Boltzmann: ensembles and coarse graining in and out of equilibrium.

Giovanni Gallavotti
Fisica and I.N.F.N. Roma 1
13 October 2006

At the beginning of Boltzmann's work (1866) kinetic theory was quite clearly founded.

Pioneer works of Bernoulli, Herapath, Waterstone; but
Avogadro's law,
First law of Thermodynamics (Meyer, Joule)
Average kinetic energy $\equiv$ absolute temperature (Krönig) and
free path (Clausius) were established.

Challenge: found the second law on Mechanics.

Ergodic hypothesis tool to insure existence of the average values

At first "naive": .. *this explanation is nothing else but the mathematical formulation of the theorem according to which the paths that do not close themselves in any finite time can be regarded as closed in an infinite time*, 1866.

Later: about atoms in a molecule
it is clear that the various gas molecules will go through all possible states of motion, (1871),
possibly different from the ergodic hypothesis (BRUSH, FLAMM)

Considering a collection of copies of the system alike to a large molecule ⇒ ergodicity property of the entire gas, (1871).

Ergodic hyp. could not possibly say that every point visits all points in a *recurrence time*: attributed to BOLTZMANN and criticized over and over (by Physicists; enlightened Math.s saw better).

However for BOLTZMANN phase space was discrete:

Die Zahl der lebendigen Kräfte ist eine diskrete, (1877)

points ⇒ finite size *cells* Δ, $h > 0$.

Thus time evolution is a permutation of cells: ergodicity meant that the permutation was a *one cycle permutation.*

On this assumption BOLTZMANN derived, in a series of 3 papers in (1871) the canonical and microcanonical distributions and the related second law,

The discrete conception ⇒ "combinatorial" derivation of canonical distribution, (1877a).

In (1877b) example of "mechanical analogy" of Thermodynamics: a concept *later* explicitly introduced by HELMHOLTZ, 1884, for systems with only periodic motions (*monocyclic*); no mention of B. 1877-example.

"mechanical analogy" $\Rightarrow$ existence of quantities (typically defined as time averages), which could be given thermodynamic names like

equilibrium state	$\rightarrow$	probabilty distribution for time averages
pressure	$\rightarrow$	p
volume	$\rightarrow$	V
energy	$\rightarrow$	U
temperature	$\rightarrow$	T
entropy	$\rightarrow$	S

and be related by

$$dS = \frac{dU + p\,dV}{T}$$

New perspective and guiding idea:
thermodynamic relations hold in every mechanical system, from the small and simple to the large and complex

trivial identities, just *thermodynamic analogies* for small sys.
nontrivial and interesting relations of general validity for large: *Thermodynamic* $\sim$ *symmetry property of Mechanics*

In 1-dim energy U, confined by potential $\varphi_V(\kappa)$

$$\begin{aligned} T(U,V) =& \tfrac{2}{\tau_0} \int_{\kappa_-(U,V)}^{\kappa_+(U,V)} \tfrac{m}{2} \sqrt{\tfrac{2}{m}(U-\varphi_V(\kappa))} d\kappa, \\ p(U,V) =& \tfrac{2}{\tau_0} \int_{\kappa_-(U,V)}^{\kappa_+(U,V)} \frac{\partial_V \varphi_V(\kappa)}{\sqrt{\frac{m}{2}(U-\varphi_V(\kappa))}} d\kappa \\ S(U,V) =& 2 \log \oint \pi d\kappa \end{aligned}$$

B. extends thermodynamic analogy: families of thermodynamic equilibrium states $\equiv$ families $\mathcal{E}$ of *invariant distributions.*
Proves existence (microcanonical, "*ergode*", canonical "*Holode*") and second law and equivalence (derivation without erg. hyp.)

Given striking success of B: $\Rightarrow$ attempt to extend method to stationary nonequil.

Equations of motion will be symbolically written as

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) \quad \rightarrow \quad \sigma(\mathbf{x}) = -\sum_i \partial_{x_i} f_i(\mathbf{x}) \neq 0$$

Ensembles? states $\equiv$ *statistics* μ of observables namely

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t F(S_\tau \mathbf{x})\, d\tau = \int F(\mathbf{y})\, \mu(\mathbf{y})$$

Follow B's method: derive consequences by conceiving phase space as discrete, time evolution as a permutation of its "discrete" volume elements.

Key obstacle: time average σ_+ of σ will be a *positive quantity*, (RUELLE) $\Rightarrow$ statistics μ on a *attracting set of* 0*-volume.*

Phase space discretized and $\sigma_+ > 0 \Rightarrow$ faithful representation of motion cannot be permutation of discrete volume elements.

Nevertheless simulations exist: analysis of their meaning can help in understanding difference between coarse grain cells and discrete volume elements and even provide a rigorous definition of both as distinct entities at least in the case of strongly chaotic systems.

Simulation $\Rightarrow$ microscopic configurations of the N particles are small "parallelepipeds" defined by the *machine precision*: their number is huge. Evolution is by construction a map of such "microscopic cells" into themselves.

BUT map is far from a permutation: unavoidably merges distinct microscopic cells (violating uniquess)': contraction unavoidable.

Contraction cannot go on forever and any map is *eventually* a permutation of a subset $\mathcal{A}$ of the microscopic cells.

Difference btwn equilibrium and nonequilibrium will then be:

Difference btwn equilibrium and nonequilibrium will then be:

in equilibrium microscopic cells out of $\mathcal{A}$ neglected and motion is permutation;

but in nonequilibrium a substantial number of microscopic cells become irrelevant after a transient time.

*Motion is still a permutation of microscopic cells but only cells which lie on the attracting set have statistical relevance.*c

In both cases an *ergodic hypothesis* can be formulated: the permutation of the microscopic cells that are on $\mathcal{A}$ is *a one cycle permutation*: every microscopic cell visits any other $\Rightarrow$ *chaotic hypothesis*, CH:

asymptotic motions of a confined chaotic mechanical system develop on an attracting set on which motion can be considered mixing and hyperbolic.

hyperbolic transitive $\longleftrightarrow$ *paradigm of chaos* $\Rightarrow$ definition of coarse graining.

$\exists$ partition of Σ into $P_1, P_2, \ldots, P_n$. Symbolic dynamics histories $\boldsymbol{\sigma} = \{\sigma_i\}_{i=-\infty}^{\infty}$ on the sets P_σ, $\sigma = 1, \ldots, n$ has *Markov property*:

(1) there is $M_{\sigma,\sigma'}$ with entries 0 or 1 and "transitive" (*i.e.* $M^m_{\sigma,\sigma'} > 0$), such that if $M_{\sigma_i,\sigma_{i+1}} \equiv 1$ for all i then $\exists! \, \xi \in \Sigma$ such that $S^i\xi \in P_{\sigma_i}$ "coded" by $\boldsymbol{\sigma}$
(2) *compatible* sequences $\boldsymbol{\sigma}$, *i.e.* with $M_{\sigma_i,\sigma_{i+1}} \equiv 1$, then for all points $\xi \in \Sigma$ there is at least one compatible sequence $\boldsymbol{\sigma}$ which codes ξ and for all but a set of zero volume $\boldsymbol{\sigma}$ is unique (*i.e.* just as in binary representation of numbers).

$P_1, P_2, \ldots, P_n$ is a *Markov partition*: statistics of random data ξ will be an S–invariant prob. dist. on compatible sequences $\boldsymbol{\sigma}$.

$P_1, P_2, \ldots, P_n$ generate *coarse grained* cells: given a precision $h > 0$, let N_h be s.t.

$$\Delta = P_{\sigma_{-N_h},\ldots,\sigma_{N_h}} = \bigcap_{j=-N_h}^{N_h} S^{-j} P_{\sigma_j}$$

with diameter $< h$: $\Delta = P_{\sigma_{-N_h},\ldots,\sigma_{N_h}}$ are *coarse cells*. Increasing N_h increases the grain precision: limit $h \to 0$ is not necessary.

S stretches Δ along unstable manifolds of Σ and compresses along stable: maps $\Delta = P_{\sigma_{-N_h},\ldots,\sigma_{N_h}}$ inside the $\cup_\sigma P_{\sigma_{-N_h+1},\ldots,\sigma_{N_h},\sigma}$.

Imagine Δ filled by *far smaller* boxes, *microscopic cells*, of equal volume. Action of S transforms them into boxes contained in only 1 of the n sets $\Delta_\sigma = P_{\sigma_{-N_h+1},\ldots,\sigma_{N_h},\sigma}$.

Such microcells, in a simulation *identified with the integers defining the system points coordinates and momenta in the computer memory*, are thought of as arranged in layers adjacent to unstable manifolds of S and are mapped into microcells of the corresponding layers in the n cells Δ_σ.

Under action of S *layers will merge*; number of microcells will initially decrease; but eventually in each coarse cell Δ will "survive" a collection of layers of microcells. And the microcells will be mapped one to one into themselves: a representation of the attractor, within the precision h.

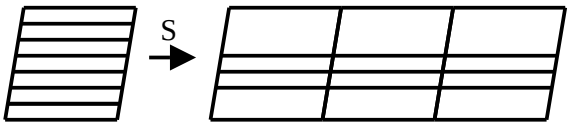

Fig.1: The lines symbolize arrays of microcells Δ: the ones in the left drawing are stretched and merged by the time evolution into arrays that end up exactly fitting into a few new boxes (three in the drawing), a consequence of (2) above.

Consistency $\Rightarrow$ eventually number of microcells in each coarse cell Δ inversely prop. to expansion $\Lambda_e(\Delta)^{-1}$ of surface elements on unstable manifold in Δ: denoted $\mathcal{N}_h\Lambda_e(\Delta)^{-1}$. Transitivity $\Rightarrow$ motion periodic and the statistics μ will be (*SRB dist.*)

$$\pi(\Delta) = \frac{\Lambda_e^{-1}(\Delta)}{\sum_{\Delta'} \Lambda_e^{-1}(\Delta')}$$

Evolution S is *faithfully* a permutation of $\mathcal{N}_h$ microcells on attractor.

The CH interpreted in the same spirit of the viewpoint by RUELLE, 1973, "*... while one would be very happy to prove ergodicity because it would justify the use of Gibbs' microcanonical ensemble, real systems perhaps are not ergodic but behave nevertheless in much the same way and are well described by Gibbs' ensemble...*".

Identification between SRB distributions and distributions giving equal probability to the microcells in the attractor suggests entropy for a stationary state (equilibrium or not) could be the logarithm of the number of microcells on the attractor.

Arbitarily assigning a probability $p(c)$ with each microcell c the distrib. will evolve towards the SRB distribution (if the chaotic hypothesis is accepted) so that $-\sum_c p(c) \log p(c)$ will tend to a maximum $\log \mathcal{N}_h$ (no matter the value of h, if small enough): *i.e.* one-cycle property of microcells evolution $\Rightarrow$ any distribution evolves towards giving equal weight to each microcell, *i.e.* SRB.

This is a kind of H-theorem, *i.e.* a Lyapunov function for the evolution towards the stationary state exists.

However $\mathcal{N}_h$ can be estimated, (2001): its logarithm depends on h. *Remarkably* as an additive constant solely depending on h *only in the case of equilibrium states*, while for nonequilibrium states it depends nontrivially also on other parameters. This seems to indicate impossibility of defining entropy of a stationary state unambiguously, *although a kind of H-theorem could still hold.*

It seems wiser to look for a physical interpretation of the divergence $\sigma(\mathbf{x})$ aiming at experimentally testable predictions.
In the notes I enter into some details about this matter.

Conclusion

BOLTZMANN's contribution to the theory of ensembles and to the mechanical interpretation of Heat and Thermodynamics was based on a discrete conception of the continuum: his staunch coherence on this view has been an essential aspect of the originality of his thought. It hands us a legacy that is often obscured by his other major contribution represented by the B-equation and its very acute interpretation (see for instance his reply to BURBURY).

It is a method still very fruitful and used when "cut–offs" or "regularizations" are employed in the most diverse fields. It is, I believe, still important in the recent theories of nonequilibrium stationary states. The Fluctuation theorem and its various interpretations, extensions and applications (to Onsager reciprocity at non zero forcing, to Green-Kubo formulae, to Fluid Mechanics, Turbulence and Intermittency is, hopefully, only an example.

On discrete conceptions BOLZMANN wrote

"*Through the symbols manipulations of integral calculus, which have become common practice, one can temporarily forget the need to start from a finite number of elements, that is at the basis of the creation of the concept, but one cannot avoid it*", see p. 227 in ([Bo74]);

"*Differential equations require, just as atomism does, an initial idea of a large finite number of numerical values and points ..... Only afterwards it is maintained that the picture never represents phenomena exactly but merely approximates them more and more the greater the number of these points and the smaller the distance between them. Yet here again it seems to me that so far we cannot exclude the possibility that for a certain very large number of points the picture will best represent phenomena and that for greater numbers it will become again less accurate, so that atoms do exist in large but finite number*, see p. 227 in ([Bo74]);

"*This naturally does not exclude that, after we got used once and for all to the abstraction of the volume elements and of the other symbols* [of calculus] *and once one has studied the way to operate with them, it could look handy and luring, in deriving certain formulae that Volkmann calls formulae for the coarse phenomena, to forget completely the atomistic significance of such abstractions. They provide a general model for all cases in which one can think to deal with* 10^{10} *or* $10^{10^{10}}$ *elements in a cubic millimeter or even with billions of times more; hence they are particularly invaluable in the frame of Geometry, which must equally well adapt to deal with the most diverse physical cases in which the number of the elements can be widely different. Often in the use of all such models, created in this way, it is necessary to put aside the basic concept, from which they have overgrown, and perhaps to forget it entirely, at least temporarily. But I think that it would be a mistake to think that one could become free of it entirely.*", see p. 55 in ([Bo74])
`http://ipparco.roma1.infn.it`