{Z_{\beta,\mu}} z^N e^{-\beta H}$$

with respect to the Lebesgue measure on each $\Gamma^{(N)}$. The *partition functions* Z_E , Z_β , and $Z_{\beta,\mu}$ are normalization constants.

In probability theory one often regards normalization constants as somewhat insignificant. But in the Gibbsian approach to statistical mechanics, the normalization constants Z are almost the whole ball game. The partition functions Z for the three ensembles provide us with the three representations for the fundamental

thermodynamic relations discussed earlier. In fact, we can more or less identify the (equilibrium) Boltzmann entropy $S_B = \log |\Gamma_{eq}|$ with $S_B(E)$, the logarithm of Z_E ,

$$S_B \simeq S_B(E) = \log Z_E,$$

since

$$Z_E = \int_{\Gamma_E} d\sigma_E / |\nabla H| \equiv |\Gamma_E| \simeq |\Gamma_{eq}|.$$

Similarly, $\log Z_\beta$ and $\log Z_{\beta,\mu}$ are simply related to the free energy $F(T, V, N)$ and the grand canonical potential $\mathcal{G}(T, V, \mu)$, in fact as follows:

Consider the canonical partition function

$$Z_\beta = Z_\beta(V, N) = \int_{\Gamma} e^{-\beta H} dX, \quad (27)$$

the normalization constant in (25). We have that

$$Z_\beta = \int_{E_{min}}^{\infty} e^{-\beta E} |\Gamma_E| dE = \int_{E_{min}}^{\infty} e^{-\beta E} e^{S_B(E)} dE. \quad (28)$$

The integrand is concentrated near an energy E_β , with small energy spread ΔE ($\sim \sqrt{E_\beta}$). Thus

$$Z_\beta \simeq e^{-\beta E} e^{S_B(E)} \Delta E$$

so that

$$-\log Z_\beta \simeq \beta E_\beta - S_B(E_\beta) = \beta \left(E_\beta - T_\beta S_B(E_\beta) \right). \quad (29)$$

The quantity inside $\left(\right)$ looks like the free energy. But one should check that $T_\beta \equiv 1/\beta$ is indeed the temperature T as defined by thermodynamics, namely given by

$$\partial S / \partial E = 1/T. \quad (30)$$

To do so, let's maximize the exponent on the right hand side of (28), namely $S_B(E) - \beta E = -\beta(E - T_\beta S_B(E))$, and hence in the canonical probability distribution for the energy, by setting its derivative with respect to E to zero. We obtain

$$1 - T_\beta \partial S_B / \partial E = 1 - T_\beta / T = 0$$

so that $T = T_\beta$, as desired.

In passing, note that we have just found that the thermodynamic temperature T , defined by (30), also gives the kinetic energy per unit particle ($\frac{3}{2}kT$) in the familiar way, since this is given by T_β (as a consequence of the form of the kinetic energy term in the formula for the canonical ensemble).

It follows from (29) that, with negligible error for large systems, we have the following formula for the individualist free energy:

$$F = -T \log Z_\beta. \quad (31)$$

We emphasize that $-T \log Z_\beta$, from an individualist perspective, is only a very good approximation for F , valid for macroscopic systems. From a pure ensemblist perspective, however, in which the Gibbs entropy is used instead of the Boltzmann entropy, it is exact: We have, writing $\langle \cdot \rangle_\beta$ for the expected value with respect to μ_β , that the ensemblist free energy is

$$\begin{aligned} \langle H \rangle_\beta + T \left\langle \log \frac{e^{-\beta H}}{Z_\beta} \right\rangle_\beta &= \langle H \rangle_\beta + \beta^{-1} \left(-\beta \langle H \rangle_\beta - \log Z_\beta \right) \\ &= -\beta^{-1} \log Z_\beta = -T \log Z_\beta. \end{aligned} \quad (32)$$

The first term on the left is the ensemblist energy, namely, its expected value, while the second term is minus the temperature T times the Gibbs entropy

$$S_G = S_G(\varrho_\beta) = - \int_\Gamma \varrho_\beta \log \varrho_\beta dX = - \left\langle \log \frac{e^{-\beta H}}{Z_\beta} \right\rangle_\beta. \quad (33)$$

(Note that it follows from (29) and (32) that for a large system, in equilibrium at temperature T , the Boltzmann and Gibbs entropies are more or less the same.)

Similarly, for the grand canonical partition function

$$Z_{\beta,\mu} = \int_{\bigcup_N \Gamma^{(N)}} e^{-\beta(H-\mu N)} dX, \quad (34)$$

the normalization constant in (26), we have that

$$Z_{\beta,\mu} = \sum_N \int_{E_{min}^{(N)}}^\infty e^{-\beta(E-\mu N)} |\Gamma_E^{(N)}| dE = \sum_N \int_{E_{min}^{(N)}}^\infty e^{-\beta(E-\mu N)} e^{S_B(E,N)} dE \quad (35)$$

with the integrand concentrated near $(E_{\beta,\mu}, N_{\beta,\mu})$ with small spread in both. We thus have that

$$\begin{aligned} \log Z_{\beta,\mu} &\simeq -\beta(E_{\beta,\mu} - \mu N_{\beta,\mu}) + S_B(E_{\beta,\mu}, N_{\beta,\mu}) \\ &= -\beta \left(E_{\beta,\mu} - T_\beta S_B(E_{\beta,\mu}, N_{\beta,\mu}) - \mu N_{\beta,\mu} \right), \end{aligned} \quad (36)$$

with the quantity inside $\left(\right)$ looking like \mathcal{G} . Maximizing the exponent on the right hand side of (35), or, equivalently, minimizing $E - T_\beta S_B(E, N) - \mu N$, we find that

$$\partial S_B / \partial E = 1/T_\beta \text{ and } \partial S_B / \partial N = -\mu/T_\beta, \quad (37)$$

so that β and μ are indeed the quantities appropriate to the thermodynamic formalism.

It follows from (36) that, with negligible error for large systems, we have the following formulas for the individualist grand canonical potential and the pressure:

$$\mathcal{G} = -T \log Z_{\beta,\mu} \quad (38)$$

and

$$p = T(\log Z_{\beta,\mu})/V. \quad (39)$$

Moreover, just as $-T \log Z_\beta$ is exactly the ensemblist free energy, so too are $-T \log Z_{\beta,\mu}$ and $T(\log Z_{\beta,\mu})/V$, respectively, exactly the ensemblist grand canonical potential and pressure.

From the very form of the canonical and grand canonical ensembles, it follows that taking logarithmic derivatives yields the expected values of basic macrovariables: For example, from (25), (26), (27), and (34) we have that

$$\partial \log Z_\beta / \partial \beta = \frac{\partial Z_\beta / \partial \beta}{Z_\beta} = \frac{-\int H e^{-\beta H} dX}{Z_\beta} = -\langle H \rangle_\beta, \quad (40)$$

and

$$\partial \log Z_{\beta,\mu} / \partial \mu = \beta \langle N \rangle_{\beta,\mu}. \quad (41)$$

Thus

$$\partial \log Z_\beta / \partial T = (\partial \log Z_\beta / \partial \beta) \partial \beta / \partial T = \langle H \rangle_\beta / T^2, \quad (42)$$

so that, from (31),

$$\partial F / \partial T = -\log Z_\beta - \langle H \rangle_\beta / T = -S_G \simeq -S_B \quad (43)$$

using (32). Similarly, from (38),

$$\partial \mathcal{G} / \partial \mu = -T \beta \langle N \rangle_{\beta,\mu} = -\langle N \rangle_{\beta,\mu} \simeq -N_{\beta,\mu}. \quad (44)$$

We thus obtain the first equation of (19), and the last equation of (21), exactly as demanded by the thermodynamic formalism.

4 Summary

Here is the overall picture that we obtain: The thermodynamic formalism involves relations among a variety of macrovariables and intensive variables, with partial derivatives playing an important role. These relations are formally modeled by the ensembles, with the partition functions, parameters such as μ and β , and expected values

with respect to the ensembles yielding the mock thermodynamic quantities. Even for small systems one has, I believe, a mock thermodynamics. But for large systems, because of the law of large numbers, the idealized mock thermodynamics becomes an enormously good approximation to the exact relations obeyed by the physical variables (individualist variables) with which we are primarily concerned in physics.

The exact relationships among these variables are presumably extremely complicated. Their ensemblist approximations, however, are much more simply related. Thus does the Gibbsian ensemblist approach yield a clean idealization describing, for large systems, the individualist relations in which we are primarily interested.

5 Conclusion

I began by mentioning a similarity between the foundations of statistical mechanics and the foundations of quantum mechanics: that they are both rather controversial. Here is another similarity. Most physicists don't have a clear understanding of what they themselves really think about these foundational issues. When it comes to quantum mechanics, by far most would say they accept the Copenhagen interpretation. For statistical mechanics, most would insist on the fundamentality of ensembles.

One reason this might be so is this. Physicists tend to be busy people; they want and need to obtain practical results about experiments and observable phenomena, and in order to do so they use the best tools at their disposal. For quantum mechanics, the tools are those of the quantum formalism, i.e., what they think of as the Copenhagen interpretation. For statistical mechanics, the tools are ensemblist tools, pioneered by Gibbs and, some would argue, also by Boltzmann. So it is quite natural for physicists to imagine that they are Copenhagenists and ensemblists.

It is of little practical value—for their research and for their professional advancement—for physicists to worry about why their tools work as well as they do and what those tools actually have to do with the reality that lies beneath observation. Nonetheless, I do think physicists should do a bit better in this regard. And philosophers of physics should as well: they should pay more careful attention to why physicists do in fact express the views that they do.

References

- [1] J. von Neumann: *Mathematische Grundlagen der Quantenmechanik* (Springer Verlag, New York-Heidelberg-Berlin 1932); English translation by R.T. Beyer, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, N.J., 1955)

- [2] J.W. Gibbs: *Elementary Principles in Statistical Mechanics* (Yale University Press, New Haven 1902) Chapter XII
- [3] E. Schrödinger: *What is Life? and Other Scientific Essays* (Doubleday Anchor Books, New York 1965) section 6
- [4] E. Schrödinger: *What is Life? : The Physical Aspect of the Living Cell with Mind and Matter & Autobiographical Sketches* (Cambridge University Press, Cambridge 1992)
- [5] L. Boltzmann: *Annalen der Physik* **57**, 773 (1896); reprinted and translated as Chapter 8 in [7]
- [6] L. Boltzmann: *Annalen der Physik* **60**, 392 (1897); reprinted and translated as Chapter 10 in [7]
- [7] S.G. Brush: *Kinetic Theory* (Pergamon, Oxford 1966)