Statistical ensembles, entropy and probability in statistical mechamics

Personne n'ignore que la chaleur peut être la cause du mouvement, qu'elle possède même une grande puissance motrice: les machines à vapeur, aujourd'hui si répandues, en sont une preuve parlant à tous les yeux.[1]

Thus S. Carnot (1824): and his theorem, on maximal efficiency of a thermal machine, led Clausius, 40y later, to II principle and definition of *entropy*, its essence, [2, 3, 4].

Immediately atomists (atomism was still debated btwn Chemists and Physicists) looked for the mechanical meaning of the new observable, Boltzmann (1866).

Boltzmann supposed that the atoms (of a gas in a container, say) moved as a whole periodically, at least with good approximation.[5]

He shewed that the principle of least action (extended to apply to periodic motions) implies existence of a function of the state (*i.e.* of the periodic orbit) whose variations depend only on the initial and final states of a tranformation.

A concrete example is a 1-dim system with Hamiltonian $U = \frac{1}{2}p^2 + \Phi(q)$. The least action principle shows that if the periodic orbit is changed from O to O' accompanied by a change δF in Φ and δU in U due to a change in few parameters $d\alpha, d\beta, dV, ..$, let

$K =$ (time) average kinetic energy in O,
$p =$ (time) average of $-\partial_V \Phi$ in O
dU the variation of U
then there is a function $S(O)$ s.t.

$$dS = \frac{dU + pdV}{K} \equiv \frac{dQ}{T}$$

and $S(O)$ is just the the action of O

$$S(O) = \log \oint_O p dq = \log 2 \int_{q_-}^{q_+} \sqrt{2(U - \Phi(q))} dq$$

$4y$ later Clausius presented, [6], the very same idea and in the priority discussion, [7, 8], reproached B that he had not included the possibility that work, *i.e.* pdV was $\neq 0$: B replied that his argument would not have changed ...

However it can be said that, by that time, B had further developed the idea that the state of a system was to be identified with the average values of the observables and starting from the earlier work of Maxwell was computing averages via a probability dist. on phase space., [9, 10],

He concluded (rederiving Maxwell's Gaussian for an infinite gas) that the distribution to be used was, in modern language, the *microcanonical*, (1868), [9, p.96].

This relied on the Liouville theorem (that B did not know as such and devoted large parts of his papers to derive it every time it was needed): it seems that he was implicitly assuming the distribution as given by a density and that the only possibility was a density function of the energy.

The conclusion (on which later was not entirely confident, [11]) was also supported by the count of occupation of cells in phase space, [9, Sec.II], and by an example of what we would call an "ergodic system", [12]

This was a form of the ergodic hypothesis: in later papers (1871) the hypothesis returns *à propos* the distribution of the atoms in a molecule where the atoms, referred to the center of mass visit all possible points because of the collisions with the other molecules, and at the end he jumps to think of the whole gas as a giant molecule obtaining the microcanonical ensemble, [13, 14, 11, 15].

The remarkable work of 1868 seems to have been recognized first by Maxwell in one of his last papers [16, p.734], as:
"*The only assumption which is necessary for the direct proof [of the microcanonical distribution by Boltzmann] is that the system, if left to itself in its actual state of motion, will, sooner or later, pass through every phase which is consistent with the equation of energy. Now it is manifest that there are cases in which this does not take place*

...

But if we suppose that the material particles, or some of them, occasionally encounter a fixed obstacle such as the sides of a vessel containing the particles, then, except for special forms of the surface of this obstacle, each encounter will introduce a disturbance into the motion of the system, so that it will pass from one undisturbed path into another..."

It might take a long time to do the travel but eventually it will be repeated.

Maxwell interprets Boltzmann without invoking external random actions, and if interpreted in this way this is essentially the Ergodic Hypothesis in its classical form.

Often B. stresses a discrete world view, [17, 18]:
Therefore if we wish to get a picture of the continuum in words, we first have to imagine a large, but finite number of particles with certain properties and investigate the behavior of the ensemble of such particles. Certain properties of the ensemble may approach a definite limit as we allow the number of particles ever more to increase and their size ever more to decrease. Of these properties one can then assert that they apply to a continuum, and in my opinion this is the only non-contradictory definition of a continuum with certain properties.

In 1884 B. writes a paper where the canoncial, microcanonical and other distributions are considered, summarizing his earlier works and showing their equivalence.[19]

At this point it is manifest that the representation of the averages of the observables is unique only if observables are restricted to depend only on the microscopic configurations realized in a "local" region. Furthermore the probabilistic description is well established only for the equilibrium states.

After a century of discussions on the ergodic hypothesis attention eventually shifted to non equilibrium: the major novelty is that in such system dissipation cannot be neglected.

The first works on non equilibrium were based on regarding the nonequilibrium steady states as perturbations of equilibrium states and led to the fundamental reciprocity of Onsager and to Green-Kubo teory for the transport coefficients.[20, 21]

In the 70's the new "law" has been that the stationary states in equilibrium as well as out of equilibrium arise from initial states which are constructed via "protocols" which generate the initial states via a probability distribution μ which
1) is unknown (because the initial data are generated by macroscopic apparata subject to unconrollable influences)
2) it has a density on phase space
furthermore
motions are chaotic, at least microscopically .

Just as in the case of the simple mechanical motion (like celestial evolutions) which were studied using as a paradigm the harmonic motions, it has been necessary to find a corresponding paradigm for the chaotic motions. There exist simple cases of chaotic motions well known

a) geodesic flow on surfaces of constant negative curvature
b) random number generators (*e.g.* unstable maps of circle)
c) axiom A systems (*i.e.* hyperbolic systems)
d) Anosov maps (*i.e.* smooth hyperbolic maps)
e) Markov chains (or more generally Gibbs processes)

actually (b–e) are conceptually equivalent being isomorphic as dynamical systems, up to 0-volume sets.

Ruelle proposed ('70s), that Axiom A systems are a good paradigm,[22, 23].

At the same time massive simulations became possible ('80s) and empirically Ruelle's view was implicitly used: which is not easily admitted as this shows [24, 25]:

... has discussed the possibility that the useful properties exhibited by certain oversimplified and quite rare dynamical systems, termed "Anosov systems", have counterparts in the more usual thermostatted systems studied with nonequilibrium simulation methods. Anosov systems are oversimplifications, like square clouds or spherical chickens...

and more: "*Theoretical constructs such as "measures", should be viewed with a healthy suspicion until algorithms for evaluating them are supplied. The chaos inherent in* *interesting* differential equations guarantees that our only access to the "strange sets" which constitute attractors and repellers will be representative time series from dynamical simulations. In no way can we construct, or even conceive of constructing, a Sinai-Ruelle-Bowen measure for an interesting system."

Disregarding such comments, Ruelle proposal opened the way to the theory of what continued being empirically studied via computers.
Theoretically the proposal of Ruelle was important because

1) unified the approach to equilibrium and non equilibrium replacing the ergodic hypothesis with the assumption that the system is Axiom A (recall that the ergodic hyp. arose from the assumption that motions were periodic!)

2) the key assumption on the protocols to generate initial data leads, for axiom A evolutions, to a unique stationary distribution (without having to invoke Liouville's theorem)

3) reduces the statistical properties of systems with chaotic motions to very well understood stochastic processes in dimension 1 (short range Gibbs).

4) the role of the microscopic time reversal symmetry is naturally reintroduced in nonequilibrium statistics and can lead to the identification of universal properties of various fluctuations

The next question is now "is there an extension of the theory of the equivalent ensembles" also to the non equilibrium cases? Or is it possible that different equations generate statistics which attribute the same averages to the "important" observables?
For instance in equilibrium the microcanonical distribution

$$\mu_{u,\rho}(dpdq) = Z^{-1}\delta(K+V-Nu)dpdq, \qquad u = \frac{K+V}{N},\ \rho = \frac{N}{V}$$

is invariant for the Hamiltonian dynamics, while the isokinetic distribution

$$\mu_{\beta,\rho} = \widetilde{Z}^{-1}\delta(K - \frac{3}{2}\beta_0^{-1}N)e^{-\beta V(q)}dpdq$$

is invariant for the evolution

$$\begin{aligned} \dot{q} &= p \\ \dot{p} &= -\partial_q V(q) - \alpha\, p \end{aligned}\ , \qquad \alpha = -\frac{p\cdot\partial_q V(q)}{\sum p^2}$$

if $\beta_0 = \beta\frac{3N}{3N+1} = \beta + O(\frac{1}{N})$.

The two evolutions are very different, yet the local observables have exactly the same averages in the thermodynamic limit.

The time reversal, *i.e.* the map $I(p,q) = (-p,q)$, is a symmetry for the two evolutions: they are time reversible.

In the next hour the extension of the theory of ensembles to the nonequilibrium stationary states will be discussed resticting the analysis to simple fluid motions and the question about the possibility (or impossibility) of defining an entropy function for steady nonequilibrium states with properties reminiscent of the equilibrium entropy.

This will bring up the role of viscosity and the question about compatibility between microscopic reversibility and macroscopic irreversibility.

As a preparation it is useful to recall the definition of thermodynamic limit.

Consider N particles in a volume $V \subset R^3$ cubic interacting via a pair potential $\Phi(q-q')$ with $F(r) > \frac{1}{r^{d+\varepsilon}}$ for $r < r_0$ and $|\Phi(r)| < c\frac{1}{r^{d+\varepsilon}}$, $\varepsilon > 0$. The Hamiltonian is

$$H = \sum_i \frac{1}{2} p_i^2 + \sum_{i<j} \Phi(q_1 - q_j)$$

Given $\boldsymbol{\beta > 0}, \boldsymbol{\rho = \frac{N}{V} > 0}$, let $Z = \int_{V^N \times R^{3N}} e^{-\beta H(p,q)} dq dp$ and define the probability distribution

$$\mu^V_{\beta,\rho}(dpdq) = Z^{-1} e^{-\beta H(p,q)} dq dp$$

Let $O_{\Lambda_0}(p,q)$ an observable localized in $\Lambda_0 \ll V$ observable, $i.e.$ depending only on the q, p of particles with $q_i \in \Lambda_0$. The therm. limit is def.

$$\lim_{V\to\infty} \mu^N_{\beta,\rho}(O_{\Lambda_0}) = \langle\, O \,\rangle_\beta$$

Likewise the microcanonical therm. limit average is obtained by replacing $\mu^V_{\beta,\rho}(dpdq)$ with

$$\widetilde{\mu}^V_{U,\rho}(dpdq) = \widetilde{Z}^{-1}\delta(H(p,q) - U)dqdp$$

and the limit $V \to \infty$ is considered with fixed $u = \frac{U}{N}$. On general grounds it is proved that the limits exist simultaneously for all local observables in suitable domains for the parameters.

Even when they do not exist and it becomes necessary to select convergent subsequences it can be proved (in many similar models), or it is believed, that still it is possible to establish a $1 \longleftrightarrow 1$ correspondence between $(\beta, \rho) \longleftrightarrow (u, \rho)$ and the subsequences choices to maintain equivalence, *i.e.* identity of the averages.

Quoted references

[1] S. Carnot.
Réflections sur la puissance motrice du feu et sur les machines propres à développer cette puissance.
Online in https://gallica.bnf.fr; original Bachelier, 1824; reprinted Gabay, 1990., Paris, 1824.

[2] R. Clausius.
Ueber eine veränderte form des zweiten hauptsatzes der mechanischen wärmetheorie.
Annalen der Physik und Chemie, 93:481–506, 1854.

[3] R. Clausius.
On the application of the theorem of the equivalence of transformations to interior work.
Philosophical Magazine, 4-XXIV:81–201, 1862.

[4] R. Clausius.
Über einige für Anwendung bequeme formen der Hauptgleichungen der mechanischen Wärmetheorie.
Annalen der Physik und Chemie, 125:353–401, 1865.

[5] L. Boltzmann.
Über die mechanische Bedeutung des zweiten Hauptsatzes der Wärmetheorie.
Wiener Berichte, 53, (W.A.,#2):195–220, (9–33), 1866.

[6] R. Clausius.
Ueber die Zurückführung des zweites Hauptsatzes der mechanischen Wärmetheorie und allgemeine mechanische Prinzipien.
Annalen der Physik, 142:433–461, 1871.

[7] L. Boltzmann.
Zur priorität der auffindung der beziehung zwischen dem zweiten hauptsatze der mechanischen wärmetheorie und dem prinzip der kleinsten wirkung.
Poggendorf Annalen, 143, (W.A.,#17):211–230, (228–236), 1871.

[8] R. Clausius.
Bemerkungen zu der priorität reclamation des hrn. boltzmann.
Annalen der Physik, 144:265–280, 1872.

[9] L. Boltzmann.
Studien über das gleichgewicht der lebendigen kraft zwischen bewegten materiellen punkten.
Wiener Berichte, 58, (W.A.,#5):517–560, (49–96), 1868.

[10] J.C. Maxwell.
On the dynamical theory of gases.
Philosophical Magazine, XXXV:129–145, 185–217, 1868.

[11] L. Boltzmann.
Analytischer Beweis des zweiten Hauptsatzes der mechanischen Wärmetheorie aus den Sätzen über das Gleichgewicht des lebendigen Kraft.
Wiener Berichte, 63, (W.A.,#20):712–732,(288–308), 1871.

[12] L. Boltzmann.
Lösung eines mechanischen problems.
Wiener Berichte, 58, (W.A.,#6):1035–1044, (97–105), 1868.

[13] L. Boltzmann.
Über das Wärmegleichgewicht zwischen mehratomigen Gasmolekülen.
Wiener Berichte, 68, (W.A.,#18:397–418, (237–258), 1871.

[14] L. Boltzmann.
Einige allgemeine sätze über Wärmegleichgewicht.
Wiener Berichte, 63, (W.A.,#19):679–711, (259–287), 1871.

[15] L. Boltzmann.
Zusammenhang zwischen den Sätzen über das Verhalten mehratomiger Gasmoleküle mit Jacobi's Prinzip des letzten Multiplicators.
Wiener Berichte, 63, (W.A.,#19):679–711, (259–287), 1871.

[16] J. C. Maxwell.
On Boltzmann's theorem on the average distribution of energy in a system of material points.
Transactions of the Cambridge Philosophical Society, 12:547–575, 1879.

[17] L. Boltzmann.

Über die Beziehung zwischen dem zweiten Hauptsatze der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung, respektive den Sätzen über das Wärmegleichgewicht.
Wiener Berichte, 76, (W.A.,#42):373–435, (164–223), 1877.

[18] L. Boltzmann.
Theoretical Physics and philosophical writings, ed. B. Mc Guinness.
Reidel, Dordrecht, 1974.

[19] L. Boltzmann.
Über die Eigenshaften monozyklischer und anderer damit verwandter Systeme.
Crelles Journal, 98, (W.A.,#73):68–94, (122–152), 1884.

[20] L. Onsager and S. Machlup.
Fluctuations and irreversible processes.
Physical Review, 91:1505–1512, 1953.

[21] L. Onsager and S. Machlup.
Fluctuations and irreversible processes.
Physical Review, 91:1512–1515, 1953.

[22] D. Ruelle.
Chaotic motions and strange attractors.
Accademia Nazionale dei Lincei, Cambridge University Press, Cambridge, 1989.

[23] D. Ruelle.
Turbulence, strange attractors and chaos.
World Scientific, New-York, 1995.

[24] W. Hoover.
Time reversibility Computer simulation, and Chaos.
World Scientific, Singapore, 1999.

[25] W. Hoover and C. Griswold.
Time reversibility Computer simulation, and Chaos.
Advances in Non Linear Dynamics, vol. 13, 2d edition. World Scientific, Singapore, 1999.

Also: http://arxiv.org & http://ipparco.roma1.infn.it