

# On the foundations of statistical mechanics: ergodicity, many degrees of freedom and inference

Sergio Chibbaro<sup>1,6</sup> and Lamberto Rondoni<sup>2,3,6</sup> and Angelo Vulpiani<sup>4,5,6</sup>

<sup>1</sup> Sorbonne Universités, UPMC Univ Paris 06,  
CNRS, UMR7190, Institut Jean Le Rond d'Alembert,  
F-75005 Paris, (France)

<sup>2</sup> Dip. Scienze Matematiche, Politecnico di Torino,  
C. Duca degli Abruzzi 24, Torino.

<sup>3</sup> INFN, Sezione di Torino, Via P. Giuria 1, Torino (Italy).

<sup>4</sup> Dipartimento di Fisica, Università "La Sapienza" Roma (Italy)

<sup>5</sup> ISC-CNR, Piazzale A. Moro 2 - 00185 Roma (Italy)

<sup>6</sup> Kavli Institute for Theoretical Physics China, CAS, Beijing 100190, China

March 9, 2014

## 1 Introduction

Statistical mechanics has been founded by Maxwell, Boltzmann and Gibbs who intended to describe in terms of microscopic interactions the properties of macroscopic objects, *i.e.* systems made of so many particles that a statistical treatment is required. For almost all practical purposes one can say that the whole subject of statistical mechanics consists in evaluating a few suitable functions and quantities, such as the partition function, the free energy and the correlation functions.

Since the discovery of deterministic chaos it is plain that statistics (e.g. over ensembles of objects following the same dynamical rules, or over time series of fluctuating quantities) is unavoidable and useful even in systems with a few degrees of freedom. On the other hand, no universal agreement has been reached so far about the fundamental ingredients for the validity of statistical mechanics [1, 2].

The wide spectrum of positions includes Landau's and Khinchin's belief that the main ingredient in producing the common macroscopic behaviours is the vast number of microscopic degrees of freedom, which makes (almost) completely irrelevant details of the microscopic dynamics such as the validity of the standard notion of ergodicity, developed in the mathematical literature. But it includes also the opinions of those, like Prigogine and his school, who consider chaos, in the sense of the positivity of Lyapunov exponents, as a fundamental ingredient

[3], although this (to us untenable) position has been strongly criticized and dismissed in numerous essays, *e.g.* Ref.[4].

On the other hand, for decades, the ergodic problem has been often completely neglected and the (so-called) Gibbs *ensemble* approach has been accepted on the grounds that “it works”. Moreover, this point of view has been commonly associated with supposedly universally acceptable principles, such as the maximum entropy principle (MEP), according to which, in good substance, all the statistical properties of a given object can be inferred simply by maximizing a kind of information theoretic entropy, under the constraints that are known to act on the system itself, [5]. But also this approach is hardly satisfactory, especially if one thinks that understanding the mechanisms underlying macroscopic physics is not less important than computing useful quantities.

The present paper is meant to give a simple introduction to the problem of the connection between microscopic dynamics and statistical laws. For sake of simplicity, we mostly refer to non-dissipative dynamics, since dissipation adds technical difficulties to the conceptual issues, although part of our discussion extends beyond this limit. In particular, the relevance of chaos and ergodicity is here confronted with that of the large number of degrees of freedom. In Section 2, we review the microscopic connection, along the lines of Boltzmann’s approach, and of its further developments. In Section 3, we discuss the falsifiability of statistical mechanics and its role as statistical inference. In Section 4, we draw our conclusions.

## 2 From microscopic to macroscopic

Once the atomistic view is accepted, the identification of the issues that may benefit from it still poses a challenge. No doubt, matter is made of atoms, and what happens to any piece of matter depends on what happens to its atoms; but which aspects of the behaviour of that object may be understood in terms of what we know about atoms? Which phenomena may be elucidated by the laws which we use to describe the behaviour of atoms? The fact is that our description of the atomic world is a direct extension of classical (Newtonian) mechanics,<sup>1</sup> which has been very successfully developed to describe systems made of a small number of objects, while the properties of matter result from the cooperation of exceedingly large numbers of atoms. At the fundamental level the main problem of statistical mechanics is to establish a link between the thermodynamics description and the dynamics ruling the time evolution of the system. Some books devoted to foundational aspects of statistical mechanics are the following: [1, 2, 7].

To treat the main conceptual problems of statistical physics, let us start from classical mechanics, considering the space  $\mathcal{M}$  of all microscopic configurations of a given system of interest, known as the *phase space*. In the case of the simple monatomic gases,  $\mathcal{M}$  may be taken to be the  $6N$  dimensional phase space  $\mathbb{R}^{6N}$ , i.e. the space of all coordinates and momenta  $\Gamma = (\mathbf{q}, \mathbf{p})$  of the  $N$

---

<sup>1</sup>Conceptually, quantum mechanics does not change the picture [6].

atoms. If the elementary constituents of the system cannot be approximated by point particles, but still obey classical mechanics, one would have to refer to a larger phase space, which has more dimensions, but the basic idea does not change. Given a vector  $\Gamma \in \mathcal{M}$  representing an initial microstate, the state at a later time may be denoted by the evolved vector  $S^t\Gamma \in \mathcal{M}$ , corresponding to the changes in positions and velocities of the atoms, where  $S^t$  is the evolution operator for a time  $t$ .

In the case of an isolated system of particles, this evolution is expressed by Hamilton's equations of motion, which can be compactly written as:

$$\dot{\mathbf{q}} = \frac{\partial H}{\partial \mathbf{p}} ; \quad \dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{q}} \quad (1)$$

where  $H$  is the Hamiltonian of the system, and  $H(\mathbf{q}, \mathbf{p}) = E$  determines the energy surface in which the time evolution  $S^t\Gamma$  is confined for all  $t$ . Then, one assumes that each microstate  $\Gamma \in \mathcal{M}$  implies a particular value  $\mathcal{O}(\Gamma) \in \mathbb{R}$  for each observable quantity  $\mathcal{O}$ .

Macroscopic measurements do not occur instantaneously, but take a certain amount of time: as a consequence of the very rapid and perpetual motion of the atoms, they do not yield a precise value  $\mathcal{O}(\Gamma)$ , because  $\Gamma$  changes in that time. Instead, a measurement yields the average of the values that  $\mathcal{O}(S^t\Gamma)$  takes while  $S^t\Gamma$  explores the phase space. For instance, the pressure of a gas measured by a manometer is determined by the average of the variations of the molecular momenta, which occur when the molecules hit the surface of the sensor.

## 2.1 Boltzmann's grand vision

A mathematical expression for a macroscopic measurement, i.e. for the average over the myriad microscopic events performed by our senses or by our measurement tools, is given by:

$$\overline{\mathcal{O}}^T(\Gamma) = \frac{1}{T} \int_0^T \mathcal{O}(S^t\Gamma) dt , \quad (2)$$

where  $T$  is the duration of the measurement and  $\Gamma$  is the initial microstate.

In general, the result of a measurement could depend on both  $T$  and on  $\Gamma$ . The dependence on  $T$  makes the result of the measurement subjective, as  $T$  may be varied at will by the experimentalist, while the dependence on  $\Gamma$  makes it stochastic, because the initial microstate can neither be controlled nor identified. For a scientific theory to be conceived, these dependencies must be kept under control.

In particular, for the connection with thermodynamics, the microscopic events must occur on time scales much shorter than the observation scales. If that is the case, one *may hope* that the quantity  $\mathcal{O}(S^t\Gamma)$  explores the interval of all its possible values so rapidly that the time required by a measurement makes  $\overline{\mathcal{O}}^T(\Gamma)$  indistinguishable from its asymptotic value almost independently of  $\Gamma$ , washing away subjectivity and stochasticity. The appropriate value of  $T$

depends on the sensitivity of the measurement tool: it is shorter for higher sensitivity, but it nevertheless remains virtually infinite with respect to molecular characteristic times, such as the average time between two molecular collisions. Therefore, a theory of measurements relies on the existence of the following limit:

$$\overline{\mathcal{O}}(\Gamma) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \mathcal{O}(S^t \Gamma) dt \quad (3)$$

and on the irrelevance of  $\Gamma$  for its value. Will that ever be the case for given microscopic dynamics?

The existence of the time limits was proven by Birkhoff in 1931 for a wide class of systems, while the dependence on  $\Gamma$  poses many more problems, from a mathematical perspective. Indeed, if thermodynamics applies, the properties of a macroscopic system can be predicted without any knowledge of  $\Gamma$ . This is essentially the content of Boltzmann's celebrated *ergodic hypothesis*. Roughly speaking, in its modern standard formulation, this hypothesis states that each hyper-surface with constant energy is completely accessible to every motion with the given energy. As a corollary, for each trajectory, the residence-time in any region is proportional to the volume of the region.

When the ergodic hypothesis holds,  $\mathcal{M}$  may be endowed with a probability density  $\rho$  reflecting the frequency with which the different regions of  $\mathcal{M}$  are visited during the evolution. This probability distribution, also called *ensemble* [8], associates higher weights to the regions which are visited more often, and lower weights to those which are visited less frequently so that:

$$\overline{\mathcal{O}}(\Gamma) = \int_{\mathcal{M}} \mathcal{O}(\Gamma) \rho(\Gamma) d\Gamma \equiv \langle \mathcal{O} \rangle_{\rho} \quad (4)$$

for all initial conditions  $\Gamma \in \mathcal{M}$ , except a set of vanishing probability. The quantity  $\langle \mathcal{O} \rangle_{\rho}$  is called the phase average, and a given system is called *ergodic* with respect to the probability distribution  $\rho d\Gamma$  if the time averages of *all* its observables obey Eq.(4).<sup>2</sup>

To explain the physical meaning of Eq.(4), in his well known book on thermodynamics, Fermi says that: *Studying the thermodynamical state of a homogeneous fluid of given volume at a given temperature (the pressure is then defined by the equation of state), we observe that there is an infinite number of states of molecular motion that correspond to it. With increasing time, the system exists successively in all these dynamical states that correspond to the given thermodynamical state. From this point of view we may say that a thermodynamical state is the ensemble of all the dynamical states through which, as a result of the molecular motion, the system is rapidly passing.*<sup>3</sup>

<sup>2</sup>More precisely, given an invariant probability measure  $\mu$  on  $\mathcal{M}$ , the statement must hold for all functions  $\mathcal{O}$  which constitute a suitable function space, e.g. Birkhoff chose  $L^1(\mathcal{M}, \mu)$ . Also, note that the equality  $\overline{\mathcal{O}}(\Gamma) = \langle \mathcal{O} \rangle_{\mu} = \int_{\mathcal{M}} \mathcal{O}(\Gamma) \mu(d\Gamma)$ , for  $\mu$ -almost-all  $\Gamma$ , defines ergodicity in full generality, hence even for the singular measures  $\mu$  required to describe steady states of dissipative dynamics.

<sup>3</sup>Fermi's statement must not be interpreted literally: it must be intended, as above, in terms

The practical importance of the ergodic hypothesis is evident: knowledge of the initial microstate  $\Gamma$  and of the solution of Hamilton's equations, which yields  $S^t\Gamma$ , are unnecessary if the system is ergodic.

Given all this, a question arises: how do the averages of the microscopic properties of a system of particles relate to the thermodynamic properties of the object composed of those particles? The first, in fact, consist of a myriad of mechanical parameters, such as the particles' energy and momentum, while the others consist of just a few measurable quantities, like temperature and pressure.

The conceptual (and technical) law which associates thermodynamics with the classical mechanics of atoms was proposed by Boltzmann and it is engraved in his tombstone:

$$S = k \log W . \quad (5)$$

This celebrated relation connects the thermodynamic entropy  $S$  of an object in the macroscopic state  $X$ , to the volume  $W$  of all microstates in  $\mathcal{M}$  which correspond to the same  $X$ . For example, considering the macrostate  $X$  corresponding to a given energy  $E$ , one typically considers the energy shell  $E - \delta E \leq H(\mathbf{q}, \mathbf{p}) \leq E$ , with small  $\delta E$ , and obtains:

$$W = \int_{E - \delta E \leq H(\mathbf{q}, \mathbf{p}) \leq E} d\mathbf{q}d\mathbf{p} .$$

The microcanonical probability distribution is constant in the energy shell so that:

$$\rho(\Gamma) = \begin{cases} \frac{1}{W} & \text{if } E - \delta E \leq H(\mathbf{q}, \mathbf{p}) \leq E \\ 0 & \text{otherwise} \end{cases}$$

In the terms of philosophy of science, Equation (5) qualifies as a bridge law[9], because  $S$  is a thermodynamic quantity, while  $W$  is a microscopic entity. Once it has been introduced, further mechanical properties of our description of the microscopic dynamics may be related to as many other thermodynamic quantities, thus bridging the gap between micro- and macro-descriptions. In particular, in the *entropy representation* of thermodynamics, one obtains the temperature as:

$$\frac{1}{T} = \frac{\partial S}{\partial E} , \quad (6)$$

and the free energy as  $F = E - TS$ , where  $E$  is the mechanical energy of the system. Note that some authors, including Nagel[9], take the relation between the temperature  $T$  and the average kinetic energy  $K$  as the bridge law. This is incorrect, as thermodynamics requires a thermodynamic potential, such as the free energy, which is a function of the other relevant thermodynamic variables. In addition, the expression of  $T$  in terms of  $K$ , which had already been guessed by Bernoulli, holds only for a special class of phenomena.

---

of the values that the observables of interest can take, and not in terms of trajectories  $\{S^t\Gamma\}$  densely exploring the phase space. This would require completely unphysical exceedingly long time scales.

The bridge law (5) has important consequences supported by empirical evidence, including, in particular, those derived from Einstein's version:

$$P(\alpha) \sim e^{(S(\alpha) - S_{\text{eq}})/k} \quad (7)$$

where, apart from the normalization constant,  $W$  has turned into the probability  $P(\alpha)$  of the collection of microstates which correspond to the same value for the macroscopic variable  $\alpha$ ,  $S_{\text{eq}}$  is the equilibrium entropy and  $S(\alpha)$  is the entropy of a spontaneous fluctuation of  $\alpha$ , produced by the cooperation of many microscopic motions. Formula (7) is meant to represent the probability of fluctuations about equilibrium states of microscopic mechanical quantities, such as the energy  $E$  of a system in contact with a heat bath at temperature  $T$ . As these fluctuations are related to observable quantities, they can be characterised by macroscopic equilibrium experiments. For instance, if  $\alpha$  is the energy, one identifies the average  $\langle E \rangle$  of the energy with the internal energy  $U$  of the system, and introduces the standard deviation  $\sqrt{\langle E^2 \rangle - \langle E \rangle^2}$ , which measures the size of the fluctuations. We obtain:

$$\langle E^2 \rangle - \langle E \rangle^2 = kT^2 C_v \quad (8)$$

where  $C_v$  is the heat capacity at constant volume. Because  $C_v$  is extensive, hence proportional to the number  $N$  of particles in the system, the relative size of the energy fluctuations is negligible in large systems:

$$\frac{\sqrt{\langle E^2 \rangle - \langle E \rangle^2}}{\langle E \rangle} \sim O\left(\frac{1}{\sqrt{N}}\right) \rightarrow 0, \quad \text{for } N \rightarrow \infty \quad (9)$$

In other words, the fluctuations of the microscopic mechanical quantity  $E$  grow with the system size proportionally to  $\sqrt{N}$ , but are negligible with respect to the observable internal energy  $\langle E \rangle = U$ , which is of order  $O(N)$ . This should not lead one to relegate fluctuations to the set of only marginally interesting phenomena. Indeed, in his search of an ultimate proof of the existence of atoms, Einstein realised that Eq.(8):

*would yield an exact derivation of the universal constant [ $k$  or, equivalently, Avogadro's number  $N_A$ ] if it were possible to determine the average of the square of the energy fluctuations of the system.*

He successfully applied this idea to describe the Brownian motion[10].

## 2.2 Beyond the mathematical limitations of ergodic theory

The effort to turn the physical notion of measurement into an appropriate mathematical one has led to the issue of ergodicity, which seems to cleverly frame the connection between mechanical and thermodynamical quantities. Unfortunately, apparently modest consideration of real-life systems, such as the insensitivity of thermodynamic quantities to microscopic states, raises deep mathematical questions. In particular, one faces the problem of identifying the probability

density  $\rho$  that describes systems in equilibrium, or evolving towards equilibrium, and the fact that requiring Eq.(4) to hold for *all* phase functions is too demanding, compared to the needs of thermodynamics.<sup>4</sup> In addition, the time scales over which the ergodicity of a system of many degrees of freedom would be obtained are astronomically larger than the physically relevant time scales.

The celebrated work by Fermi, Pasta and Ulam, concerning a chain of non-linear oscillators, further showed that ergodicity may be violated even by the simplest particles systems. Indeed, in their numerical simulations, known as the FPU experiment, Fermi and coworkers showed that a typical Hamiltonian system is not ergodic. This fact was totally unexpected, at that time, and was only later explained in the sophisticated mathematical terms of KAM theory [11].

Quite surprisingly, the very simple probability distributions known as *microcanonical*, *canonical* and *grand-canonical* ensembles describe very well most equilibrium situations. When legitimate, this is an extremely powerful way of proceeding, whose success seems to rest on our limited knowledge of the microscopic dynamics. However one should be wary of possible misunderstandings. In particular, ensembles are often described as *fictitious collections of macroscopically identical copies of the object of interest, whose microstates differ from each other*. While this maybe a convenient perspective, one should not forget that their purpose is to describe the properties of a single system, whose microstate evolves forever. We can say that the word “statistical ensemble” is nothing but a way to indicate the probability density of  $\Gamma$ .

But why? The foundations of the ergodic hypothesis look shaky, and its success puzzling, if no further explanation is given. To address these issues, Khinchin pioneered an approach based on the following premises[12]:

- a) statistical mechanics concerns systems with a large number of degrees of freedom;
- b) the physical observables are but a few and quite special functions;
- c) it is physically acceptable that ensemble averages do not coincide with time averages, on a small set of phase space trajectories;

As appropriate for rarefied systems, he considered dynamics whose Hamiltonians are the sum of single particle contributions:

$$H = \sum_{n=1}^N H_n(\mathbf{q}_n, \mathbf{p}_n)$$

and restricted the space of observables to the *sum functions* – functions defined as sums of single particle contributions  $f_n$ :

$$f(\Gamma) = \sum_{n=1}^N f_n(\mathbf{q}_n, \mathbf{p}_n) .$$

---

<sup>4</sup>We state that a physical system is in *equilibrium* if all currents – of mass, momentum, heat, etc. – vanish, and the system is uniquely described by a (typically quite small) set of *state variables* which do not change with time.

The pressure and the kinetic energy are examples of such functions. Then, denoting by  $\langle \cdot \rangle$  the microcanonical ensemble average, Khinchin demonstrated that:

$$\text{Prob} \left( \frac{|\bar{f} - \langle f \rangle|}{|\langle f \rangle|} \geq K_1 N^{-1/4} \right) \leq K_2 N^{-1/4} ,$$

where  $K_1$  and  $K_2$  are constants. This means that the microcanonical averages of sum functions differ from their time averages by more than a (small) relative tolerance only along a set of trajectories whose probability vanishes in the  $N \rightarrow \infty$  limit. The problem is that the initial conditions of the microstate must be taken within a proper subset of the phase space, but Khinchin showed that the fraction of volume of phase space which lies outside this subset vanishes in the  $N \rightarrow \infty$  limit.

Ultimately, from various standpoints, Khinchin's theory ascribes the good statistical properties required for normal thermodynamic behaviour to the fact that  $N$  is very large. From his perspective, the details of the microscopic dynamics appear practically irrelevant for the physics of rarefied gases. An important extension of this approach, which goes beyond the low density gas, was obtained by Mazur and van der Linden, [13]. These authors did not require the Hamiltonian to be separable in single particle contributions, but admitted particles to interact only through short-range interaction potentials and, like Khinchin, considered only sum variables. They proved that their systems can be treated as consisting of many non-interacting parts.

Although even this theory is not completely satisfactory, because the set of sum variables is too limited for dense systems, the works of Khinchin and of Mazur and van der Linden clarify why one should not be surprised that the ergodic hypothesis applies so generally in physics: macroscopic systems are made of very many particles.

### 3 Remarks

Let us conclude with some provocative remarks and a brief summary.

#### 3.1 Is Statistical Mechanics falsifiable?

The question may seem provocative but deserves a brief excursion.

In statistical mechanics tests, the typical exercise that students have to solve is the following: given a box of volume  $V$  containing  $N$  particles interacting according to a known potential  $U(|\mathbf{q}_i - \mathbf{q}_j|)$ , assuming thermal equilibrium at temperature  $T$ , is it possible to calculate specific heat, state equation, etc.? For instance, for dilute gases it is possible to write the virial expansion

$$\frac{p}{k_B T} = \rho + b_2(T)\rho^2 + b_3(T)\rho^3 + \dots \quad (10)$$

where  $\rho = N/V$  and virial coefficients  $b_1, b_2, \dots$  can be expressed in terms of  $U(r)$ . Calculations should be then compared against experiments.

In practice, however, this approach cannot be pursued. Even if the dynamics is classical, the potential  $U(r)$  has quantum origins and is not known. It should be computed from the Schrödinger equation, but a good approximation can only be obtained in a few lucky cases. In general, it is necessary to tackle the problem in a very different way.

First of all, one chooses a given form for  $U(r)$ , for instance, for simple liquids the Lennard-Jones potential is very popular:

$$U(r) = 4\epsilon \left[ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^6 \right], \quad (11)$$

where  $\epsilon$  and  $\sigma$  are two parameters. Then, one computes the virial coefficients  $b_2(T), b_3(T), \dots$  as functions of  $U(r)$ , i.e. of the parameters  $\epsilon$  and  $\sigma$  and, finally,  $\epsilon$  and  $\sigma$  are determined by comparison with the experimental data [14].

Since this procedure is self-consistent, one could conclude that statistical mechanics cannot be falsified. While this discussion sheds light on the complex relation between microscopic and macroscopic worlds, the conclusion is incorrect. Indeed, statistical mechanics predicts many other non-trivial properties that can be checked experimentally. Among the others, the Maxwell-Boltzmann distribution for the molecules, phase transitions and universality of critical phenomena[15, 16]. Interestingly, in all these cases, the knowledge of the potential  $U(r)$  is not needed: the Maxwell-Boltzmann distribution is completely independent of  $U(r)$ , whereas critical phenomena are influenced only by some qualitative properties of this potential, which delimit the universality class.

### 3.2 Statistical mechanics as statistical inference?

According to a radically anti-dynamical point of view, statistical mechanics is but a form of statistical inference rather than a theory of objective physical reality. Under this light, probabilities measure the degree of truth of a logical proposition, rather than describing the state of a system.

In this context, Jaynes proposed the MEP as a general rule for finding the probability of a given event when only partial information is available. If the mean values of  $m$  independent functions  $f_i(\mathbf{x})$  are given:

$$c_i = \langle f_i \rangle = \int f_i(\mathbf{x}) \rho(\mathbf{x}) d\mathbf{x} \quad i = 1, \dots, m ,$$

the MEP rule determines the probability density  $\rho$  of the events compatible with these mean values, by maximising the “entropy”

$$H = - \int \rho(\mathbf{x}) \ln \rho(\mathbf{x}) d\mathbf{x} ,$$

under the constraints  $c_i = \langle f_i \rangle$ . Using the Lagrange multipliers one easily obtains

$$\rho(\mathbf{x}) = \frac{1}{Z} \exp \sum_{i=1}^m \lambda_i f_i(\mathbf{x})$$

where  $\lambda_1, \lambda_2 \dots \lambda_m$  depend on  $c_1, c_2, \dots, c_m$ . For instance, for systems with a fixed number of particles subjected to the unique constraint that their mean energy is fixed, the MEP leads to the canonical distribution in a very simple fashion.

As a technical but rather important detail, we note that this holds only if  $\mathbf{x}$  is the vector of the canonical coordinates (i.e. positions and momenta of the particles). Analogously, for systems of varying numbers of particles, the grand canonical distribution is obtained by additionally constraining the mean number of particles. Many find in these facts an unquestionable proof of the validity of the MEP, but we will see that it is just a matter of fortunate coincidence, related to the choice of canonical coordinates.

The most frequent objection to the MEP is summarised by the maxim *Ex nihilo nihil*, or citing the title of one of Peres's papers, "Unperformed experiments have no results" [17], i.e. our ignorance cannot be credited for inferences about real phenomena, [15]. In spite of the optimistic claims of the MEP enthusiasts, to the best of our knowledge MEP has only produced different, sometime more elegant, derivations of previously known results[16]. Apart from this very general observation, the weakest technical aspect of the MEP approach is the dependence of the results on the choice of the variables.

For simplicity's sake, consider a scalar random variable  $X$ , ranging over a continuum, whose probability distribution function is  $p_X$ . It is easy to realise that the "entropy"  $H_X = - \int p_X(x) \ln p_X(x) dx$  is not an intrinsic quantity of the phenomena concerning  $X$ . With a different parametrisation, i.e. using the coordinates  $y = f(x)$  with an invertible function  $f$ , rather than  $x$ , the entropy of the same phenomenon would now be given by

$$H_Y = - \int p_Y(y) \ln p_Y(y) dy$$

with  $p_Y(y) = p_X(f^{-1}(y))/|f'(x = f^{-1}(y))|$ . Therefore, one has

$$H_Y = H_X + \int p_X(x) \ln |f'(x)| dx$$

The MEP gives different solutions if different variables are adopted to describe the very same phenomenon.

In order to avoid this dependence on the choice of variables, Jaynes later proposed a more sophisticated version of the MEP, in terms of the relative entropy:

$$\tilde{H} = - \int \rho(\mathbf{x}) \ln \left[ \frac{\rho(\mathbf{x})}{q(\mathbf{x})} \right] d\mathbf{x} ,$$

where  $q$  is a known probability density. Of course,  $\tilde{H}$  depends on  $q$ ; but, at variance with the entropy, it does not depend on the chosen variables. On the other hand, one must decide how to select  $q$ , and this issue is equivalent to the problem of choosing the "proper variables". Therefore, even this more elaborate method is non-predictive, and we see no reason to pursue the MEP approach further in the field of statistical mechanics. The interested reader is referred to the existing extensive literature, e.g. [18, 19], for further details.

## 4 Summary

Let us summarise the fundamental concepts discussed above. The purpose of Statistical Mechanics is the understanding of macroscopic behaviour in terms of the properties of the microscopic constituents of matter. This has been achieved adopting:

- the ergodic hypothesis
- the Boltzmann's principle (5) as a bridge law,
- the huge number of degrees of freedom ( $N \gg 1$ )

However, strictly speaking, the ergodic hypothesis cannot be verified, except in a few exceptional cases, as evidenced by the FPU numerical experiment and by the KAM theorem. On the other hand, Khinchin's strong mathematical results showed that Eq.(4) holds in the  $N \rightarrow \infty$  limit, for a physically relevant class of observables. Then, the success of Statistical Mechanics in describing macroscopic systems may be attributed to the following facts:

- although ergodicity is not exactly verified in realistic models, it does hold in a weak sense, which is sufficient for the purposes of physics;
- the bridge law  $S = k \log W$  links the microscopic mechanical quantity  $W$  with the emerging thermodynamic quantity  $S$ , through the Boltzmann constant  $k$ . We stress the fact that the relation (5) is a fundamental assumption of the same nature (and importance) as Newton's principles for mechanics.
- macroscopic objects are made of very large numbers of microscopic constituents. The number of particles in macroscopic bodies is of the order of the Avogadro's number ( $N_A \approx 6.02 \cdot 10^{23}$ ). Boltzmann's constant  $k = \mathcal{R}/N_A$ , where  $\mathcal{R}$  is the universal gas constant that takes the mind-boggling value  $k \approx 1.38 \cdot 10^{-23} J/K$ , is an astonishingly physically powerful element of the bridge law. Because it constitutes a unit of entropy (energy divided by temperature),  $k$  binds mechanics and thermodynamics together. The very small numerical value of  $k$  measures the "distance" between the microscopic world and the macroscopic world.
- Chaos is not a fundamental ingredient for the validity of statistical mechanics: the naive idea that chaos implies good statistical properties is inconsistent[7, 20]. Sometimes, even in the absence of chaos one can have good agreement between the time averages and their values predicted by statistical mechanics; such a result is in agreement with Khinchin's ideas[20].
- How large  $N$  must be for the framework outlined above to properly describe physical systems is hard to tell. No doubt,  $N = N_A$  is largely sufficient for the vast majority of equilibrium phenomena, meaning phenomena on the scale of our daily life. Reducing the values of  $N$ , violations

of the thermodynamic behaviours are bound to be met, but when that happens depends very strongly on the systems and on the observables under consideration. One may have thermodynamic-like behaviours at the nanometric scale for certain observables, but not for others. A thumb rule is that fluctuations must be negligible compared to the averages, but the onset of the non-thermodynamic behaviours is still largely to be understood. This is why the statistical physics of small systems is a thriving subject.

## References

- [1] L. Sklar. *Physics and chance*. Cambridge University Press, 1995.
- [2] G. Emch and C. Liu. *The Logic of Thermo-Statistical Physics*. Springer Verlag, 2002.
- [3] I. Prigogine and I. Stengers. *Order out of Chaos*. Bantam Books, Toronto-New York-London-Sydney., 1984.
- [4] J. Bricmont. Science of chaos or chaos in science? *Annals of the New York Academy of Sciences*, 775(1):131, 1995.
- [5] S. Pressé, K. Ghosh, J. Lee, and K. A. Dill. Principles of maximum entropy and maximum caliber in statistical physics. *Rev. Mod. Phys.*, 85(3):1115, 2013.
- [6] J.L. Lebowitz. Boltzmann's entropy and time's arrow. *Physics Today*, 46:32, 1993.
- [7] P. Castiglione, M. Falcioni, A. Lesne, and A. Vulpiani. *Chaos and coarse graining in statistical mechanics*. Cambridge University Press, 2008.
- [8] E. Fermi. *Thermodynamics*. Dover Publications, New York, 1956.
- [9] E. Nagel. *The structure of science*. Hackett Publishing Company, 1979.
- [10] A. Einstein. *Investigations on the Theory of the Brownian Movement*. Dover Publications, 1956.
- [11] M. Cencini, F. Cecconi, and A. Vulpiani. *Chaos: from simple models to complex systems*. World Scientific, Singapore, 2009.
- [12] A I Khinchin. *Mathematical Foundations of Statistical Mechanics*. Dover Publications, 1949.
- [13] P. Mazur and J. van der Linden. Asymptotic form of the structure function for real systems. *Journal of Mathematical Physics*, 4:271, 1963.
- [14] P. Attard. *Thermodynamics and Statistical Mechanics*. Academic Press, 2002.

- [15] S. K. Ma. *Statistical Mechanics*. World Scientific, Singapore, 1985.
- [16] L. Peliti. *Statistical Mechanics in a Nutshell*. Princeton University Press, 2011.
- [17] A. Peres. Unperformed experiments have no results. *Am. J. Phys*, 46(7):745, 1978.
- [18] E.T. Jaynes. Foundations of probability theory and statistical mechanics. In M. Bunge, editor, *Delaware Seminar in the Foundations of Physics*, volume 1, page 76. Springer-Verlag, Berlin, 1967.
- [19] J. Uffink. Can the maximum entropy principle be explained as a consistency requirement? *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 26(3):223, 1995.
- [20] R. Livi, M. Pettini, S. Ruffo, and A. Vulpiani. Chaotic behavior in nonlinear hamiltonian systems and equilibrium statistical mechanics. *Journal of Statistical Physics*, 48:539, 1987.